



SIMULAÇÃO ESCALÁVEL DE EPIDEMIAS EM REDES BASEADAS EM
PASSEIOS ALEATÓRIOS COM CARACTERIZAÇÃO DE TRANSIÇÕES DE
FASE

João Vitor Barbosa Tavares

Dissertação de Mestrado apresentada ao Programa de Pós-graduação em Engenharia de Sistemas e Computação, COPPE, da Universidade Federal do Rio de Janeiro, como parte dos requisitos necessários à obtenção do título de Mestre em Engenharia de Sistemas e Computação.

Orientadores: Daniel Ratton Figueiredo
Giulio Iacobelli

Rio de Janeiro
Abril de 2018

SIMULAÇÃO ESCALÁVEL DE EPIDEMIAS EM REDES BASEADAS EM
PASSEIOS ALEATÓRIOS COM CARACTERIZAÇÃO DE TRANSIÇÕES DE
FASE

João Vitor Barbosa Tavares

DISSERTAÇÃO SUBMETIDA AO CORPO DOCENTE DO INSTITUTO
ALBERTO LUIZ COIMBRA DE PÓS-GRADUAÇÃO E PESQUISA DE
ENGENHARIA (COPPE) DA UNIVERSIDADE FEDERAL DO RIO DE
JANEIRO COMO PARTE DOS REQUISITOS NECESSÁRIOS PARA A
OBTENÇÃO DO GRAU DE MESTRE EM CIÊNCIAS EM ENGENHARIA DE
SISTEMAS E COMPUTAÇÃO.

Examinada por:

Prof. Daniel Ratton Figueiredo, Ph.D.

Prof. Giulio Iacobelli, Ph.D.

Prof. Valmir Carneiro Barbosa, Ph.D.

Prof. Daniel Sadoc Menasche, Ph.D.

RIO DE JANEIRO, RJ – BRASIL

ABRIL DE 2018

Tavares, João Vitor Barbosa

Simulação Escalável de Epidemias em redes Baseadas em Passeios Aleatórios com Caracterização de Transições de Fase/João Vitor Barbosa Tavares. – Rio de Janeiro: UFRJ/COPPE, 2018.

XIII, 51 p.: il.; 29, 7cm.

Orientadores: Daniel Ratton Figueiredo

Giulio Iacobelli

Dissertação (mestrado) – UFRJ/COPPE/Programa de Engenharia de Sistemas e Computação, 2018.

Referências Bibliográficas: p. 49 – 51.

1. Epidemia. 2. Redes. 3. Transição de fase. I. Figueiredo, Daniel Ratton *et al.* II. Universidade Federal do Rio de Janeiro, COPPE, Programa de Engenharia de Sistemas e Computação. III. Título.

À minha querida esposa.

Agradecimentos

Primeiramente, agradeço ao meu Deus pela sua graça e infinitas bênçãos em minha vida.

Agradeço à minha família e sobretudo meus pais por todo apoio e incentivo que sempre me deram. Obrigado por me amarem e cuidarem de mim.

Principalmente, à minha amada esposa, por sempre acreditar em mim, me amar e estar comigo em todos os momentos desta dissertação e da vida. Sem dúvidas você é a melhor esposa e companheira do mundo, amo você.

Agradeço ao meus orientadores, Prof. Daniel Ratton Figueiredo e Prof. Giulio Iacobelli por compartilharem seus conhecimentos comigo.

Agradeço à COPPE e seus professores por contribuírem para meu aprendizado.

Resumo da Dissertação apresentada à COPPE/UFRJ como parte dos requisitos necessários para a obtenção do grau de Mestre em Ciências (M.Sc.)

SIMULAÇÃO ESCALÁVEL DE EPIDEMIAS EM REDES BASEADAS EM
PASSEIOS ALEATÓRIOS COM CARACTERIZAÇÃO DE TRANSIÇÕES DE
FASE

João Vitor Barbosa Tavares

Abril/2018

Orientadores: Daniel Ratton Figueiredo
Giulio Iacobelli

Programa: Engenharia de Sistemas e Computação

Entender como redes e processos dinâmicos se relacionam é um tema central de pesquisa nos dias de hoje, e em particular no contexto de epidemias que se desdobram sobre redes. Neste trabalho, consideramos o cenário onde indivíduos se movimentam por uma rede, que representa a estrutura do espaço de movimentação. O contágio pode ocorrer quando dois (ou mais) indivíduos se encontram em um mesmo local (vértice). O objetivo deste trabalho é projetar e implementar um simulador de eventos discretos eficiente para este tipo de epidemia e caracterizar o comportamento da epidemia em função da estrutura da rede e parâmetros do modelo. Em particular, a avaliação teórica e empírica indicam que o simulador é escalável no tamanho da rede, tempo de simulação e número de indivíduos. Além disso, resultados obtidos sobre o comportamento de epidemias indicam uma transição de fase em diversos parâmetros do modelo, de forma que epidemias ou terminam rapidamente ou perduram por muito tempo.

Abstract of Dissertation presented to COPPE/UFRJ as a partial fulfillment of the requirements for the degree of Master of Science (M.Sc.)

SCALABLE SIMULATIONS OF EPIDEMICS ON NETWORKS BASED ON
RANDOM WALKS AND A PHASE TRANSITIONS CHARACTERIZATION

João Vitor Barbosa Tavares

April/2018

Advisors: Daniel Ratton Figueiredo
Giulio Iacobelli

Department: Systems Engineering and Computer Science

Understanding how networks and dynamic processes relate is an important topic of research, and particularly in the context of epidemics spreading through networks. In this work, we consider the scenario where individuals move through a network. Contagion occurs when two (or more) individuals are in the same location (vertex). The aim of this work is to build a scalable simulator for this type of epidemic and characterize the behavior of the epidemic in function of the network structure. In particular, our results indicate that the simulator is scalable in network size, simulation time and number of individuals. In addition, results on the epidemic indicate a phase transition in several parameters of the model, so that epidemics terminate quickly or last for a long time.

Sumário

Lista de Figuras	x
Lista de Tabelas	xiii
1 Introdução	1
1.1 Contribuição	2
1.2 Organização	3
2 Epidemia em redes	5
2.1 Modelos epidêmicos	5
2.1.1 Modelo SI	7
2.1.2 Modelo SIS	8
2.2 Modelos epidêmicos em redes	9
2.2.1 Modelo SI em redes	10
2.2.2 Modelo SIS em redes	13
2.2.3 Duração de epidemias SIS e estrutura das redes	14
2.2.4 Epidemia em redes onde indivíduos se movimentam	15
2.3 Simuladores de epidemias	16
2.3.1 GLEAM	16
2.3.2 EpiSimdemics	17
3 Modelo epidêmico em redes com passeios aleatórios	18
3.1 Modelo	18
4 Simulador	22
4.1 Arquitetura	22
4.2 Implementação	24
4.3 Medidas de Interesse	25
4.4 Utilização	27
5 Avaliação Numérica	30
5.1 Eficiência computacional	30

5.2	Caracterização da epidemia no tempo	33
5.3	Transição de fase no tempo de simulação da epidemia	37
5.4	População inicial de infectados	42
5.5	Validação	43
6	Conclusões e Trabalhos Futuros	47
6.1	Conclusão	47
6.2	Trabalhos Futuros	48
	Referências Bibliográficas	49

Lista de Figuras

2.1	Transições entre níveis de contágio de diferentes modelos epidêmicos	6
2.2	Regime de diferentes tipos de epidemias ao longo do tempo	7
2.3	Propagação da peste negra na Europa durante idade média. Fonte: Roger Zener, Wikimedia Commons.	10
2.4	Propagação da epidemia de SARS pelo mundo de novembro de 2002 a abril de 2003. Fonte: WHO, http://www.who.int/csr/sars/map2003_04_28.gif	10
2.5	Fração de vértices infectados para cada classe de grau k no modelo SI. Extraída de [1]	12
3.1	Transição entre os estados epidêmicos	19
3.2	Indivíduo infectado (em vermelho) entra em um vértice ($v \rightarrow u$) com indivíduos suscetíveis.	19
3.3	Indivíduo no estado contraído (em amarelo) torna-se infectado em um vértice (u) com indivíduos suscetíveis.	19
3.4	Indivíduo suscetível (em vermelho) entra em um vértice ($v \rightarrow u$) com indivíduos infectados.	21
3.5	Indivíduo infectado torna-se suscetível em um vértice (u) com indivíduos infectados.	21
4.1	Diagrama de classes do simulador	23
4.2	Fluxo de processamento dos eventos do simulador	25
4.3	Log do simulador para $I(t)$, $C(t)$ e $S(t)$	27
4.4	Log do simulador para o indivíduo 0	27
4.5	Parâmetros do simulador	28
5.1	Tempo execução em função do tamanho da rede, n , para diferentes redes e número de indivíduos ($K = 5$, $K = 50$) em escala log-log. Parâmetros: $T = 10000$, $\lambda = 1$, $\gamma = 1$, $\tau = 1$	31
5.2	Tempo execução em função do tempo de simulação, em escala log-log. Parâmetros: $n = 25$, $\lambda = 1$, $\gamma = 1$, $\tau = 1$, $T = 1000$	32

5.3	Tempo de execução em função dos passeios aleatórios, em escala log-log. Parâmetros: $T = 1000, n = 50, \lambda = 1, \gamma = 1, \tau = 1$	33
5.4	Fração de indivíduos infectados ao longo do tempo para diferentes taxas de caminhar, λ . Parâmetros: $n = 100, K = 150, k_i = 2, \gamma = 1, \tau = 1$	33
5.5	Fração de indivíduos infectados ao longo do tempo para diferentes taxas de infectar, τ . Parâmetros: $n = 100, K = 10000, \lambda = 5, \gamma = 1$. .	34
5.6	Fração de indivíduos em cada estado epidêmico ao longo do tempo para $\lambda = 10, \gamma = 1, \tau = 1, n = 100$ e $K = 10000$	35
5.7	Fração de indivíduos em cada estado epidêmico ao longo do tempo para $\lambda = 10, \gamma = 1, \tau = 0.5, n = 100$ e $K = 10000$	35
5.8	Fração de indivíduos em cada estado epidêmico ao longo do tempo para o grafo completo ($\lambda = 1, \gamma = 1, \tau = 1, n = 100$ e $K = 10000$). .	36
5.9	Fração de indivíduos em cada estado epidêmico ao longo do tempo para o grafo em anel ($\lambda = 1, \gamma = 1, \tau = 1, n = 100$ e $K = 10000$). . .	36
5.10	Fração de indivíduos em cada estado epidêmico ao longo do tempo para o grafo anel ($\lambda = 1, \gamma = 1, \tau = 1, n = 1000$ e $K = 10000$).	36
5.11	Número de indivíduos infectados em função do tempo de simulação, para epidemia de curta duração. Parâmetros: $n = 30, K = 38, \lambda = 1, \gamma = 1, \tau = 1, T = 100000$	37
5.12	Número de indivíduos infectados em função do tempo de simulação, para epidemia de longa duração. Parâmetros: $n = 30, K = 50, \lambda = 1, \gamma = 1, \tau = 1, T = 100000$	37
5.13	Transição de fase para o grafo completo. Parâmetros: $\lambda = 1, \gamma = 1, \tau = 1, T = 100000$	38
5.14	Transição de fase para o anel. Parâmetros: $\lambda = 1, \gamma = 1, \tau = 1, T = 100000$	39
5.15	Transição de fase para o torus no parâmetro K . Parâmetros: $\lambda = 1, \gamma = 1, \tau = 1, T = 100000$	39
5.16	Transição de fase para o grafo completo na relação λ/γ . Parâmetros: $K = 60, \tau = 1, T = 100000$	40
5.17	Transição de fase para o anel na relação λ/γ . Parâmetros: $K = 60, \tau = 1, T = 100000$	41
5.18	Transição de fase para o torus na relação λ/γ . Parâmetros: $K = 60, \tau = 1, T = 100000$	42
5.19	Transição de fase com diferentes valores de k_i para diferentes tamanhos de rede. Parâmetros $p = 0.5, \lambda = 1, \gamma = 1, \tau = 1$	43
5.20	Desvio padrão para diferentes tamanhos de rede, n , da duração da epidemia variando k_i . Parâmetros $p = 0.5, \lambda = 1, \gamma = 1, \tau = 1$	43

- 5.21 Probabilidade empírica e teórica de um indivíduo permanecer um tempo $T \geq t$ em um vértice. $\lambda = 1$. ($n = 25, K = 40, \gamma = 1, \tau = 1$) . . 45
- 5.22 Probabilidade empírica e teórica de um indivíduo permanecer um tempo $T \geq t$ no estado infectado, $\gamma = 1$. ($n = 25, K = 40, \lambda = 1, \tau = 1$) 45
- 5.23 Probabilidade empírica e teórica de um indivíduo permanecer um tempo $T \geq t$ no estado contraído, $\tau = 1$. ($n = 25, K = 40, \lambda = 1, \gamma = 1$) 45

Lista de Tabelas

3.1	Variáveis do modelo	21
5.1	Comparação entre os dados gerados pelo simulador e valores teóricos das exponenciais para diferentes eventos do simulador ($n = 35, K = 40, \lambda = 1, \gamma = 1, \tau = 1, T = 100000$).	44
5.2	Comparação entre os dados gerados pelo simulador e valores teóricos das exponenciais para diferentes eventos do simulador ($n = 35, K = 40, \lambda = 2, \gamma = 0.5, \tau = 3, T = 100000$).	44
5.3	Comparação entre o valor amostral e esperado do parâmetro p ($n = 35, K = 40, \lambda = 1, \gamma = 1, \tau = 1, T = 100000$).	46

Capítulo 1

Introdução

Ao longo de toda sua existência, a humanidade convive com epidemias, como a gripe espanhola que matou milhões de pessoas no início do século XX, e as epidemias de dengue que ocorrem nos países tropicais todos os anos. Mais recentemente, cascatas de informação atingem milhões de pessoas, como vídeos virais no Facebook e mensagens no Twitter. Desta forma, epidemias são tema central de pesquisa em diversas áreas. O primeiro modelo matemático surgiu com Bernoulli em 1760 [2, 3] para explicar a epidemia de varíola da época. Mais tarde, em 1927, Kermack *et al.* [4] propuseram o primeiro modelo genérico para epidemias, utilizando equações diferenciais ordinárias, que foi e vem sendo amplamente aplicado e generalizado. Mais recentemente, modelos de epidemias em redes vem sendo utilizados para melhor capturar a estrutura que governa o processo de contágio [5–7].

Modelos epidêmicos em geral atribuem estados epidêmicos aos indivíduos, sendo os mais comuns: *Suscetível*, *Infectado* e *Recuperado*. O estado suscetível representa os indivíduos que não estão contagiados mas que ainda podem ser contagiados. No estado infectado estão os indivíduos contagiados que são capazes de propagar a epidemia. Já no estado recuperado estão os indivíduos que se recuperaram, que estiveram no estado infectado, mas que agora estão imunes a novos contágios. Alguns modelos acrescentam um estado intermediário entre os estados suscetível e infectado, denotado por *Contraído*, para representar indivíduos que foram contagiados mas ainda não são capazes de transmitir a epidemia.

O modelo clássico de epidemia em redes representa indivíduos por vértices da rede e a possibilidade de contágio entre dois indivíduos por arestas. Neste cenário, a epidemia é um processo sobre a estrutura da rede, no qual vértices podem ser contagiados através de seus vizinhos [8, 9]. Modelos mais recentes consideram a rede como sendo uma estrutura de locais sobre a qual indivíduos se movimentam [10, 11]. Neste cenário, cada vértice da rede representa um local e cada aresta indica a possibilidade de locomoção direta entre dois locais. Indivíduos que se movimentam pela rede possuem um estado epidêmico e o contágio pode ocorrer quando dois (ou

mais) indivíduos se encontram em um mesmo local. Este trabalho estuda este tipo de modelo de epidemia em redes.

Um aspecto importante de um modelo epidêmico refere-se às possíveis transições entre os estados epidêmicos. Um dos modelos mais simples possui apenas dois estados epidêmicos, Suscetível (S) e Infectado (I), e apenas uma possível transição, $S \rightarrow I$. Um outro modelo amplamente estudado é o modelo que adiciona mais uma transição, do estado Infectado para o Suscetível, conhecido por modelo SIS. Neste caso, indivíduos podem passar pelo ciclo SIS inúmeras vezes. Este é o caso do vírus da gripe, por exemplo. Neste trabalho iremos considerar o modelo SCIS, adicionando mais um estado epidêmico (Contraído) e mais duas transições ($S \rightarrow C$ e $C \rightarrow I$). Este modelo generaliza o clássico SIS.

Modelos computacionais epidêmicos têm como objetivo simular epidemias complexas, com muitas variáveis e parâmetros e vem sendo cada vez mais usados para avaliar o comportamento de epidemias sobre diferentes condições [12, 13]. Por exemplo, a caracterização da duração de uma epidemia em função da estrutura da rede sobre a qual se desenvolve e de outros parâmetros da epidemia, como tempo de recuperação e velocidade de locomoção.

Neste trabalho, projetamos e implementamos um simulador de eventos discretos para simular epidemias SCIS em redes com indivíduos se movimentando. A análise teórica e experimental demonstra que o simulador é escalável no tamanho da rede, no tempo de simulação e também no número de indivíduos que participam da epidemia. Apresentamos resultados empíricos do tempo de execução do simulador em diferentes cenários para ilustrar sua escalabilidade. Em particular, mostramos que, para um tempo fixo de simulação T , o tempo de execução não depende do tamanho da rede, que o mesmo cresce linearmente com tempo de simulação, e que o mesmo cresce como $\mathcal{O}(K \log K)$ com o número K de indivíduos na rede.

Além disso, utilizamos o simulador para caracterizar o tempo de duração de uma epidemia. Repare que uma epidemia SIS (ou SCIS) termina quando todos os indivíduos se encontram no estado Suscetível, pois não há mais infecções neste caso. Nossos resultados empíricos indicam que a duração de uma epidemia está sujeita a uma transição de fase em diferentes parâmetros do modelo, tal que epidemias ou terminam rapidamente ou perduram na rede por muito tempo (até o tempo máximo de simulação). Isto indica a sensibilidade do modelo epidêmico aos parâmetros, tais como o número de indivíduos, tempo de recuperação e velocidade de locomoção.

1.1 Contribuição

Os modelos desenvolvidos para estudar epidemias não são recentes, vindo desde Bernoulli em 1760. No entanto, no dias atuais surgiram muitos modelos epidêmicos

em redes para representar processos de propagação, muito por conta do estudo recente de redes reais. Visando contribuir para o desenvolvimento da epidemiologia computacional, as principais contribuições deste trabalho são:

- Propor um estudo de modelo epidêmico para redes onde indivíduos movimentam-se de acordo com passeios aleatórios independentes. Cada indivíduo possui um estado epidêmico que representa o estágio de contágio. Para representar os estados epidêmicos dos indivíduos e suas transições, utilizamos o modelo SCIS, largamente utilizado na literatura, onde o estado Contraído representa indivíduos contagiados que ainda não podem propagar a epidemia.
- Projetar e implementar um simulador de eventos discretos para estudar e caracterizar epidemias SCIS. O simulador desenvolvido possui os seguintes eventos: Caminhar, Infectar e Recuperar. O evento Caminhar é responsável por movimentar um indivíduo do vértice atual para um vértice vizinho, o evento Infectar muda o estado epidêmico do indivíduo de Contraído para Infectado, e o evento Recuperar muda o estado epidêmico do indivíduo de Infectado para Suscetível. Proporcionamos um estudo da escalabilidade do simulador para mostrar que ele é escalável em diversos parâmetros do modelo. Em particular, para um tempo fixo de simulação T , o tempo de execução do simulador não depende do tamanho da rede, é linear com o tempo de simulação T , e cresce como $\mathcal{O}(K \log K)$ com o número K de indivíduos na rede.
- Caracterizar o comportamento da epidemia com um estudo de transição de fase em alguns parâmetros da rede, mais precisamente o número de indivíduos e a relação λ/γ , onde λ é a taxa de caminhar dos indivíduos e γ a taxa com que os indivíduos recuperam-se. Queremos identificar se a duração da epidemia é influenciada por estes parâmetros e se é possível identificar um ponto crítico que determine se a epidemia tem curta ou longa duração. Em particular, nossos experimentos mostram que a epidemia possui transição de fase no número de indivíduos e em λ/γ e que o ponto de transição depende da estrutura da rede.

1.2 Organização

Esta dissertação de mestrado está organizada em 6 capítulos, dos quais este é o primeiro. No capítulo 2, descreve-se com mais detalhes o estudo de epidemias em redes, assim como alguns modelos e resultados. No capítulo 3, detalha-se o modelo epidêmico utilizado nesta tese. No capítulo 4, apresenta-se, com mais detalhes, o simulador desenvolvido para estudar epidemias. No capítulo 5, apresenta-se uma

avaliação do simulador, tanto de sua escalabilidade quanto a caracterização da epidemia. Finalmente, no capítulo 6, apresentamos algumas conclusões e possíveis melhorias para o tema.

Capítulo 2

Epidemia em redes

Epidemias são um tema de pesquisa desde o primeiro modelo epidêmico criado por Bernoulli em 1760 [2]. Epidemias podem ser classificadas como um fenômeno de propagação causado por um agente que opera sobre uma determinada rede ou estrutura. A teoria de epidemias pode ser expandida para englobar qualquer fenômeno de propagação, desde propagação de vírus e patógenos como dengue e Aids até propagação de conhecimento, inovação, rumores através de ferramentas como Twitter, Facebook e Youtube.

A propagação de uma epidemia em redes é determinada por dois fatores: o mecanismo de transmissão de seu agente e a rede sobre a qual a epidemia se desdobra. Agentes epidêmicos se propagam de diversas maneiras, através do ar, contato físico, fluidos contaminados ou contato sexual. Alguns destes mecanismos favorecem a rápida propagação da epidemia enquanto outros espalham a epidemia lentamente. Por exemplo, a epidemia de SARS em 2003 [14] foi favorecida pela facilidade de propagação do vírus pelo ar, de modo que um indivíduo contaminado pudesse infectar todas as pessoas presentes em uma mesma sala. Por outro lado, doenças como Aids propagam-se mais lentamente já que é necessário contato sexual para transmissão.

A estrutura por onde a epidemia se desdobra também possui papel determinante em sua propagação. No caso da epidemia de SARS, seu comportamento foi diretamente influenciado pela presença de *super-spreads* [15] e pela estrutura da rede aérea que possibilitou disseminar o vírus para outros continentes. Em redes, os *super-spreads* são chamados de *hubs*, vértices que possuem grau bem maiores que a média, alcançando muito mais vértices diretamente.

2.1 Modelos epidêmicos

A epidemiologia surgiu como uma área da ciência dedicada a estudar surtos epidêmicos e com ela surgiu o modelo clássico de epidemia que se baseia em duas premissas: o agrupamento dos indivíduos em estados e a probabilidade de um in-

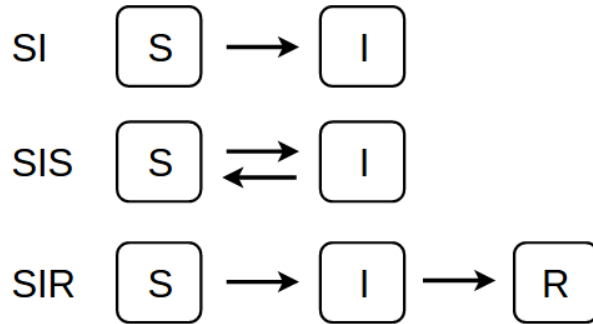


Figura 2.1: Transições entre níveis de contágio de diferentes modelos epidêmicos

divíduo infectado contaminar qualquer outro indivíduo saudável com igual probabilidade (*homogeneous mixing*) [1]. Ou seja, este modelo não leva em consideração a rede de interação entre os indivíduos.

Os estados epidêmicos mais utilizados para representar o nível de contágio dos indivíduos são: *Suscetível*, *Infectado* e *Recuperado*, onde o estado Suscetível (S) agrupa os indivíduos sem infecção, o estado Infectado (I) representam os indivíduos contagiados que podem propagar a epidemia e o estado Recuperado (R) representa indivíduos que se recuperaram e não podem transmitir a epidemia. A quantidade e tipos dos estados são determinados pelo tipo da epidemia. Neste trabalho utilizamos o estado Contraído para classificar indivíduos contagiados mas que ainda não podem transmitir a infecção.

A transição entre estados epidêmicos permite representar epidemias com diferentes comportamentos, conforme ilustrado na figura 2.1. O modelo SIS atribui dois estados epidêmicos aos indivíduos permitindo transições $S \rightarrow I \rightarrow S$. Neste modelo, os indivíduos podem passar inúmeras vezes pelo ciclo SIS. Um outro modelo com os mesmos estados mas com apenas uma transição é o $S \rightarrow I$ (SI). Neste modelo, o indivíduo uma vez infectado, não muda mais de estado. Outro modelo, com mais um estado, é o SIR, com transições $S \rightarrow I \rightarrow R$. Neste modelo, ao se recuperar, o indivíduo não sai mais deste estado.

Epidemias SI e SIS, não conferem imunização aos agentes, diferentemente de epidemias SIR. No modelo SI, ao final, todos os indivíduos estarão no estado infectado. No modelo SIS, os indivíduos são submetidos ao ciclo de transições $S \rightarrow I \rightarrow S$ e a epidemia termina quando todos os indivíduos tornam-se saudáveis ou epidemia pode entrar em um estado estacionário (estado endêmico), com uma quantidade constante de indivíduos infectados. Já a epidemia SIR termina quando não há mais indivíduos infectados, seja porque todos estão saudáveis ou recuperados. A figura 2.2 mostra o regime de cada modelo ao longo do tempo.

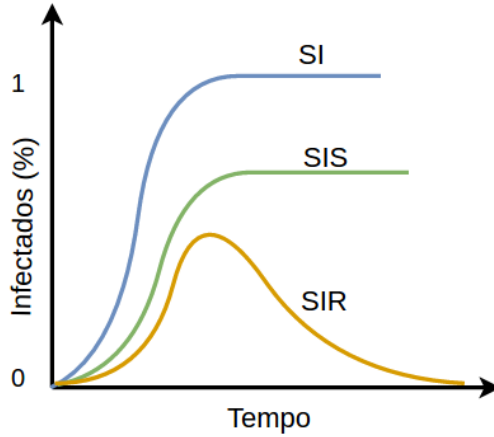


Figura 2.2: Regime de diferentes tipos de epidemias ao longo do tempo

As transições e o tempo no qual os indivíduos permanecem nos estados epidêmicos são importantes para determinar o comportamento e a evolução da epidemia. A transição $S \rightarrow I$ ocorre somente quando um indivíduo saudável encontra um infectado. Por outro lado, as transições $I \rightarrow S$ e $I \rightarrow R$ ocorrem após um intervalo de tempo, sem nenhuma interação entre indivíduos.

Albert-László Barabási apresenta em seu livro [1] um estudo dos modelos epidêmicos clássicos (sem redes) e o modelo epidêmico baseado em redes apresentado por Pastor-Satorras e Alessandro Vespignani em 2001 [9], que adiciona ao modelo clássico a característica topológica da rede de contato dos indivíduos. As próximas subseções apresentam detalhes destes modelos.

2.1.1 Modelo SI

Sejam $S(t)$ e $I(t)$ o número de indivíduos suscetíveis e infectados no tempo t , respectivamente, e N o número total de indivíduos do modelo. Chamemos de β a taxa de transmissão da infecção de um indivíduo contagiado para um indivíduo suscetível e l o número de contatos que um indivíduo possui. Lembrando que para o modelo SI clássico (sem rede) qualquer indivíduo tem probabilidade igual de encontrar qualquer outro e por isso a chance de um indivíduo infectado encontrar outro suscetível é $S(t)/N$. Com l contatos, a chance de um indivíduo infectado encontrar outro suscetível é $lS(t)/N$. Considerando a taxa de transmissão da infecção β , o número médio de novas infecções em um intervalo de tempo é:

$$\frac{di(t)}{dt} = \beta l \frac{S(t)I(t)}{N} = \beta l s(t)i(t), \quad (2.1)$$

onde $s(t) = S(t)/N$ (fração de indivíduo suscetíveis no tempo t) e $i(t) = I(t)/N$ (fração de indivíduos infectados no tempo t). A partir da equação 2.1 (vide [1]) chega-se na fração de indivíduos infectados em função do tempo:

$$i(t) = \frac{i_0 e^{\beta l t}}{1 - i_0 + i_0 e^{\beta l t}}, \quad (2.2)$$

onde i_0 é o número de indivíduos infectados no tempo zero.

No começo da epidemia, existem poucos indivíduos infectados e portanto estes encontram majoritariamente indivíduos suscetíveis fazendo a fração de indivíduos infectados no tempo crescer exponencialmente. Deste modo a epidemia se espalha rapidamente na população. Por outro lado, a medida que a epidemia se desenrola, os encontros entre indivíduos infectados e suscetíveis são mais escassos e por isso $i(t)$ cresce mais devagar até chegar ao regime de saturação, onde todos os indivíduos estão infectados (figura 2.2).

Além disso, outro ponto importante é o tempo característico do modelo:

$$\tau = \frac{1}{\beta l}, \quad (2.3)$$

que é o inverso da taxa de propagação da epidemia na população. Ou seja, quando a probabilidade de infecção aumenta, o tempo até a epidemia chegar em seu regime de saturação diminui. Repare que para um vírus com mecanismo de transmissão rápido (alto valor para β) ou quando indivíduos infectados estão em contato com muitos indivíduos suscetíveis (alto valor para L) a epidemia se propaga mais rapidamente.

2.1.2 Modelo SIS

Em certas doenças os indivíduos são capazes de se recuperar e voltar para o suscetível, tornando-se novamente suscetível ao patógeno, como é o caso do vírus da gripe. Para representar estas doenças adiciona-se ao modelo SI a transição $I \rightarrow S$, gerando o modelo SIS. Neste modelo os indivíduos recuperam-se a uma taxa μ . Para representar o número de indivíduos infectados por intervalo de tempo, estende-se a equação 2.1 para adicionar o termo μ :

$$\frac{di(t)}{dt} = \beta l s(t) i(t) - \mu i(t) \quad (2.4)$$

Repare que $\mu i(t)$ é a taxa em que os indivíduos se recuperam da infecção e quando $\mu i(t)$ é maior que $\beta l s(t) i(t)$, o número de indivíduos infectados por intervalo de tempo é negativo e portanto a epidemia não se espalha.

A solução da equação 2.4 (vide [1]) determina a fração de indivíduos infectados

ao longo do tempo:

$$i(t) = \left(1 - \frac{\mu}{\beta l}\right) \frac{C e^{(\beta l - \mu)t}}{1 + C e^{(\beta l - \mu)t}}, \quad (2.5)$$

onde $C = i_0 / (1 - i_0 - \mu / \beta l)$.

Note que de acordo com a equação 2.5 uma epidemia SIS possui dois destinos possíveis: ou a epidemia termina rapidamente (e $i(t)$ vai a zero) ou ela se espalha e persiste em uma fração constante de indivíduos por um longo período. Quando a taxa de recuperação é grande em relação a taxa de infecção ($\mu > \beta l$) o expoente da equação 2.5 é negativo, ou seja, temos mais indivíduos se recuperando do que indivíduos sendo infectados por unidade de tempo. Neste cenário todos os indivíduos tornam-se suscetíveis rapidamente pois a fração de indivíduos infectados decresce exponencialmente.

Por outro lado, quando $\mu < \beta l$ a epidemia SIS comporta-se semelhantemente a uma epidemia SI. No entanto, a fração de indivíduos infectados cresce exponencialmente até encontrar uma contante $i^* < 1$, permanecendo no estado endêmico (figura 2.2). O tempo característico da infecção no modelo SIS é $\tau = 1 / (\beta l - \mu)$. Reescrevendo de outra forma temos:

$$\tau = \frac{1}{\mu(R_0 - 1)}, \quad (2.6)$$

onde $R_0 = \beta l / \mu$ é comumente chamada de valor de reprodução básico. Este parâmetro indica o número médio de indivíduos suscetíveis que foram contaminados por um indivíduo infectado durante seu período contagioso. Repare que quando $R_0 > 1$ cada indivíduo infectado contamina em média mais de um indivíduo suscetível levando a epidemia ao estado endêmico. Por outro lado quando $R_0 < 1$ a epidemia não persiste na população já que τ é negativo, ou seja, um indivíduo infectado não contamina mais do que um indivíduo suscetível em média.

2.2 Modelos epidêmicos em redes

Os modelos SI, SIS apresentados não levam em consideração a estrutura de contato dos indivíduos. De fato, eles baseiam-se na premissa de que um indivíduo infectado tem probabilidade igual de encontrar qualquer indivíduo suscetível da população. Esta hipótese descarta o fato de que indivíduos estão organizados em redes, ou seja, um indivíduo interage apenas com uma pequena parcela da população. Além disso, outra premissa é de que todos os indivíduos possuem o mesmo número de contatos k , desprezando o fato de que indivíduos possuem número de interações diferentes.

Isso não é suficiente para estudar fielmente epidemias. Tome como exemplo as epidemias que ocorreram durante a idade média. A velocidade de propagação da

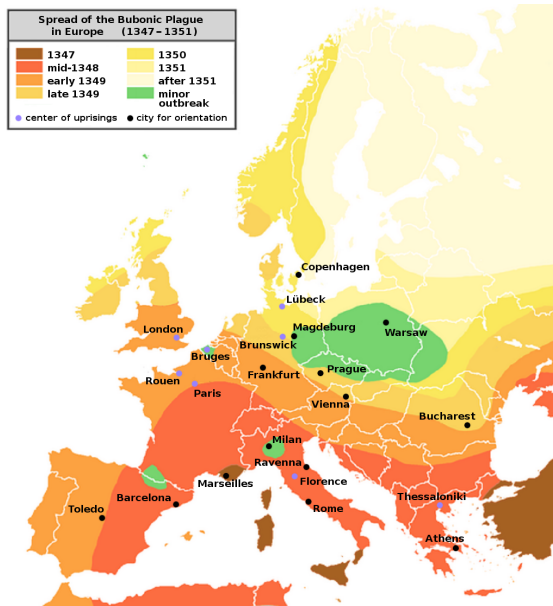


Figura 2.3: Propagação da peste negra na Europa durante idade média. Fonte: Roger Zener, Wikimedia Commons.

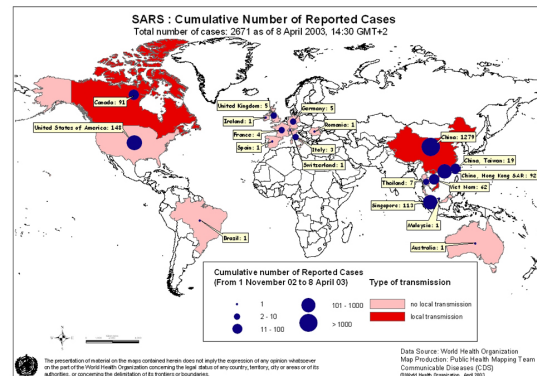


Figura 2.4: Propagação da epidemia de SARS pelo mundo de novembro de 2002 a abril de 2003. Fonte: WHO, http://www.who.int/csr/sars/map2003-04_28.gif.

peste negra era lenta (compatível com os meios de transportes da época) e demorou alguns anos para se espalhar por todo continente europeu, figura 2.3. Hoje em dia, adicionamos conexões entre cidades e continentes que não existiam naquela época, através das viagens de avião. Esta facilidade de transporte impactou diretamente a maneira como um vírus pode se propagar, como por exemplo a epidemia de SARS que em apenas meses já estava presente em vários continentes, figura 2.4.

A criação de modelos epidêmicos utilizando redes permite utilizar a estrutura da rede para representar a estrutura de interação entres os indivíduos, enriquecendo ainda mais o modelo clássico [1, 9]. Em redes reais, indivíduos possuem diferentes conexões, cuja distribuição de grau frequentemente possui cauda pesada. Indivíduos com mais conexões tem mais probabilidade de encontrarem indivíduos infectados e assim serem contaminados. Levando isso em consideração, Romualdo Pastor-Satorras e Alessandro Vespignani [9] propuseram um modelo epidêmico em redes, onde vértices representam indivíduos e arestas representam as interações entre pares de indivíduos. Para analisar este modelo, utilizaram a premissa de que indivíduos com o mesmo grau k são estatisticamente equivalentes (*degree block approximation* ou *degree-based mean field*). Isso implica que ao invés de considerar o estado de um indivíduo, podemos considerar o estado de todos os indivíduos com grau k [16].

2.2.1 Modelo SI em redes

Seguindo o estudo descrito em [9], para o modelo SI podemos determinar a fração de indivíduos com grau l que estão infectados pelo total de indivíduos com grau l ,

N_l :

$$i_l = \frac{I_l}{N_l} \quad (2.7)$$

Repare que a fração total de indivíduos infectados na rede é dada pela média da fração de infectados para todos os graus: $i = \sum_l p_l i_l$, onde p_l é a fração de indivíduos com grau l .

Como os indivíduos estão agrupados pelo grau, a taxa de contágio vai ser diferente para cada classe l :

$$\frac{di_l}{dt} = \beta(1 - i_l)l\Theta_l, \quad (2.8)$$

onde $(1 - i_l)$ é a fração de indivíduos suscetíveis com grau l e Θ_l é a fração de vizinhos infectados de um vértice suscetível com grau l , chamada de função de densidade. Repare que Θ_l era a fração total de indivíduos infectados para o modelo anterior (premissa do *homogeneous mixing*). Além disso, neste modelo teremos l_{max} equações (uma para cada grau) que conjuntamente determinam o comportamento da epidemia.

No início da epidemia, a fração de indivíduos infectados para qualquer classe l é pequena. Por isso a equação 2.8 pode ser aproximada por:

$$\frac{di_l}{dt} \approx \beta l \Theta_l, \quad (2.9)$$

A função densidade Θ_l para uma rede sem correlação de graus (nas pontas de uma aresta) é independente de l , e como é mostrado em [1], $\Theta_l = i_0 \frac{\langle l \rangle - 1}{\langle l \rangle} e^{t/\tau^{SI}}$. Assim teremos:

$$\frac{di_l}{dt} \approx \beta l i_0 \frac{\langle k \rangle - 1}{\langle k \rangle} e^{t/\tau^{SI}}, \quad (2.10)$$

onde $\langle k \rangle$ é o grau médio da rede e τ^{SI} é:

$$\tau^{SI} = \frac{\langle k \rangle}{\beta(\langle k^2 \rangle - \langle k \rangle)} \quad (2.11)$$

Para obter a fração de indivíduos infectados com grau l , basta integrar a equação 2.10:

$$i_k = i_0 \left(1 + \frac{k(\langle k \rangle - 1)}{\langle k^2 \rangle - \langle k \rangle} (e^{t/\tau^{SI}} - 1) \right) \quad (2.12)$$

Repare que a fração de indivíduos infectados para a classe l depende de l , como esperado. Isso significa que quanto maior o grau de um vértice, maior a chance dele ser infectado. A figura 2.5 mostra que o crescimento da fração de indivíduos infectados depende de seu grau e que quanto maior o grau, mais rapidamente cresce esta fração. Isso mostra a importância dos *hubs* em redes já que eles são os principais responsáveis por disseminar a epidemia na rede.

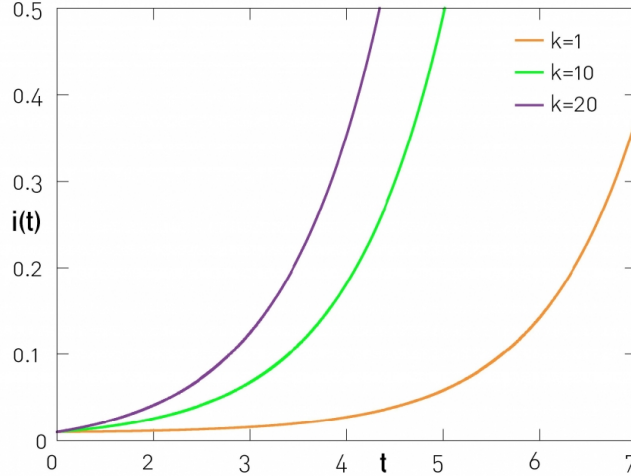


Figura 2.5: Fração de vértices infectados para cada classe de grau k no modelo SI. Extraída de [1]

O tempo característico τ^{SI} mostra que a velocidade de propagação da epidemia depende tanto do grau médio $\langle k \rangle$ da rede quanto do segundo momento da distribuição de grau, através de $\langle k^2 \rangle$. Por isso é interessante entender o comportamento de τ^{SI} para diferentes redes. Por exemplo, para o modelo $G(n, p)$ de grafo aleatório, onde os vértices tem aproximadamente a mesma quantidade de vizinhos, $\langle k^2 \rangle = \langle k \rangle (\langle k \rangle + 1)$, e portanto:

$$\tau^{SI} = \frac{1}{\beta \langle k \rangle}, \quad (2.13)$$

é exatamente igual ao tempo característico para o modelo SI sem redes. Para modelos de redes que apresentam distribuição de grau livre de escala o comportamento da epidemia depende do expoente α . Para $\alpha > 3$ tanto $\langle k \rangle$ quanto $\langle k^2 \rangle$ são finitos e portanto τ^{SI} comporta-se de maneira semelhante a um grafo aleatório divergindo apenas do ponto característico.

Por outro lado, para $\alpha \leq 3$ e no limite em que $N \rightarrow \infty$, onde N é a população total, então $\langle k^2 \rangle \rightarrow \infty$ levando $\tau^{SI} \rightarrow 0$. Isto mostra que a propagação da epidemia em redes livres de escala é praticamente instantânea, mostrando a importância dos *hubs*, que funcionam como *super-spreads*.

Outras redes com distribuição de grau heterogênea também reduzem o tempo característico τ^{SI} , afetando diretamente a velocidade com que a epidemia se espalha na rede. De fato, basta que $\langle k^2 \rangle > \langle k \rangle (\langle k \rangle + 1)$ para aumentar a velocidade média de propagação da infecção. No modelo SI, ao final da epidemia todos os indivíduos estão infectados, mas a velocidade com que a epidemia chega no regime de saturação depende do tempo característico τ^{SI} que é fundamentalmente influenciado pela estrutura da rede.

2.2.2 Modelo SIS em redes

Como no modelo SIS os indivíduos são capazes de se recuperar da infecção, precisamos adicionar a equação a taxa de recuperação, obtendo:

$$\frac{di_l}{dt} = \beta(1 - i_l)l\Theta_l - \mu i_l \quad (2.14)$$

A partir de 2.14 chega-se no tempo característico para epidemias SIS (vide [1]):

$$\tau^{SIS} = \frac{\langle k \rangle}{\beta \langle k^2 \rangle - \mu \langle k \rangle} \quad (2.15)$$

Repare que na epidemia SIS a taxa de recuperação μ possui um papel importante pois quanto maior seu valor, τ^{SIS} torna-se negativo e portanto a fração de indivíduos infectados decresce exponencialmente levando ao término da epidemia. Vemos novamente que a estrutura da rede possui papel fundamental em τ^{SIS} através de $\langle k^2 \rangle$. Para entender o comportamento da epidemia seja

$$\varphi = \frac{\beta}{\mu} \quad (2.16)$$

a taxa de propagação da epidemia que depende somente das características do patógeno. Esta relação indica que quanto maior φ mais chances a epidemia tem de se espalhar pela rede. Através de φ podemos definir o *threshold* da epidemia, φ_c , que determina que se $\varphi > \varphi_c$, então a epidemia vai se espelhar pela rede e atingir um estado endêmico.

Para o modelo $G(n, p)$ de grafo aleatório, $\langle k^2 \rangle = \langle k \rangle (\langle k \rangle + 1)$ e utilizando as equações 2.15 e 2.16 chegamos ao seguinte *threshold*:

$$\varphi_c = \frac{1}{\langle k \rangle + 1} \quad (2.17)$$

Repare que para este modelo de grafo aleatório o *threshold* não depende da distribuição de grau e sempre terá um valor finito já que $\langle k \rangle$ é sempre finito. Isso implica que se $\varphi > \varphi_c$ então a epidemia sempre se espalha na rede até alcançar seu estado endêmico. Mas caso $\varphi < \varphi_c$ então a epidemia não se espalha, terminando rapidamente.

Para modelos de redes livres de escala, o *threshold* da epidemia é:

$$\varphi_c = \frac{\langle k \rangle}{\langle k^2 \rangle} \quad (2.18)$$

Desta vez, o *threshold* é diretamente influenciado pelo segundo momento da distribuição de grau da rede. Se o expoente α da lei de potência for $\alpha \leq 3$ então, $\langle k^2 \rangle$ diverge quando $N \rightarrow \infty$, e com isto não temos *threshold* para epidemias SIS.

Ou seja, neste regime a epidemia sempre se espalha até atingir o estado endêmico. Isto é resultado direto da presença de *hubs* em redes livres de escala pois mesmo para patógenos com baixa taxa de transmissão, eles são capazes de contaminar um grande número de indivíduos.

Note que as equações 2.11 e 2.17 mostram que o tempo característico da epidemia τ e o *threshold* φ_c dependem fundamentalmente da estrutura da rede através de $\langle k \rangle$ e $\langle k^2 \rangle$, e por isso qualquer rede com distribuição de grau heterogênea afeta o comportamento da epidemia. De fato, basta que $\langle k^2 \rangle$ seja maior que $\langle k \rangle(\langle k \rangle + 1)$ para que isso ocorra.

Em suma, o trabalho de Romualdo *et al.* [9] mostra que a estrutura da rede é fundamental para a propagação da epidemia e que não pode ser negligenciada quando estudamos processos de propagação.

2.2.3 Duração de epidemias SIS e estrutura das redes

Ganesh *et al.* propõem um estudo para entender a influência da estrutura da rede na duração de epidemias [17]. Mais precisamente, identificar propriedades estruturais da rede que determinam a persistência da epidemia SIS por um período muito longo. O modelo clássico de epidemia em redes é considerado, onde os vértices representam os indivíduos e as arestas o contato entre eles, e a probabilidade de um indivíduo ser contaminado, β , é proporcional ao número de vizinhos infectados. O indivíduo recupera-se a uma taxa γ .

Em particular, o estudo busca determinar o comportamento do valor esperado ($E[\tau]$) do tempo até a epidemia terminar, τ . Para determinar uma condição suficiente para a epidemia terminar rapidamente, foi utilizado o raio espectral $\rho(A)$ da matriz de adjacência A que representa a rede, que é seu maior autovalor. Os autores mostram que para uma epidemia terminar rapidamente, a seguinte condição é suficiente:

$$\rho(A) < \frac{\gamma}{\beta}, \quad (2.19)$$

e que neste caso $E[\tau] = \mathcal{O}(\log n)$, ou seja, o tempo médio de duração da epidemia é logarítmico no tamanho n da rede.

Por outro lado, uma condição suficiente para uma epidemia ter longa duração é:

$$\frac{\gamma}{\beta\eta(m)} < 1, \quad (2.20)$$

onde m é o número de arestas e $\eta(m)$ é a *generalized isoperimetric constant* da rede e que neste caso $E[\tau] = \Omega(e^{n^a})$ para uma constante $a > 1$, ou seja, a duração da epidemia é super-exponencial no tamanho da rede.

2.2.4 Epidemia em redes onde indivíduos se movimentam

O modelo de Romualdo *et al.* representa os indivíduos como vértices da rede que permanece estático ao longo da epidemia constituindo o modelo clássico de epidemia em redes. Assim sendo, este modelo não captura a mobilidade dos indivíduos, que permite trocar suas interações ao longo do tempo. Para capturar esta característica, modelos que atribuem mobilidade aos indivíduos vêm sendo propostos, utilizando os vértices da rede como localidades e arestas representam a possibilidade de movimentação.

Draief *et al.* [10] consideram um modelo epidêmico em redes com agentes móveis que executam passeios aleatórios independentes visando determinar a influência da estrutura da rede na propagação da epidemia. Um grafo $G(V, E)$ representa as localidades por onde os indivíduos podem caminhar. Cada indivíduo pode ser contaminado a uma taxa β e isso ocorre quando um indivíduo suscetível encontra outro infectado em um mesmo vértice. A contribuição deste trabalho é mostrar que o tempo que os indivíduos permanecem juntos em um vértice depende da rede e difere para grafo regulares e redes livres de escala.

Em particular, para grafos regulares (e.g. completo) o valor esperado do tempo $\tau(t)$ que um indivíduo suscetível permanece no mesmo vértice que um indivíduo infectado durante o período de tempo $[0, t]$ é dado por:

$$E[\tau(t)] = \frac{t}{n}, \quad (2.21)$$

onde n é o número de vértices da rede, indicando que $\tau(t)$ é inversamente proporcional ao número de vértices. Além disso, os autores mostram que a probabilidade $\varphi(t)$ do indivíduo permanecer infectado até o tempo t é:

$$\varphi(t) \leq 1 - e^{(-\beta t)/n}, \quad (2.22)$$

ou seja, a probabilidade de um indivíduo não ser infectado decresce exponencialmente com o tempo. Estes resultados estão de acordo com a intuição já que quanto maior a rede, os encontros são mais raros e por isso o tempo $\tau(t)$ decresce. Os resultados são os mesmo para o modelo $G(n, p)$ e redes livre de escala com expoente maior que 3. Por outro lado, para redes livre de escala expoente $\alpha \leq 3$ os indivíduos passam mais tempo juntos, principalmente nos vértices de maior grau permitindo que a infecção ocorra mais rapidamente [10].

2.3 Simuladores de epidemias

Além do trabalho de modelagem matemática de epidemia em redes e do tratamento analítico destes modelos, outra abordagem para caracterizar epidemias é o desenvolvimento de simuladores de epidemias. Esta abordagem permite capturar aspectos mais reais de uma epidemia e responder diferentes perguntas, assim como auxiliar o processo de tomada de decisões para prevenção e contenção de epidemias.

A epidemiologia computacional [18] surge como uma área interdisciplinar com o objetivo de desenvolver e utilizar modelos computacionais para entender e caracterizar o comportamento de epidemias. Estes modelos permitem simular diversos cenários epidêmicos e possibilitam intervenções em tempo real para avaliar diversas estratégias de contenção de epidemias, como vacinação e quarentena.

2.3.1 GLEAM

GLEAM (Global epidemic mobility model) [19] é um simulador de epidemias global cujo objetivo é analisar e prever o comportamento de epidemias permitindo aplicar estratégias de intervenção para reduzir o impacto das epidemias. Para produzir epidemias realísticas, o modelo é baseado em três camadas: a primeira é composta por dados reais da população global, na segunda são utilizados dados reais de mobilidade da população e última camada representa um modelo matemático com a dinâmica de infecção. As duas primeiras camadas são utilizadas para determinar onde e quando os indivíduos interagem e potencialmente transmitem a infecção.

A população é agrupada em células, representada por vértices, que estão atreladas ao aeroporto comercial mais próximo da célula. Utilizando dados globais de vôos comerciais, construiu-se uma rede de aeroportos em cima das células. Além disso, os autores utilizaram dados de censos populacionais para atribuir densidade populacional a cada uma das células, variando de 10 pessoas a 100 mil pessoas. O simulador é capaz de parametrizar o modelo de estados epidêmicos dos indivíduos e o mecanismo de contágio do patógeno de modo que possamos simular diversos tipos de epidemias.

O GLEAM foi utilizado em 2009 para estudar o comportamento da epidemia H1N1 e investigar as medidas de intervenção que foram implementadas a época. O simulador foi carregado com dados reais da epidemia e foi capaz de prever o pico da epidemia, no outono do hemisfério norte, contrariando os picos sazonais em janeiro e fevereiro. O pico da epidemia coincidiu justamente com o início das campanhas de vacinação, tendo a mesma pouca efetividade em conter o início e espalhamento da epidemia global.

2.3.2 EpiSimdemics

Barret *et al.* descrevem um simulador de epidemia escalável para redes reais que permite realizar políticas de contenção em tempo real, *EpiSimdemics* [20]. Uma rede de contatos representada por um grafo bipartido que varia no tempo é avaliado, onde uma classe de vértices representa as pessoas e a outra representa as localidades. Uma aresta é adicionada entre os vértices caso uma pessoa decida visitar uma localidade, permanecendo um intervalo de tempo t .

Os indivíduos podem estar no estado epidêmico Suscetível, Latente, Infectado ou Recuperado. O estado epidêmico Latente representa o estado intermediário entre o estado Suscetível e Infectado, onde o indivíduo não pode disseminar a infecção. A transmissão da infecção é possível quando um ou mais indivíduos decidem visitar a mesma localidade e pelo menos um deles está infectado.

Para construir a rede de interação entre os indivíduos foi utilizado dados do censo populacional dos EUA, de modo que foi possível capturar residências, escolas e padrões de deslocamento. No entanto, o modelo proposto não permite representar a estrutura de movimentação entre as localidades, pois indivíduos trocam de qualquer localidade para qualquer outra localidade (de forma aleatória).

Capítulo 3

Modelo epidêmico em redes com passeios aleatórios

Modelos tradicionais de epidemia em rede utilizam os vértices para representar os indivíduos de uma epidemia e as arestas para representar a interação (contato) entre dois indivíduos, ou seja, a possibilidade de contágio. Neste modelo, os indivíduos sempre interagem com os mesmo vértices vizinhos. Neste trabalho, consideramos um modelo de epidemia em redes no qual indivíduos se movimentam entre localidades, e a rede representa as localidades (vértices) e a possibilidade de deslocamento entre elas (arestas). Repare que este modelo pode tanto representar a propagação de vírus em dispositivos móveis através de conexões como Bluetooth ou epidemias entre pessoas como a de SARS.

3.1 Modelo

Seja $G(V, E)$, onde V é o conjunto de vértices que representam as localidades e E é o conjunto de arestas representando a possibilidade de mobilidade direta entre as localidades. Definimos também K como o número de indivíduos que se movimentam pela rede.

Iremos considerar que os indivíduos se movimentam pela rede de acordo com passeios aleatórios independentes. Ou seja, quando um indivíduo está no vértice v , seu próximo vértice é escolhido aleatoriamente entre os vizinhos de v , com probabilidade uniforme. O tempo que um indivíduo permanece em um vértice antes de se locomover para um vizinho é dado por uma variável aleatória exponencial, de densidade $P(t) = \lambda e^{-\lambda t}$, onde $1/\lambda$ é o tempo médio de permanência no vértice. Desta forma, ao entrar em um vértice, o indivíduo permanece um tempo aleatório até se movimentar para um vértice vizinho. Iremos assumir que o tempo efetivo para realizar a movimentação para o vértice vizinho é zero.

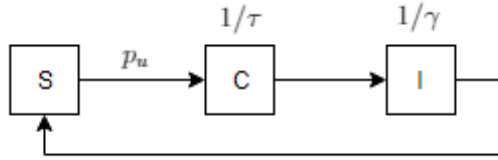


Figura 3.1: Transição entre os estados epidêmicos

Os estados epidêmicos dos indivíduos e suas transições seguem o modelo SCIS. Neste modelo generalizamos o modelo SIS para acrescentar um estado intermediário chamado Contraído (C) que representa indivíduos que estão contaminados mas não podem propagar a epidemia. Este estado epidêmico é comum em doenças que apresentam forma latente ou o período de incubação de vírus. As transições entre estados epidêmicos para este modelo são: $S \rightarrow C \rightarrow I \rightarrow S$, onde um indivíduo pode passar por este ciclo diversas vezes.

A medida que os indivíduos se movimentam pela rede, ocorrem encontros entre indivíduos infectados e suscetíveis. Em cada encontro deste tipo o contágio ocorre de acordo com uma variável aleatória de Bernoulli de parâmetro p_u , onde u é um vértice da rede. Ou seja, com probabilidade p_u o indivíduo susceptível é contagiado pelo indivíduo infectado no instante do encontro. A associação da probabilidade de transmissão p_u ao vértice permite modelar localidades da rede com diferentes taxas de infecção.

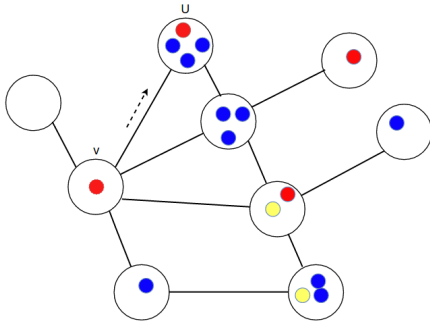


Figura 3.2: Indivíduo infectado (em vermelho) entra em um vértice ($v \rightarrow u$) com indivíduos suscetíveis.

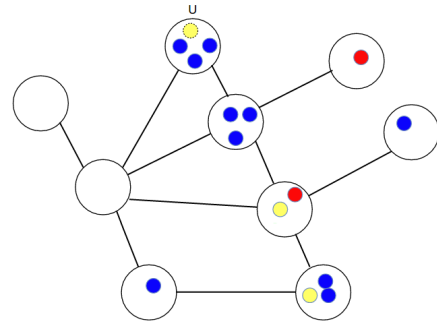


Figura 3.3: Indivíduo no estado contraído (em amarelo) torna-se infectado em um vértice ($u \rightarrow v$) com indivíduos suscetíveis.

Os indivíduos permanecem no estado suscetível até que sejam infectados e quando isso ocorre, eles passam imediatamente para o estado contraído. O tempo de permanência no estado contraído possui distribuição exponencial $P_{cont}(t) = \tau e^{-\tau t}$. Ao final deste tempo o indivíduo muda para o estado Infectado. Uma vez infectado, o indivíduo permanece neste estado por um tempo aleatório com distribuição

exponencial $P_{inf}(t) = \gamma e^{-\gamma t}$. A Figura 3.1 ilustra a transição e os tempos médios entre os estados epidêmicos. Repare que como o tempo de permanência nos estados são gerados por variáveis aleatórias exponenciais, o tempo médio de permanência nos estados contraído e infectado são $1/\tau$ e $1/\gamma$, respectivamente. Além disso, a transição entre o estado suscetível e contraído determinado por p_u ocorre em tempo zero.

O tempo médio de permanência dos indivíduos nos estados epidêmicos Contraído e Infectado obedece as variáveis aleatórias de cada evento. No caso do estado epidêmico suscetível, o indivíduo permanece neste estado enquanto ele não for contagiado, e quanto isso ocorre ele passa para o estado contraído em tempo zero, pois a transição é imediata. Em particular, quando um indivíduo infectado se recupera em um vértice com muitos indivíduos infectados, ele tem alta probabilidade de ser novamente infectado, passando imediatamente para o estado contraído. Quando essa situação se repete frequentemente na epidemia (para K grande em relação a n), o tempo médio que o indivíduo fica no estado epidêmico suscetível tende a zero. Repare que nestas situações, o indivíduo passa para o estado suscetível e depois para o estado contraído no mesmo instante de tempo t .

Repare que a mobilidade dos indivíduos tem um papel fundamental na propagação da epidemia pois um indivíduo pode caminhar pela rede enquanto está infectado. Mais precisamente, esta mobilidade influencia a velocidade com a qual a epidemia alcança a maior parte da população. Por isso a relação entre λ e γ é fundamental. Note que $1/\lambda$ é o tempo médio de permanência de um indivíduo no vértice e $1/\gamma$ é o tempo médio que o indivíduo fica infectado. Por exemplo, se $\lambda = 10$ e $\gamma = 1$, em média o indivíduo vai visitar 10 localidades enquanto infectado.

Outro parâmetro importante do modelo é a taxa de infectar τ que determina o tempo que o indivíduo passa no estado contraído, em média é $1/\tau$. Repare que quanto mais tempo o indivíduo passa no estado contraído, mais lento será a propagação da epidemia pois enquanto ele está neste estado ele não pode infectar outros indivíduos suscetíveis. Este período no estado contraído funciona como uma barreira, impedindo que a epidemia se alastre rapidamente.

Os encontros entre indivíduos infectados e suscetíveis estão divididos em dois tipos: um indivíduo infectado encontra um suscetível, I-S, ou um indivíduo suscetível encontra um infectado, S-I. O primeiro tipo ocorre quando um indivíduo infectado entra em um vértice com pelo menos um indivíduo suscetível, Figura 3.2, ou quando um indivíduo, no estado contraído, torna-se infectado em um vértice com pelo menos um indivíduo suscetível, Figura 3.3. Para cada um destes encontros I-S, o indivíduo suscetível contrai a infecção com probabilidade p_u e permanece suscetível com probabilidade $(1 - p_u)$. Caso haja mais de um indivíduo suscetível, cada um deles tem chance de ser infectado, independentemente dos outros.

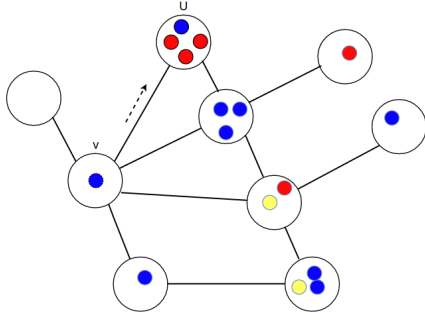


Figura 3.4: Indivíduo suscetível (em vermelho) entra em um vértice ($v \rightarrow u$) com indivíduos infectados.

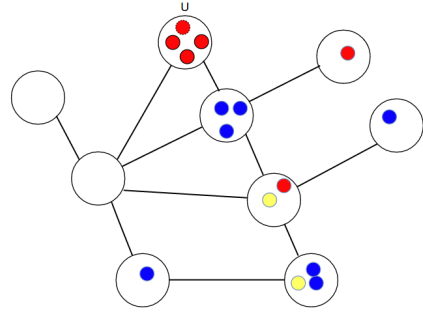


Figura 3.5: Indivíduo infectado torna-se suscetível em um vértice (u) com indivíduos infectados.

Já os encontros S-I ocorrem quando um indivíduo suscetível entra em um vértice com pelo menos um indivíduo infectado, Figura 3.4, ou quando um indivíduo infectado se recupera em um vértice que possui ao menos outro indivíduo infectado, Figura 3.5. A probabilidade de infecção nestes casos é dada por:

$$1 - (1 - p_u)^{k_u}, \quad (3.1)$$

onde p_u é a probabilidade de infecção associada ao vértice do encontro, e k_u é o número de indivíduos infectados no vértice u . Esta probabilidade é a complementar da probabilidade do indivíduo suscetível não ser infectado por nenhum dos infectados.

A Tabela 3.1 apresenta os parâmetros do modelo assim como as respectivas definições.

Tabela 3.1: Variáveis do modelo

Variável	Descrição
n	número de vértices da rede
K	número de indivíduos (passeios aleatórios)
k_i	população inicial de infectados (passeios aleatórios)
λ	taxa de caminhar (distribuição exponencial)
γ	taxa de recuperar (distribuição exponencial)
τ	taxa de infectar (distribuição exponencial)
p_u	probabilidade de contágio no vértice u

Capítulo 4

Simulador

Modelos computacionais epidêmicos tem sido desenvolvidos para estudar e caracterizar o comportamento de epidemias [6, 21]. Estes modelos capturam o comportamento de epidemias através de diversos parâmetros, permitindo simular diferentes cenários epidêmicos.

4.1 Arquitetura

Para estudar o modelo de epidemia em redes proposto, projetamos e implementamos um simulador de eventos discretos para epidemias SCIS em redes. Os eventos do simulador são *Caminhar* (C), *Infectar* (I) e *Recuperar* (R), sendo cada evento aplicado a um indivíduo em um determinado instante de tempo. Ou seja, cada evento é representado por um tipo de evento $\{C, I, R\}$, o indivíduo a quem pertence o evento $i \in \{1, \dots, K\}$ e o seu tempo de ocorrência t no relógio de simulação. Os eventos da lista estão ordenados crescentemente de acordo com t . O evento $\{C, i, t\}$ é utilizado para representar o deslocamento do indivíduo i do vértice atual para um vértice vizinho. O evento $\{I, i, t\}$ representa a transição do estado contraído para o infectado, para o indivíduo i . Por fim, o evento $\{R, i, t\}$ representa a recuperação do indivíduo i , ou seja, a transição do estado infectado para o suscetível.

A figura 4.1 apresenta o diagrama de classes do simulador. A classe principal do simulador é a *Simulator* onde implementamos o laço principal do simulador responsável por processar os eventos. A classe *Simulator* armazena todos os eventos dos indivíduos em uma lista de prioridade mínima que está organizada em ordem crescente do tempo de ocorrência dos eventos de modo que sempre o primeiro evento a ocorrer será o primeiro a ser processado.

Os indivíduos da simulação são modelados pela classe *RandomWalk* que está associada à classe *Vertex* e à classe *Simulator*. A associação do indivíduo a um vértice indica em qual vértice da rede o indivíduo está no instante de tempo t . Para isso, o indivíduo armazena o identificador do seu vértice atual, alterando sempre

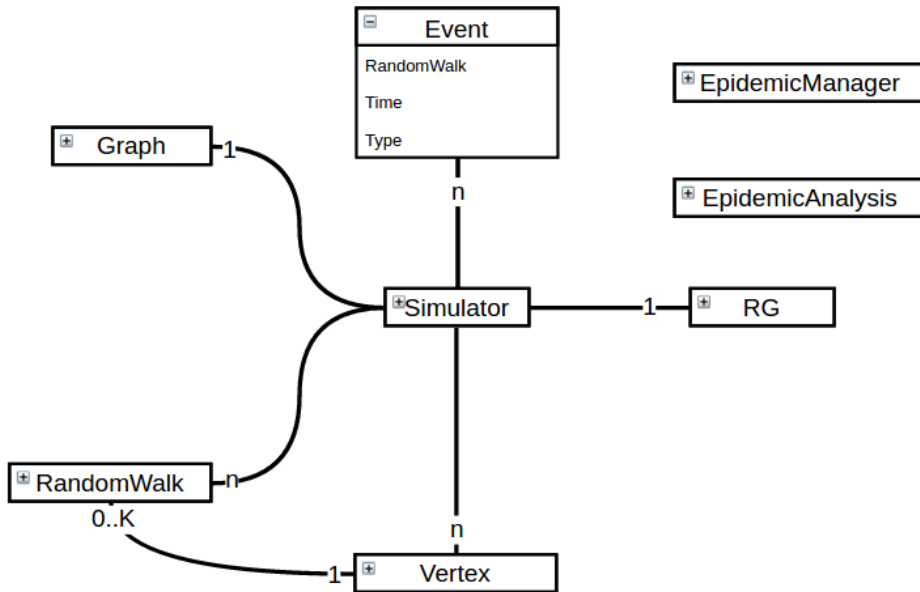


Figura 4.1: Diagrama de classes do simulador

que ele se movimenta. A classe *Simulator* pode ter um ou mais indivíduos que participarão da simulação e que são armazenados em um vetor onde cada posição representa o identificador do indivíduo, permitindo recuperar o indivíduo em tempo constante.

A classe *Vertex* modela os vértices da rede e está associada à classe *RandomWalk* e *Simulator*. Repare que um vértice possui um número variável de indivíduos já que eles se movimentam livremente pela rede. Os indivíduos são armazenados nos vértices por meio de uma lista duplamente encadeada para ter custo constante nas operações inserir e remover. A operação inserir adiciona um indivíduo no topo da lista, ao custo constante. Para remover um indivíduo da lista, cada indivíduo carrega consigo um ponteiro para sua posição na lista do seu vértice atual. Logo, quando um indivíduo muda de vértice, removemos ele da lista do vértice origem utilizando o ponteiro para a sua posição e adicionamos ele no topo da lista do vértice destino, sempre atualizando o ponteiro. Desta forma, a movimentação dos indivíduos não impacta a complexidade do simulador.

A classe *RG* é responsável por gerar todas as variáveis aleatórias utilizadas no simulador que são: variáveis aleatórias exponenciais para os eventos Caminhar, Infec-tar e Recuperar, a Bernoulli que determina se o indivíduo será infectado e a Uniforme que indica qual o próximo vértice que o indivíduo irá visitar ao movimentar-se.

A classe *Graph* foi criada para armazenar e manipular o grafo $G(V, E)$. O grafo é criado utilizando vetores de adjacência onde cada índice do vetor está associado a um vértice da rede contendo um vetor com seus vizinhos. O simulador possui um

gerador de grafos capaz de gerar as seguintes redes: grafo completo, anel, torus e grafo bipartido. Além de criar o próprio grafo, o simulador pode carregar um grafo já existente lendo um arquivo de texto contendo os vértices e suas arestas.

O simulador gera algumas medidas de interesse por meio de logs durante a simulação. Estes logs são utilizados para validar o simulador e para caracterizar o comportamento epidemia e a classe responsável por isso é a *EpidemicAnalysis*. Por fim, a classe *EpidemicManager* é responsável por instanciar o simulador e determinar qual operação será executada de acordo com os parâmetros fornecidos.

4.2 Implementação

A lista de eventos sempre possui um evento *Caminhar* para cada indivíduo, uma vez que o processamento deste evento gera outro evento *Caminhar* para o mesmo indivíduo. Ao processar o evento caminhar $\{C, i, t\}$, caso o indivíduo i esteja no estado suscetível e seja contagiado ao entrar no vértice de destino, é gerado um evento *Infectar* para o indivíduo (repare que a transição para o estado contraído é imediata, ocorrendo no mesmo instante do encontro). Por outro lado, se o indivíduo que se movimenta estiver no estado infectado, é gerado um evento *Infectar* para todo indivíduo que ele contagiar no vértice destino. Lembrando que o contágio ocorre com probabilidade p_u quando o encontro ocorre no vértice u . Ao final, é gerado um novo evento *Caminhar* cujo tempo de ocorrência é $t_1 = t + \Delta_1$, sendo Δ_1 uma variável aleatória exponencial com parâmetro λ , definida na seção 3.

O processamento do evento $\{I, i, t\}$ muda o estado do indivíduo de contraído para infectado. Ao tornar-se infectado, o indivíduo pode infectar indivíduos suscetíveis (caso haja) que estão no mesmo vértice que ele, gerando um novo evento *Infectar* para cada indivíduo contagiado. O processamento do evento *Infectar* gera um evento *Recuperar* para o mesmo indivíduo, ou seja, é gerado um evento $\{R, i, t_1\}$, cujo tempo de ocorrência é $t_1 = t + \Delta_2$, sendo Δ_2 uma variável aleatória exponencial com parâmetro γ .

Ao processar o evento $\{R, i, t\}$, o indivíduo i muda do estado infectado para o estado suscetível. Quando isso ocorre, o indivíduo pode ser novamente contagiado por indivíduos infectados que estejam no mesmo vértice que ele. Deste modo, no mesmo instante de tempo t , o indivíduo passa do estado infectado para o suscetível e do suscetível para o contraído, e um novo evento *Infectar* é gerado. Todo evento *Infectar* é gerado no instante em que o indivíduo muda para o estado epidêmico Contraído. Em particular, o evento $\{I, i, t_1\}$ é gerado no instante t no qual o indivíduo i é contagiado, onde $t_1 = t + \Delta_3$, sendo Δ_3 uma variável aleatória exponencial com parâmetro τ . O processamento dos eventos é ilustrado pelo diagrama de fluxos na figura 4.2.

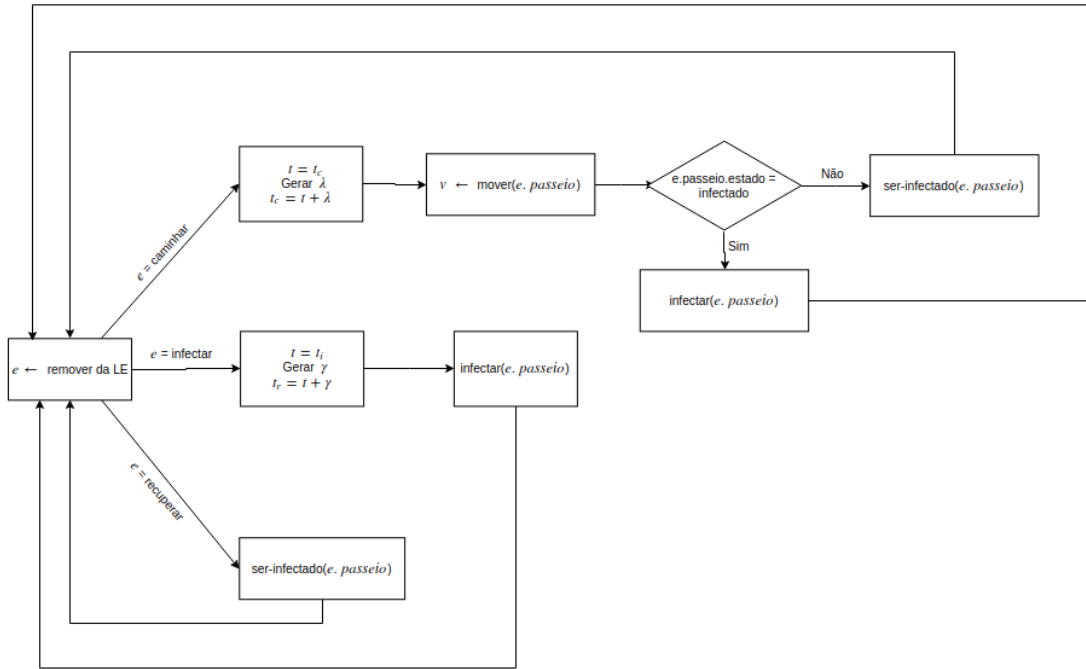


Figura 4.2: Fluxo de processamento dos eventos do simulador

O algoritmo 1 mostra o laço principal do simulador que processa o próximo evento da lista de eventos, gerando os eventos necessários. Repare que a simulação termina quando não há mais indivíduos infectados ou o tempo de simulação atinge o limite de T . As funções *ser-infectado()* e *infectar()* são fundamentais para a complexidade do simulador e serão discutidas em seguida.

O tamanho da lista de eventos do simulador é proporcional ao número de indivíduos K , havendo sempre um evento *Caminhar* para cada indivíduo, já que o processamento de um evento *Caminhar* gera outro evento do mesmo tipo. Os eventos *Infectar* e *Recuperar* são mutuamente exclusivos, pois caso o indivíduo esteja no estado contraído, existe um evento *Infectar* na lista de eventos e caso ele esteja no estado infectado, existe um evento *Recuperar* na lista de eventos. Além disso, no estado Suscetível, o Indivíduo não possui nenhum dos dois eventos. Portanto, cada indivíduo pode ter no mínimo um e no máximo dois eventos na lista de eventos. Assim, o tamanho da lista de eventos respeita a seguinte desigualdade:

$$K \leq |L| \leq 2K \quad (4.1)$$

4.3 Medidas de Interesse

Durante a simulação da epidemia, são geradas medidas de interesse para capturar o estado do sistema e permitir a análise do comportamento da epidemia. Para registrar

Algorithm 1 Laço principal do simulador

```
while  $S_t \neq K$  e tempo de simulação  $< T$  do
   $e \leftarrow$  remover da LE
  if  $e.tipo = Caminhar$  then
    Gerar Caminhar para  $e.passeio$ 
    Inserir Caminhar na LE
     $v \leftarrow$  mover( $e.passeio$ )
    if  $e.passeio.estado = infectado$  then
      infectar( $e.passeio$ )
    else if  $e.passeio.estado == suscetivel$  then
      ser-infectado( $e.passeio$ )
    end if
  else if  $e.tipo = Infectar$  then
    Gerar Recuperar
    Inserir Recuperar na LE
    infectar( $e.passeio.vertice$ )
  else if  $e.tipo = Recuperar$  then
    ser-infectado( $e.passeio.vertice$ )
  end if
end while
```

o número de indivíduos em cada estado epidêmico ao longo do tempo, definimos: S_t como o número de indivíduos suscetíveis no instante t , I_t , o número de indivíduos infectados no instante de tempo t e C_t , o número de indivíduos no estado contraído no instante t , de forma que $S_t + I_t + C_t = K$. Repare que uma epidemia termina assim que todos os indivíduos se encontram no estado suscetível. Neste caso, não há mais possibilidade de contágio e os indivíduos apenas se movimentam pela rede. Seja T_0 , o instante de tempo que a epidemia termina, dado por $T_0 = \inf\{t : S_t = K\}$.

Além do tempo de término da epidemia, o simulador registra diversas outras medidas de interesse, como a distribuição do tempo que os indivíduos permanecem em cada estado epidêmico (Suscetível, Contraído e Infectado), a distribuição do número de infectados e suscetíveis ao longo do tempo, e a fração das infecções que ocorrem em cada vértice da rede. O simulador gera arquivos de saída com dados pertinentes a cada execução, que são analisados por um outro programa para gerar tais medidas de interesse.

Para ter uma visão melhor dos dados gerados pelo simulador, a figura 4.3 mostra o arquivo que registra o número de infectados $I(t)$, o número de contraídos $I(t)$ e o número de suscetíveis $S(t)$. Estes dados são registrados em cada instante de tempo t que houve mudança em um dos estados epidêmicos.

Em outro arquivo (figura 4.4) registramos, para cada indivíduo, todos os eventos concernentes a ele e o instante de tempo que ocorreu. A partir deste arquivo podemos rastrear todos os vértices que ele visitou. Podemos ver também quando ocorreram

```

# tempo, infectados, contraídos, suscetíveis
0.005051,25,0,25
0.007739,24,0,26
0.014521,24,1,25
0.055903,24,2,24
0.055903,23,2,25
0.055903,23,3,24
0.074192,22,3,25
0.099648,22,4,24
0.123269,22,5,23
0.126623,21,5,24
0.193270,22,4,24
0.257856,22,5,23
0.257856,22,6,22
0.257856,22,7,21
0.263897,22,8,20

```

Figura 4.3: Log do simulador para $I(t)$, $C(t)$ e $S(t)$

seus eventos e em qual vértice, em especial, o indivíduo foi infectado (muda para estado contraído) e qual indivíduo o infectou.

```

# Random Walk: 0
# Parameters: Lambda: 1, Gama: 1, Tau: 1
# time,event,vertex,state,effect,infected_by
0.000000,Walk,16,Infected,-,
0.918821,Walk,12,Infected,-,
2.117847,Walk,3,Infected,-,
2.122959,Recover,3,Susceptible,-,
2.756103,Walk,17,Susceptible,-,
2.756103,-,17,Contracted,Infected,
2.820374,Infect,17,Infected,-,
2.842730,Walk,24,Infected,-,
3.332011,Recover,24,Susceptible,-,
3.855185,-,24,Contracted,Infected,26
4.277798,Infect,24,Infected,-,
4.471074,Recover,24,Susceptible,-,
4.663869,-,24,Contracted,Infected,46

```

Figura 4.4: Log do simulador para o indivíduo 0

Este logs são utilizados pelo classe *EpidemicAnalysis* para gerar algumas medidas de interesse. A partir do arquivo mostrado na figura 4.3 geramos graficamente a evolução no tempo do número de indivíduos em cada estado epidêmico e calculamos também o intervalo de tempo que o sistema teve k indivíduos em cada estado. Igualmente, a partir do arquivo mostrado na figura 4.4 calculamos a CCDF para a taxa de caminhar, recuperar e infectar de cada indivíduo.

4.4 Utilização

O simulador opera em diversos cenários por meio dos de parâmetros especificados em um arquivo json. Os parâmetros do simulador variam desde parâmetros para o grafo até os parâmetros para as variáveis aleatórias. A figura 4.5 mostra o formato do arquivo json com os parâmetros do simulador.


```

{
  "GraphParam": {
    "Type": 3,
    "path": "/path/to/graph.txt",
    "n": 20,
    "n2": 10,
    "NS": "[10,50,100]"
  },
  "RwParam": {
    "k": 30,
    "ki": 10,
    "kiPer": 0.5,
    "lambda": 1,
    "lambdaS": "[0.1,0.2,0.3]",
    "gama": 1,
    "tau": 1,
    "rwParamArray": "[[rw, [rw, vertex, lambda, gama, tau, state (0-Sus, 1-Cont, 2-Inf)]]]"
  },
  "VertexParam":{
    "p": 0.5,
    "vertexParamArray": "[[0,0.4],[1,0.9]"
  },
  "DebugLevel": 4,
  "p": 0.5,
  "cont_simulation": "false",
  "Runs": 10,
  "Time": 100000,
  "TimeS": "[1000,10000]",
  "OutputDir": "/path/to/write/logs",
  "KS": "[10,20,30]",
  "Operation": "sim"
}

```

Figura 4.5: Parâmetros do simulador

O objeto *GraphParam* define os parâmetros para a construção da rede, onde o tipo da rede pode ser especificado no parâmetro *type*. Este parâmetro pode assumir os valores: 0 para carregar um grafo a partir de um arquivo de texto, 1 para o utilizar um anel, 2 para utilizar um torus, 3 para um grafo completo e finalmente 4 para utilizar um grafo bipartido. Para a opção 0, deve ser informado, no parâmetro *path*, o caminho do arquivo contendo o grafo. Nas opções de 1 a 4 o próprio simulador gera as redes. O parâmetro *n* é o número e vértices da rede e para o caso de um grafo bipartido, *n2* é o número de vértices da segunda classe. Note que é possível construir uma estrela passando $n = 1$ e especificando em *n2* o número de vértices folhas. Já o parâmetro *NS* é utilizado para simular a epidemia variando a quantidade de vértices.

Já o objeto *RwParam* define os parâmetros para os indivíduos (passeios aleatórios), onde *K* é o número de indivíduos na simulação e *kiPer* é o percentual de indivíduos infectados no início da simulação. O parâmetro *lambda* é o parâmetro da variável aleatória exponencial que determina o tempo que um indivíduo permanece em um vértice, *gama* é o parâmetro da variável aleatória exponencial que define o tempo que o indivíduo vai ficar no estado epidêmico infectado e *tau* é o parâmetro da variável aleatória exponencial que determina o tempo que o indivíduo vai ficar no estado epidêmico contraído. Note que estes três parâmetros tem caráter global, ou seja, são definidos para todos os indivíduos da simulação. Porém, podemos especificar parâmetros individuais para cada indivíduo utilizando o array *rwParamArray* onde cada posição refere-se a um indivíduo. Neste parâmetro, juntamente com o identificador do indivíduo (*rw*), podemos especificar o vértice onde o indivíduo

irá começar a simulação, os valores de λ , γ , e τ , além de seu estado epidêmico (S,C,I).

No objeto *VertexParam* podemos definir o parâmetro da variável aleatória de Bernoulli que determina se o indivíduo será contaminado durante um encontro. Para isto, basta utilizar o parâmetro p , que será carregado para todo vértice da rede. É possível especificar valores diferentes de p para cada vértice utilizando o array *vertexParamArray*, onde cada posição contém um array com o identificador do vértice $(0, \dots, n - 1)$ e o parâmetro p .

Existem também parâmetros globais da simulação, como o nível de debug (*DebugLevel*), o número de rodadas independentes de simulação (*Runs*). O tempo limite da simulação (*Time*), que determina quando a simulação deve parar, considerando os casos onde ainda haja indivíduos infectados. E o diretório onde serão salvos os arquivos de log do simulador (*OutputDir*).

Como as rodadas de simulação são independentes, o simulador executa cada uma em paralelo utilizando a API de programação paralela do c++ chamada OpenMP.

O parâmetro *Operation* define o tipo de operação que será executada pelo simulador. As operações possíveis são:

- *sim*: Executa a simulação da epidemia com *Runs* rodadas e gera os arquivos de log.
- *sim_by_k*: Executa *Runs* rodadas de simulação variando o número de indivíduos (K) utilizando o parâmetro KS . O objetivo desta operação é gerar dados para estudar o comportamento de transição de fase da epidemia.
- *sim_by_n*: Executa *Runs* rodadas de simulação variando o número de vértices da rede utilizando o parâmetro NS .
- *scalability_k*: Executa *Runs* rodadas de simulação variando o número de indivíduos utilizando KS e o tempo limite de simulação utilizando $TimeS$. Esta funcionalidade é utilizada para medir a escalabilidade do simulação em função do número de indivíduos.
- *scalability_vertices*: Executa *Runs* rodadas de simulação variando o número de vértices em NS . Esta funcionalidade tem por objetivo analisar a escalabilidade do simulador em função do número de vértices da rede.
- *scalability_time*: Executa *Runs* rodadas de simulação variando o tempo limite de simulação utilizando $TimeS$. Esta funcionalidade tem por objetivo analisar a escalabilidade do simulador em função do tempo de simulação.

Capítulo 5

Avaliação Numérica

5.1 Eficiência computacional

O objetivo do simulador é permitir estudar o comportamento de epidemias em diferentes cenários. Portanto, é fundamental que o simulador tenha um bom desempenho computacional, não somente teórico (complexidade de pior caso) mas também prático, apresentando uma boa escalabilidade do tempo de execução em função dos parâmetros do modelo. A complexidade computacional do simulador depende tanto do algoritmo 1 quanto dos algoritmos que processam cada tipo de evento. Um aspecto central são as estruturas de dados utilizadas para armazenar a lista de eventos, a rede e os indivíduos. Os eventos são armazenados em uma fila de prioridade mínima, onde a complexidade das operações de remover e inserir tem custo $\mathcal{O}(\log K)$. O evento caminhar precisa escolher um vértice vizinho na rede, mas isto tem complexidade constante $\mathcal{O}(1)$, pois os vizinhos de cada vértice são armazenados em vetores (podemos escolher aleatoriamente o próximo vértice em tempo constante). Desta forma a complexidade do laço principal do simulador não depende da estrutura e do tamanho da rede. Por outro lado, inserir o novo evento caminhar na lista de eventos tem complexidade $\mathcal{O}(\log K)$.

Na função *ser-infectado*, é utilizada a Equação 3.1, ao custo $\mathcal{O}(1)$, para determinar se o indivíduo será contagiado por outros indivíduos infectados no vértice destino. Já na função *infectar*, é preciso iterar sobre todos os indivíduos suscetíveis no vértice destino para determinar individualmente se cada um será (ou não) contagiado, ao custo $\mathcal{O}(k_u)$, onde k_u é o número de indivíduos no vértice u . Sempre que houver um contágio, um novo evento *Infectar* é gerado para o indivíduo recém infectado e inserido na lista de eventos ao custo $\mathcal{O}(\log K)$. Como os indivíduos se movimentam aleatoriamente, $k_u = K$, no pior caso, e conseqüentemente a operação *infectar* tem custo $\mathcal{O}(K \log K)$. Por fim, o custo de cada laço principal do simulador é $\mathcal{O}(K \log K)$.

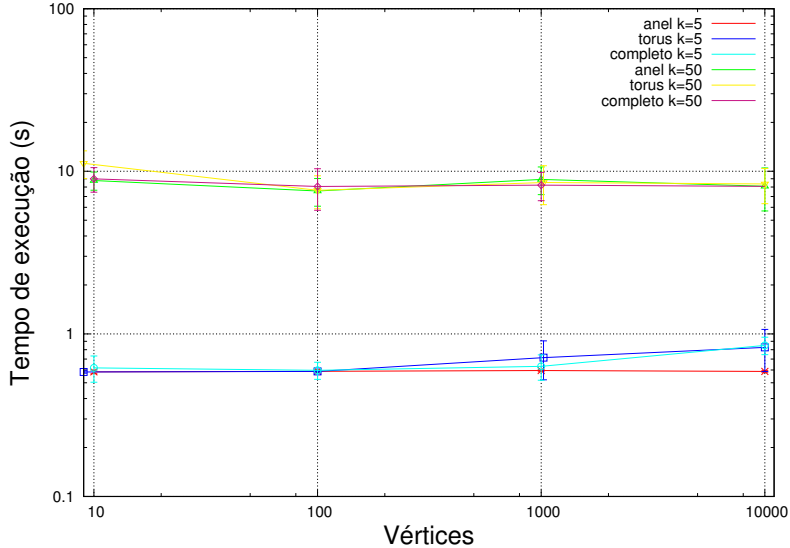


Figura 5.1: Tempo execução em função do tamanho da rede, n , para diferentes redes e número de indivíduos ($K = 5$, $K = 50$) em escala log-log. Parâmetros: $T = 10000$, $\lambda = 1$, $\gamma = 1$, $\tau = 1$.

Primeiro vamos avaliar empiricamente a escalabilidade do simulador em função do tamanho da rede. A análise teórica de pior caso indica que o simulador tem custo constante ($\mathcal{O}(1)$) no tamanho da rede, n , para um tempo de simulação fixo T . A figura 5.1 apresenta em escala log-log o tempo de execução (relógio de parede) do simulador em função do tamanho da rede, n para diferentes redes e para um tempo fixo, $T = 10000$, de simulação. Repare que o tempo de execução é praticamente constante com o aumento do tamanho da rede e que também não depende da estrutura da mesma. Este mesmo comportamento pode ser observado para $K = 5$ e $K = 50$ indivíduos, mostrando que isto não influencia a escalabilidade do simulador quanto ao tamanho da estrutura da rede. Desta forma, os resultados empíricos estão de acordo com a teórica da complexidade do simulador.

Repare que a complexidade de cada iteração do laço principal não depende do tempo de simulação, de forma que a complexidade do simulador seja linear no tempo de simulação (dado todos os outros parâmetros do modelo). Entretanto, o número de eventos processados por unidade de tempo de simulação depende dos parâmetros do modelo, tais como taxa de caminhar, taxa de recuperar, número de indivíduos e tamanho da rede. Ou seja, o tempo de execução necessário para simular um tempo T de simulação depende dos parâmetros do modelo, mas dado os parâmetros, este tempo de execução é linear em T .

A figura 5.2 apresenta o tempo de execução (tempo de relógio de parede) do simulador em função do tempo de simulação (T) para os três tipos de rede. Simulamos a epidemia com 5 e 50 indivíduos, fazendo 30 rodadas e obtendo o tempo

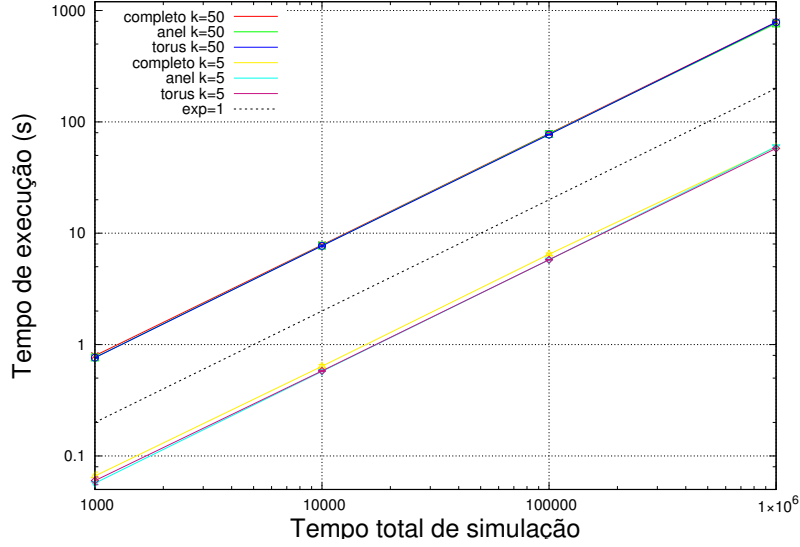


Figura 5.2: Tempo execução em função do tempo de simulação, em escala log-log. Parâmetros: $n = 25$, $\lambda = 1$, $\gamma = 1$, $\tau = 1$, $T = 1000$

médio de execução. A figura mostra que o desempenho do simulador é linear no tempo de simulação, corroborando a análise teórica. Repare que em escala log-log, a reta com inclinação 1 corresponde a uma relação linear na escala convencional, e esta é a inclinação das retas ilustradas na figura. Os resultados também indicam que este tempo não depende da estrutura da rede, pois temos retas praticamente idênticas para as três redes. Entretanto, repare que o cenário com 50 indivíduos demanda praticamente 10 vezes mais tempo de execução do que com 5 indivíduos. Este comportamento também é esperado, pois teremos aproximadamente 10 vezes mais eventos por unidade de tempo de simulação.

A Figura 5.3 apresenta o tempo de execução do simulador em função do número de indivíduos, para as três redes avaliadas. Nota-se que, para K pequeno (menor que 1000), o simulador apresenta uma escalabilidade praticamente linear (reta com inclinação 1 na escala log-log). Entretanto, a medida que K aumenta (para valores maiores que 1000), a inclinação da curva representando o tempo de execução também aumenta. Este comportamento está de acordo com a análise teórica de pior caso, que tem complexidade $\mathcal{O}(K \log K)$. Repare que para valores grandes de K , o fator $\log K$ faz diferença. Novamente, observamos que o desempenho do simulador não depende da estrutura da rede, já que as curvas são praticamente idênticas para as três redes.

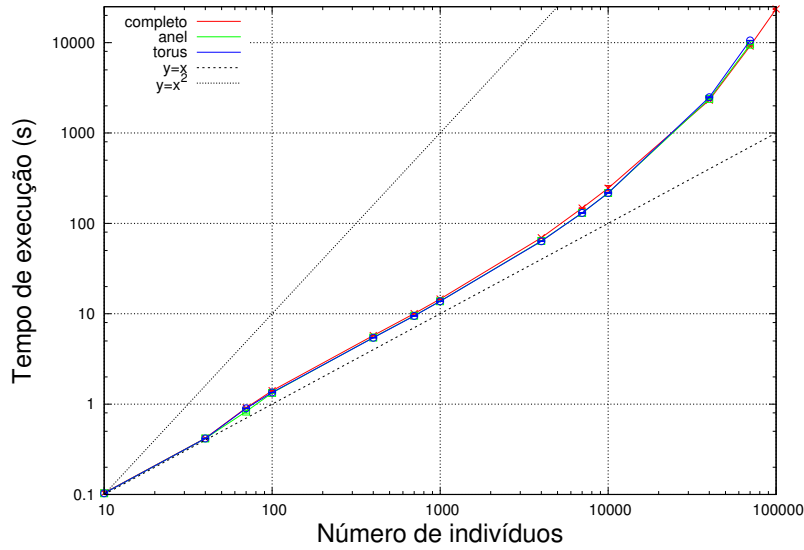


Figura 5.3: Tempo de execução em função dos passeios aleatórios, em escala log-log. Parâmetros: $T = 1000$, $n = 50$, $\lambda = 1$, $\gamma = 1$, $\tau = 1$

5.2 Caracterização da epidemia no tempo

No modelo epidêmico proposto, indivíduos movimentam-se de acordo com passeios aleatórios independentes com taxa λ . Outro parâmetro que influencia diretamente o comportamento da epidemia é a taxa de recuperação dos indivíduos, γ . Repare que a relação λ/γ representa o número médio de localidades que um indivíduo visita enquanto infectado. Um indivíduo que caminha muito mais rápido que o tempo necessário para se recupera ele irá levar a epidemia para mais vértices da rede.

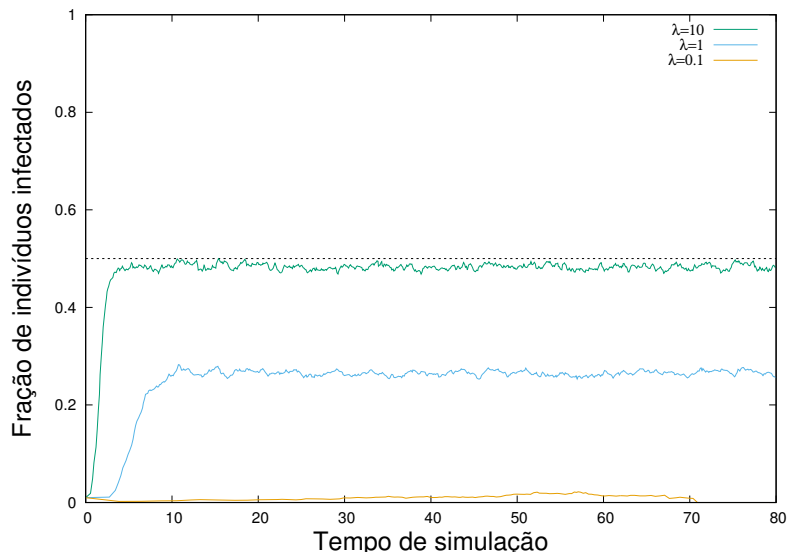


Figura 5.4: Fração de indivíduos infectados ao longo do tempo para diferentes taxas de caminhar, λ . Parâmetros: $n = 100$, $K = 150$, $k_i = 2$, $\gamma = 1$, $\tau = 1$.

Para ilustrar a importância da mobilidade na propagação da epidemia, a figura 5.4 exibe a fração de indivíduos infectados ao longo do tempo para diferentes taxas de caminhar, λ . O número de indivíduos é $K = 150$ com uma população inicial de infectados $k_i = 2$ que equivale a aproximadamente 1% da população total. A simulação foi feita para o grafo completo com $n = 100$ e a figura representa a média de 30 simulações. Além disso a taxa de recuperar é $\gamma = 1$. Assim, a quantidade de vértices (localidades) que um indivíduo visita enquanto está infectado é $\lambda/\gamma = \lambda$.

Na figura 5.4, para $\lambda = 10$ a fração de indivíduos infectados cresce mais rápido que as demais curvas e desacelera quando chega perto do estado endêmico (assim como todas as curvas). Repare que o indivíduo visita 10 localidades enquanto infectado, podendo contagiar bem mais indivíduos.

Por outro lado para $\lambda = 0.1$ a epidemia termina quando, $t = 70$, ou seja, não há mais indivíduos infectados. Isso ocorre porque $\lambda/\gamma < 1$, ou seja, o indivíduo não se locomove enquanto infectado, podendo infectar apenas indivíduos que entram ou estão no mesmo vértice quando foi infectado.

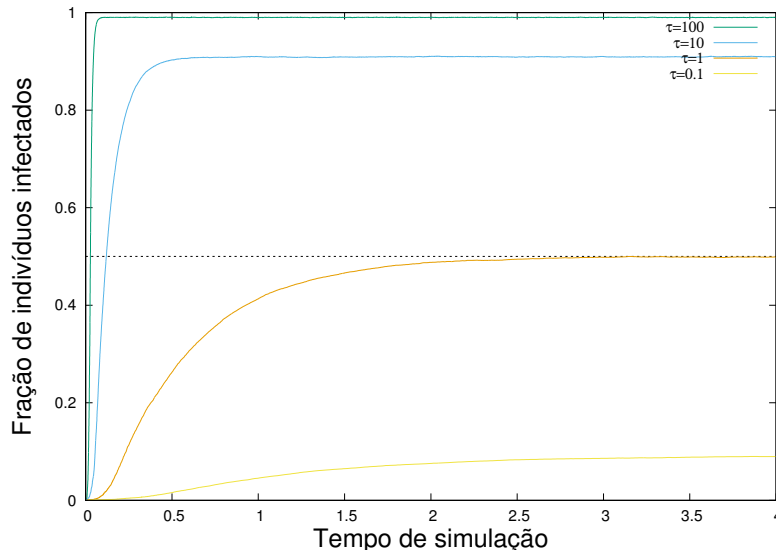


Figura 5.5: Fração de indivíduos infectados ao longo do tempo para diferentes taxas de infectar, τ . Parâmetros: $n = 100, K = 10000, \lambda = 5, \gamma = 1$

Outro parâmetro do modelo que podemos estudar é o tempo médio de incubação $1/\tau$ antes da infecção que determina o tempo que o indivíduo passa no estado contraído até ir para o estado infectado. A figura 5.5 mostra a evolução da fração de infectados ao longo do tempo. Cada curva representa uma taxa de infectar τ diferente. Os outros parâmetros são: $\lambda = 5$ e $\gamma = 1$. Como podemos observar, quanto maior τ , menos tempo o indivíduo fica no estado contraído passando direto para o estado infectado, aumentando assim a fração de indivíduos infectados.

Repare que o parâmetro τ influencia o estado endêmico da epidemia, aumen-

tando a fração de indivíduos infectados quando τ aumenta. Isto ocorre porque os indivíduos passam menos tempo no estado contraído diminuindo assim a fração de indivíduos neste estado. Além de influenciar o estado endêmico, τ também desacelera a propagação da epidemia quando pequeno ($\tau = 0.1$). Isso ocorre porque na maior parte do tempo que o indivíduo caminha pela rede ele está no estado contraído, impossibilitado de disseminar a infecção. Isso mostra que o estado contraído funciona como uma barreira, impedindo que a epidemia se alastre rapidamente pela rede.

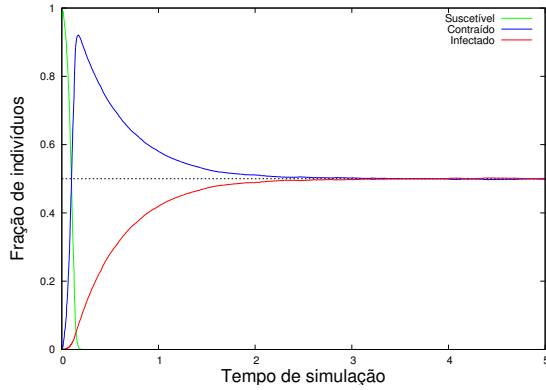


Figura 5.6: Fração de indivíduos em cada estado epidêmico ao longo do tempo para $\lambda = 10$, $\gamma = 1$, $\tau = 1$, $n = 100$ e $K = 10000$.

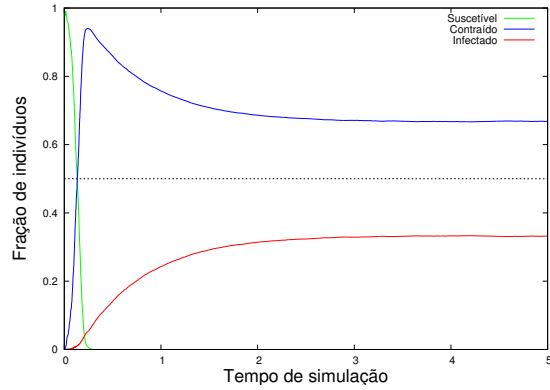


Figura 5.7: Fração de indivíduos em cada estado epidêmico ao longo do tempo para $\lambda = 10$, $\gamma = 1$, $\tau = 0.5$, $n = 100$ e $K = 10000$.

As figuras 5.6 e 5.7 mostram a fração de indivíduos em cada estado epidêmico ao longo do tempo. Nos dois casos, a fração de indivíduos no estado suscetível vai rapidamente a zero pois a densidade de indivíduos por vértice é muito alta ($n = 100$ e $K = 10000$) gerando muitos encontros. Desta forma, quando o indivíduo se recupera, ele volta, no mesmo instante de tempo, para o estado contraído. Repare que na figura 5.6 temos que $\lambda = \gamma = 1$ levando a fração de indivíduos infectados ser igual a fração de indivíduos no estado contraído. Por outro lado, na figura 5.7 $\lambda = 2\tau$ e por isso vemos uma fração maior de indivíduos no estado contraído.

Também é possível notar que a fração de indivíduos no estado contraído cresce rapidamente antes de alcançar o estado endêmico. Isso ocorre porque no início da epidemia quando os indivíduos são infectados eles passam primeiro pelo estado contraído para depois mudar para o estado infectado.

A propagação de uma epidemia é determinada pelo mecanismo de transmissão do patógeno e também pela estrutura da rede na qual a mesma se propaga. Para ilustrar, simulamos a epidemia para o grafo em anel e grafo completo, figura 5.8 e figura 5.9, respectivamente. Repare que o estado endêmico da epidemia nos dois casos é o mesmo. Entretanto a fração de indivíduos suscetíveis decresce muito mais rápido para o grafo completo que no anel, levando quatro vezes mais tempo para ir

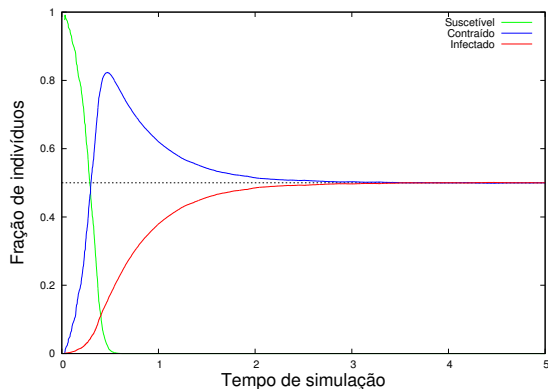


Figura 5.8: Fração de indivíduos em cada estado epidêmico ao longo do tempo para o grafo completo ($\lambda = 1$, $\gamma = 1$, $\tau = 1$, $n = 100$ e $K = 10000$).

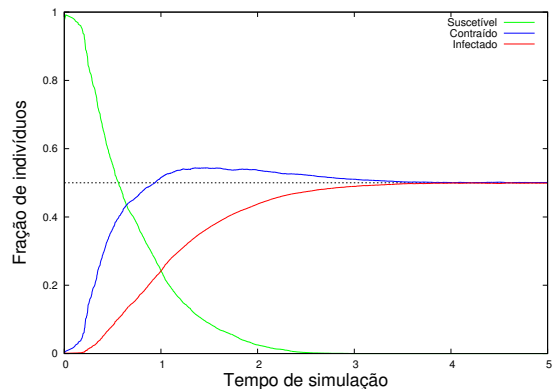


Figura 5.9: Fração de indivíduos em cada estado epidêmico ao longo do tempo para o grafo em anel ($\lambda = 1$, $\gamma = 1$, $\tau = 1$, $n = 100$ e $K = 10000$).

a zero no anel ($t = 0.5$ no grafo completo e $t = 2$ no anel). Além disso, a fração de indivíduos contraídos no grafo completo possui um pico acima de 80%, o que não ocorre no grafo em anel. A diferença na propagação da epidemia nestas redes ocorre porque o grafo completo tem muito mais arestas, permitindo mais encontros na rede a medida que os indivíduos caminham.

O comportamento da epidemia depende também do tamanho da rede, pois quanto maior for a rede, mais os indivíduos podem se espalhar, diminuindo a quantidade de encontros e por conseguinte o número de infecções.

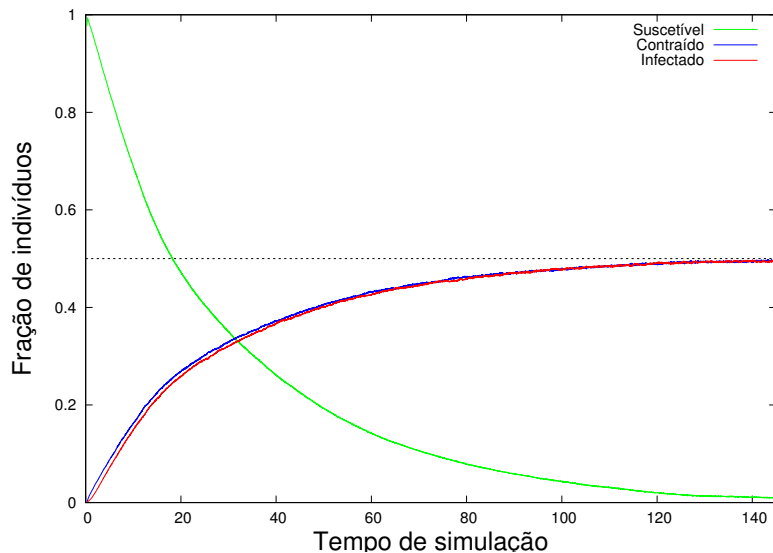


Figura 5.10: Fração de indivíduos em cada estado epidêmico ao longo do tempo para o grafo anel ($\lambda = 1$, $\gamma = 1$, $\tau = 1$, $n = 1000$ e $K = 10000$).

A figura 5.10 apresenta a fração de indivíduos em cada estado epidêmico para o grafo anel com 1000 vértices. Note que a fração de indivíduos suscetíveis vai a zero

mais lentamente do que na figura 5.9 (onde $n = 1000$). Quando o número de vértices aumentou 10 vezes (de $n = 100$ para $n = 1000$), o tempo que a epidemia leva para chegar ao estado endêmico aumentou mais de 50 vezes. Repare ainda que as curvas da fração de contraídos e infectados são bem parecidas e que em certo momento ($t = 30$) o sistema teve a mesma fração de indivíduos em cada estado epidêmico.

Estes experimentos demonstram que o comportamento da epidemia neste modelo é sensível a muitos fatores, e em particular as taxas de infecção e recuperação, e a estrutura da rede. Esta sensibilidade também foi observada nos modelos clássicos de epidemia em redes, onde indivíduos não se movimentam.

5.3 Transição de fase no tempo de simulação da epidemia

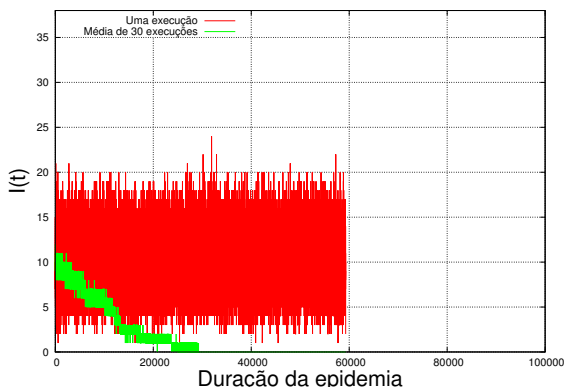


Figura 5.11: Número de indivíduos infectados em função do tempo de simulação, para epidemia de curta duração. Parâmetros: $n = 30, K = 38, \lambda = 1, \gamma = 1, \tau = 1, T = 100000$.

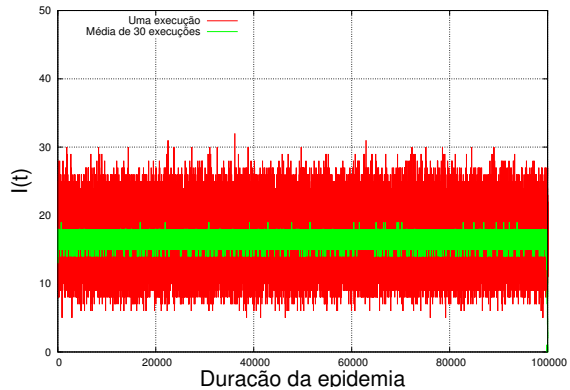


Figura 5.12: Número de indivíduos infectados em função do tempo de simulação, para epidemia de longa duração. Parâmetros: $n = 30, K = 50, \lambda = 1, \gamma = 1, \tau = 1, T = 100000$.

As figuras 5.11 e 5.12 apresentam o número de indivíduos infectados em função do tempo de simulação. A curva vermelha representa o número de infectados no tempo em uma execução de simulação, enquanto a curva a verde representa a média de indivíduos infectados no tempo para 30 rodadas de simulação. Repare que a curva vermelha oscila muito, indicando a grande variação do número de infectados no decorrer da epidemia. Entretanto, quando a curva vale zero, a epidemia para, pois não há mais novos contágios. As duas figuras representam simulações com os mesmos parâmetros, exceto pelo número de indivíduos ($K = 38$ e $K = 50$, respectivamente). Repare que na figura 5.11 a epidemia termina antes do final do tempo de simulação. Já na figura 5.12, a epidemia dura até o tempo limite de simulação. Este comportamento sugere uma *transição de fase* no parâmetro K ,

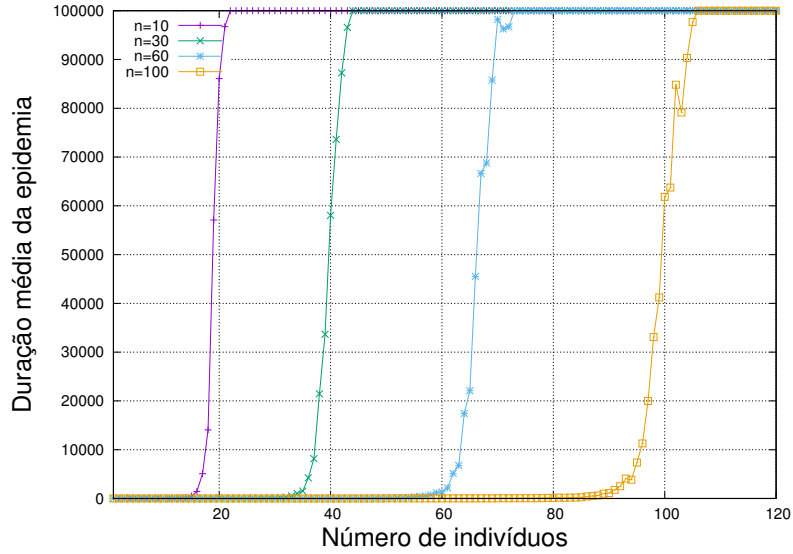


Figura 5.13: Transição de fase para o grafo completo. Parâmetros: $\lambda = 1, \gamma = 1, \tau = 1, T = 100000$.

que determina a duração da epidemia, ou seja, ou longa (até o final do período de simulação) ou curta (termina antes do tempo máximo de simulação).

Para ilustrar o comportamento de transição de fase da epidemia, realizamos experimentos com redes diferentes e de diferentes tamanhos. As figuras 5.13, 5.14, 5.15 apresentam o tempo médio de duração da epidemia em função do número de indivíduos na rede, para diferentes tamanhos de rede (diferentes curvas) e diferentes redes (figuras diferentes). Repare que em todos os tamanhos de rede e em todos os tipos de rede, observamos uma transição de fase ao aumentarmos o número de indivíduos na rede. Ou seja, temos epidemias relativamente curtas se K é menor que o ponto de transição, e relativamente longas (até o tempo máximo de simulação) se K é maior que o ponto de transição.

Para o grafo completo, a transição de fase ocorre quando $K > n$, mas a relação entre K e n diminui conforme n aumenta (figura 5.13). Por exemplo, quando $n = 10$ a transição de fase ocorre em torno de $K = 20$, sendo necessário 2 indivíduos por vértice da rede para que epidemias sejam longas. Entretanto, quando $n = 100$ são necessários de 110/100 indivíduos por vértice para que epidemias sejam longas. Este fenômeno ocorre pois em um grafo completo, o aumento do número de indivíduos gera mais possibilidades de encontro entre indivíduos suscetíveis e infectados do que o aumento do tamanho da rede.

A figura 5.14 exibe o tempo médio de duração da epidemia em função do número de indivíduos em um anel de diferentes tamanhos. O anel é uma estrutura onde as distâncias entre os vértices é $\mathcal{O}(n)$, uma vez que é necessário que os indivíduos se movimentem pelo anel. Isto faz com que os encontros sejam mais erráticos, podendo dois indivíduos permanecerem um longo tempo sem se encontrar. Desta

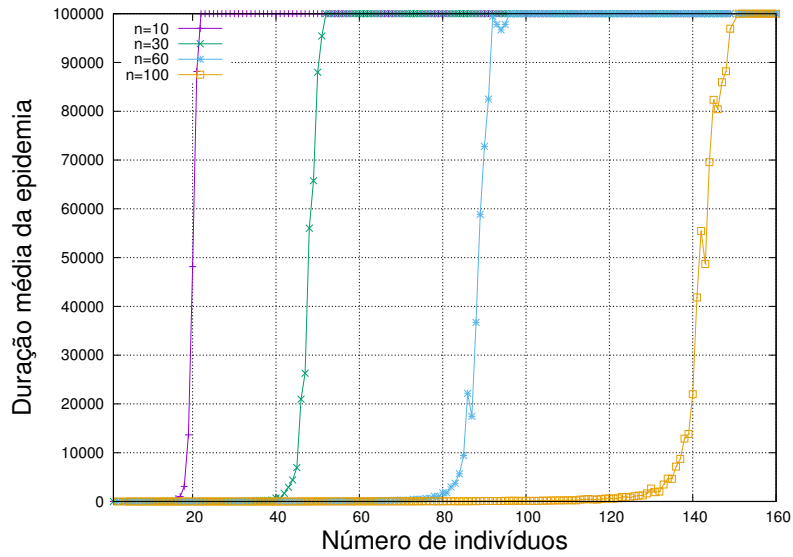


Figura 5.14: Transição de fase para o anel. Parâmetros: $\lambda = 1, \gamma = 1, \tau = 1, T = 100000$.

forma, a chance de um indivíduo suscetível encontrar um indivíduo infectado no anel é menor que no grafo completo. Como os encontros são mais difíceis, é preciso alta densidade de indivíduos para epidemias persistirem na rede. Por exemplo, com $n = 100$ são necessários ao menos $K = 150$ indivíduos para epidemias serem longas. Além disso, como podemos observar no gráfico, quando n aumenta, o número de indivíduos K onde ocorre a transição de fase também aumenta, diferentemente do grafo completo. Isso ocorre porque a distância média entre os vértices aumenta quando a rede aumenta demandando mais indivíduos para manter a epidemia em curso e não deixá-la terminar rapidamente.

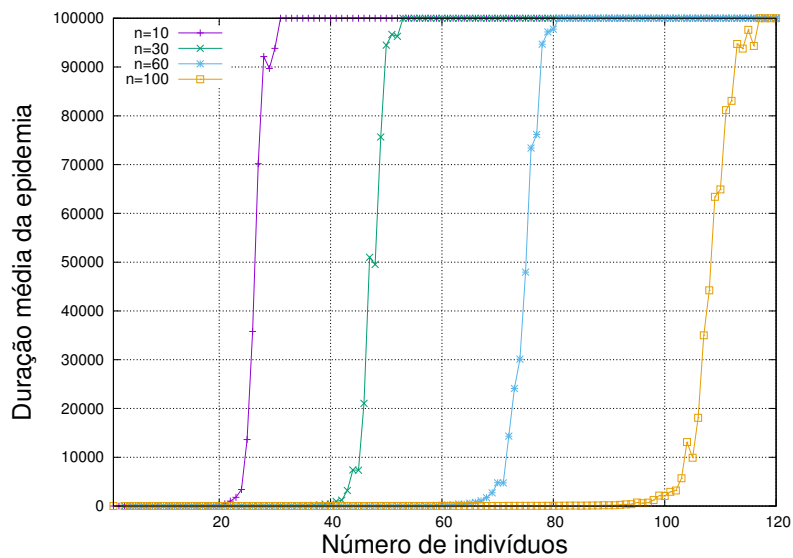


Figura 5.15: Transição de fase para o torus no parâmetro K . Parâmetros: $\lambda = 1, \gamma = 1, \tau = 1, T = 100000$.

A figura 5.15 apresenta o tempo médio de duração da epidemia em função número de indivíduos para torus de diferentes tamanhos. O torus é uma rede onde todos os vértices possuem 4 vizinhos, tendo a aparência de um *rosca*. Conforme observamos, a transição de fase ocorre quando $K > n$. Assim como o anel, a distância média entre os vértices do torus também aumenta quando a rede aumenta (distância média da ordem de $\mathcal{O}(\sqrt{n})$) e por isso a relação entre K e n não diminui conforme no grafo completo. A tendência é que a relação entre K e n aumente a medida que n aumente, conforme o anel, para torus muito grandes.

Outro fator que influencia o comportamento da epidemia é a relação entre a taxa de caminhar λ e a taxa de recuperar γ principalmente porque a fração λ/γ determina a o número médio de saltos realizados por um indivíduo infectado. Vimos anteriormente que quanto maior λ em relação a γ mais rápido a epidemia se espalha pela rede. Estamos interessados na possível transição de fase da epidemia, assim como visto para K .

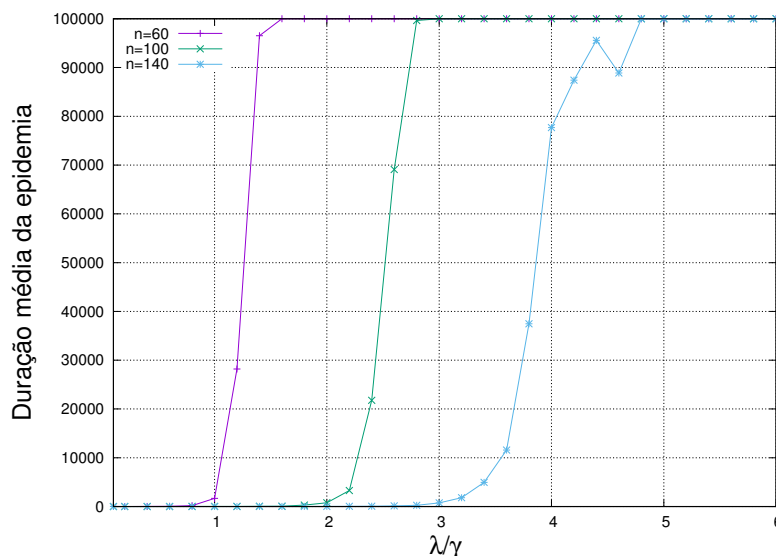


Figura 5.16: Transição de fase para o grafo completo na relação λ/γ . Parâmetros: $K = 60, \tau = 1, T = 100000$.

A figura 5.16 mostra a duração média da epidemia em função de λ/γ para o grafo completo de diferentes tamanhos ($n = 60, n = 100$ e $n = 140$, para $K = 60$). Repare que para $K = 60$, em K todos os tamanhos da rede estão antes da transição de fase, como mostra a figura 5.13, ou seja, estes resultados não são influenciados pela transição de fase em K . Note que a transição de fase em λ/γ ocorre para todos os tamanhos da rede, mas a medida que aumentamos n , a fração λ/γ também aumenta já que os indivíduos estão mais espalhados no grafo completo e um indivíduo infectado deve percorrer uma distância maior para encontrar um indivíduo suscetível. Note também que a inclinação da curva no ponto de transição de fase é menor para n maior, indicando que a transição de fase é menos sensível

aos valores críticos de λ/γ .

No anel, ao contrário do grafo completo, as distâncias entre os vértices crescem com o número de vértices, fazendo com que os indivíduos tenham que percorrer mais vértices até se encontrarem. A figura 5.17 mostra o tempo médio de duração da epidemia, para 30 rodadas, em função da relação λ/γ . Assim como no grafo completo, a medida que λ/γ aumenta existe uma transição de fase levando a epidemia a persistir na rede até o tempo máximo de simulação. Diferentemente do grafo completo, os indivíduos precisam percorrer mais vértices para a epidemia entrar no estado endêmico, mais precisamente, no grafo completo para $n = 100$ a transição de fase ocorre em $\lambda/\gamma \approx 3$ enquanto para o grafo em anel em $\lambda/\gamma = 6,5$, pouco mais que o dobro.

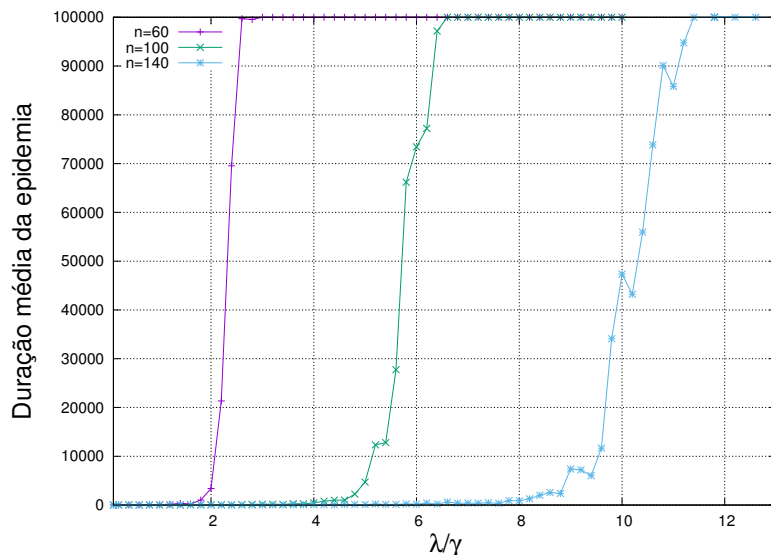


Figura 5.17: Transição de fase para o anel na relação λ/γ . Parâmetros: $K = 60$, $\tau = 1$, $T = 100000$.

Vale ressaltar que para $n = 100$ e $n = 140$, o valor de $K = 60$ está bem abaixo da transição de fase no parâmetro K , como mostra a figura 5.14. Isso mostra que embora a rede seja esparsa (com poucos indivíduos), a alta mobilidade dos indivíduos pode determinar a duração da epidemia, que depende igualmente da estrutura da rede.

Já para o torus, ilustrado na 5.18, a transição de fase ocorre conforme o esperado, com o aumento de λ/γ , e com valores intermediários entre o grafo completo e o anel, semelhante aos experimentos de transição de fase para o parâmetro K . Isto é esperado já que um vértice no torus possui menos vizinhos que no grafo completo e mais que no anel.

Esta avaliação da transição de fase entre epidemias curtas e longas indica a importância da estrutura da rede no processo epidêmico. Apesar da transição de fase ocorrer em todas as redes avaliadas, o ponto de transição depende fundamentalmente

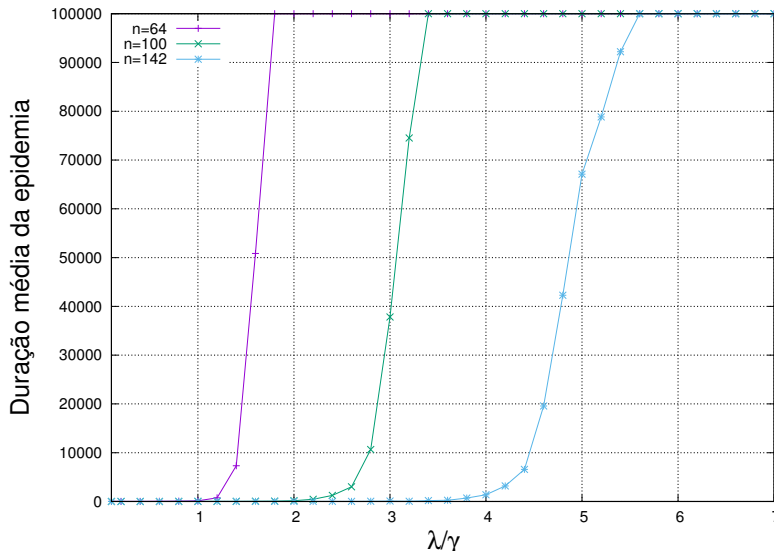


Figura 5.18: Transição de fase para o torus na relação λ/γ .
Parâmetros: $K = 60, \tau = 1, T = 100000$.

da rede. Em particular, a transição de fase em uma rede mais densa, como o grafo completo, ocorre antes (com menos indivíduos e λ/γ pequeno) que em uma rede mais esparsa, como o grafo em anel, independente do tamanho destas redes.

5.4 População inicial de infectados

Um parâmetro importante do modelo que determina o comportamento da epidemia perto do ponto de transição de fase é o número de indivíduos infectados no tempo t_0 . Este parâmetro também é importante para determinar a duração da epidemia pois se K for grande mas k_i pequeno demais, a epidemia pode terminar rapidamente mesmo que os parâmetros estejam após o ponto de transição de fase.

Para caracterizar o comportamento da epidemia em função k_i , simulamos o cenário de uma epidemia em grafos completos com diferentes tamanhos e valores $k_i = 0.1, k_i = 0.5$ e $k_i = 0.9$ (figura 5.19). Variamos K até a epidemia encontrar o ponto de transição de fase. Repare que para $k_i = 0.1$ o comportamento da epidemia é errático perto do ponto de transição de fase. Isso ocorre pois ainda que seja escolhido um K alto (após transição de fase), existe a chance da epidemia terminar rapidamente já que a mesma inicia com um percentual pequeno da população.

Para ilustrar melhor a oscilação da duração do tempo da epidemia quando próximo ao ponto de transição de fase, a figura 5.20 apresenta o desvio padrão da duração da epidemia. Repare como as curvas são oscilam para valores logo após o ponto de transição de fase e com valores maiores (maiores picos) para $k_i = 0.1$. Note que para $k_i = 0.5$ e $k_i = 0.9$ não existe muita diferença entre as curvas.

É importante notar que a transição de fase não depende do parâmetro k_i . De

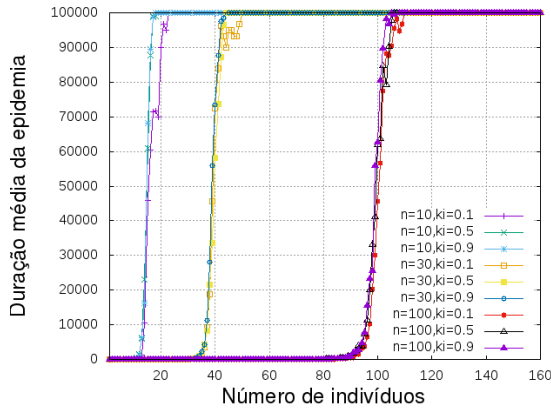


Figura 5.19: Transição de fase com diferentes valores de k_i para diferentes tamanhos de rede. Parâmetros $p = 0.5$, $\lambda = 1$, $\gamma = 1$, $\tau = 1$.

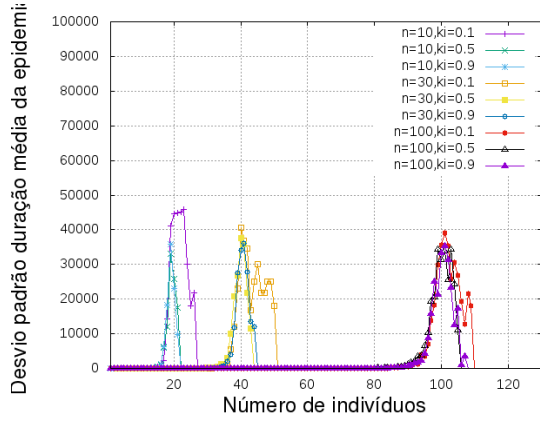


Figura 5.20: Desvio padrão para diferentes tamanhos de rede, n , da duração da epidemia variando k_i . Parâmetros $p = 0.5$, $\lambda = 1$, $\gamma = 1$, $\tau = 1$.

fato, a figura 5.19 ilustra que para valores pequenos de k_i a transição de fase ocorre um pouco depois e para valores maiores de k_i a transição de fase é mais rápida. Para uma população inicial pequena de indivíduos infectados, a epidemia pode terminar rapidamente caso os indivíduos se recuperem ou não consigam encontrar indivíduos suscetíveis. Porém, isto é fruto de oscilação estatística pois caso a epidemia se espalhe inicialmente, então ela vai persistir por muito tempo quando estivermos operando depois da transição de fase.

5.5 Validação

Assim como qualquer software, simuladores são passíveis de erros e por isso precisam ser criteriosamente verificados e validados. O simulador desenvolvido possui um módulo para verificar que seu comportamento está adequado. A partir dos dados de logs gerados pelo simulador, o módulo de validação compara os dados experimentais com os valores teóricos. O simulador funciona com diversos parâmetros e simula diferentes cenários, mas para fazer a validação podemos escolher um caso especial onde podemos facilmente comparar os dados gerados pelo simulador com os valores esperados.

Simulamos uma epidemia com 40 indivíduos em um grafo completo de 25 vértices. Por simplicidade, escolhemos os seguintes parâmetros para as exponenciais: $\lambda = \gamma = \tau = 1$. A probabilidade de infecção nos vértices é igual para todos, $p = 0.5$. O tempo de simulação limite é $T = 100000$, permitindo gerar bastante dados. Neste módulo, o simulador coleta os valores gerados pelas exponenciais calculando então a média amostral, desvio padrão amostral e mediana amostral. Neste cenário comparamos valores amostrais com os valores teóricos para os eventos caminhar,

recuperar e infectar ilustrados na tabela 5.1. Lembrando que para uma variável aleatória exponencial com parâmetro λ , a média e o desvio padrão são $1/\lambda$, e a mediana é $\ln 2/\lambda$. Os valores indicam que o comportamento dos tempos destes eventos está dentro do esperado. A tabela 5.2 apresenta a mesma comparação para valores distintos de $\lambda = 2$, $\gamma = 0.5$ e $\tau = 3$.

Tabela 5.1: Comparação entre os dados gerados pelo simulador e valores teóricos das exponenciais para diferentes eventos do simulador ($n = 35, K = 40, \lambda = 1, \gamma = 1, \tau = 1, T = 100000$).

	Caminhar		Recuperar		Infectar	
	Teórico	Amostral	Teórico	Amostral	Teórico	Amostral
Média	1	1.00355	1	0.998223	1	1.00343
Desvio Padrão	1	1.00387	1	1.00226	1	1.00903
Mediana	0.693147	0.697129	0.693147	0.69028	0.693147	0.690399

Tabela 5.2: Comparação entre os dados gerados pelo simulador e valores teóricos das exponenciais para diferentes eventos do simulador ($n = 35, K = 40, \lambda = 2, \gamma = 0.5, \tau = 3, T = 100000$).

	Caminhar		Recuperar		Infectar	
	Teórico	Amostral	Teórico	Amostral	Teórico	Amostral
Média	0.5	0.501733	2	2.0109	0.33333	0.333999
Desvio Padrão	0.5	0.501261	2	2.00788	0.33333	0.333394
Mediana	0.34657	0.346574	1.38629	1.39795	0.23104	0.232659

A partir dos dados gerados pelo simulador, o módulo de validação compara graficamente os tempos destes eventos. Para isso, calculamos a probabilidade empírica de um indivíduo permanecer um tempo $T \geq t$ (CCDF) em um vértice. A figura 5.21 compara os dados amostrais gerados pelo simulador e a CCDF da exponencial com parâmetro 1 para o evento caminhar. Repare que as curvas são praticamente idênticas até valores de $t = 6$, depois disso nota-se uma oscilação na curva amostral devido a pouca quantidade de amostras, já que a probabilidade de um indivíduo permanecer um tempo maior ou igual a 6 no vértice é $P[t \geq 6] \approx 0.001$.

A figura 5.22 apresenta a mesma comparação de CCDF para o evento recuperar que determina o tempo que o indivíduo permanece no estado epidêmico infectado. Repare que conforme a CCDF para o evento caminhar, as curvas são praticamente idênticas até $P[T \geq 6]$.

A mesma comparação para o evento infectar, que determina o tempo que o indivíduo permanece no estado epidêmico contraído, está ilustrada na figura 5.23. Repare que as três figuras apresentam curvas semelhantes, o que era esperado já que as três são variáveis aleatórias com o mesmo parâmetro.

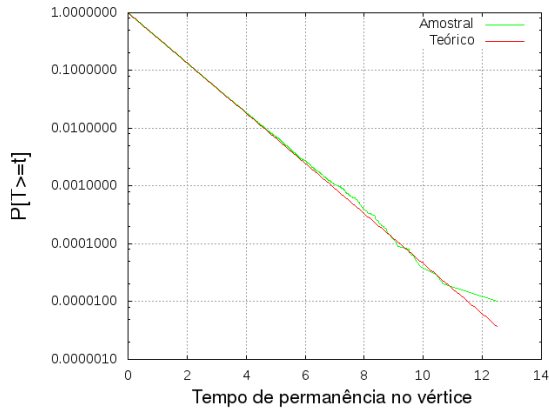


Figura 5.21: Probabilidade empírica e teórica de um indivíduo permanecer um tempo $T \geq t$ em um vértice. $\lambda = 1$. ($n = 25, K = 40, \gamma = 1, \tau = 1$)

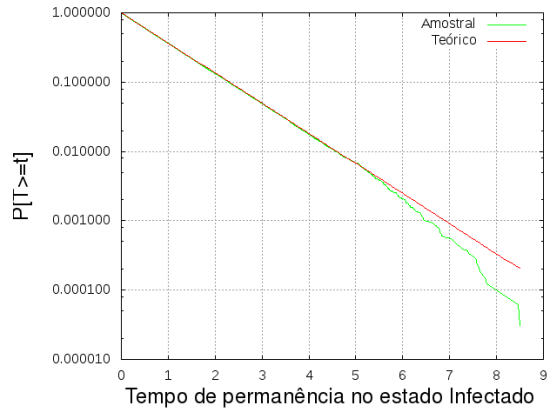


Figura 5.22: Probabilidade empírica e teórica de um indivíduo permanecer um tempo $T \geq t$ no estado infectado, $\gamma = 1$. ($n = 25, K = 40, \lambda = 1, \tau = 1$)

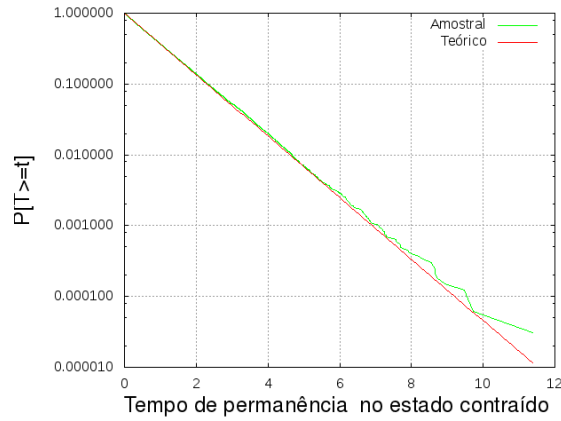


Figura 5.23: Probabilidade empírica e teórica de um indivíduo permanecer um tempo $T \geq t$ no estado contraído, $\tau = 1$. ($n = 25, K = 40, \lambda = 1, \gamma = 1$)

Tabela 5.3: Comparação entre o valor amostral e esperado do parâmetro p ($n = 35, K = 40, \lambda = 1, \gamma = 1, \tau = 1, T = 100000$).

	Esperado	Amostral
p_u	0.1	0.101462
p_v	0.5	0.498201
p_w	0.9	0.900134

Outro parâmetro validado pelo simulador é a probabilidade de infecção p associada a cada vértice da rede, dada por uma variável aleatória de Bernoulli no momento de um encontro do tipo S-I. Para validar esta variável aleatória, atribuímos probabilidades diferentes para três vértices, $p_u = 0.1$, $p_v = 0.5$ e $p_w = 0.9$ e comparamos estes valores teóricos com os resultados empíricos (fração relativa dos encontros SI que geram infecção) gerados pelo simulador, na tabela 5.3. Os resultados na tabela indicam que o simulador possui um comportamento conforme o esperado.

Capítulo 6

Conclusões e Trabalhos Futuros

6.1 Conclusão

Fenômenos de propagação em redes tem sido amplamente estudados nos últimos anos e para contribuir, propomos um modelo epidêmico em redes que diferente da abordagem tradicional, onde os indivíduos são fixos (representados pelos vértices), eles movimentam-se pela rede executando passeios aleatórios independentes. Generalizamos o modelo SIS para adicionar o estado contraído representando o período que o indivíduo esta infectado mas não pode disseminar a epidemia.

Como principal contribuição de nosso trabalho, foi projetado e implementado um simulador de eventos discretos para simular epidemias SCIS. Uma análise teórica e experimental do tempo de execução do simulador mostra que o mesmo é escalável em relação ao tamanho da rede (n), tempo de simulação (T) e número de indivíduos (K). Vimos que o tempo de execução do simulador é constante no tamanho da rede, $\mathcal{O}(1)$, linear no tempo de simulação, $\mathcal{O}(T)$, dado os parâmetros do modelo, e que cresce como $\mathcal{O}(K \log K)$ com o número de indivíduos K . Esta análise teórica de pior caso foi corroborada com o tempo real (relógio de parede) de execução do simulador em diferentes cenários.

O simulador também foi utilizado para caracterizar o comportamento da epidemia ao longo do tempo. Os resultados mostram que a propagação da epidemia depende dos parâmetros do modelo (λ , γ e τ) e também da estrutura da rede. Em particular, vimos que quando a taxa de caminhar aumenta, a epidemia se alastra mais rapidamente já que os indivíduos alcançam mais localidades enquanto infectados.

A avaliação empírica indicou uma transição de fase na duração da epidemia em função do número de indivíduos na rede e da relação λ/γ . Em particular, quando o número de indivíduos é pequeno o suficiente, epidemias tem curta duração. Por outro lado, se este número for grande o suficiente, epidemias tem longa duração

(duram até o final do tempo de simulação). Este comportamento foi encontrado em todas as redes avaliadas, e em diferentes tamanhos de rede. Entretanto, o ponto de transição de fase depende fundamentalmente da estrutura da rede, realçando a importância da estrutura para a caracterização do processo epidêmico. Já para λ/γ , a transição de fase ocorre quando a diferença λ e γ aumenta. A transição de fase em λ/γ também foi diferente para os três tipos de rede, em particular, para o anel, o indivíduo precisou caminhar muito mais, enquanto infectado, que no grafo completo para a epidemia ter longa duração.

6.2 Trabalhos Futuros

Neste trabalho, quando um indivíduo suscetível encontra outro infectado, ele pode ser contaminado de acordo com uma probabilidade associada ao vértice do encontro. Neste abordagem, não capturamos a influência do tempo de exposição do indivíduo suscetível ao patógeno. Para tal, pode-se modelar a transmissão da infecção utilizando taxas conforme feito para o evento infectar. Além disso, pode-se definir um fator de transmissão associado a cada indivíduo, representando a susceptibilidade de cada indivíduo ao patógeno.

Vimos que a relação λ/γ determina quantos vértices um indivíduo visita enquanto infectado e que influencia no comportamento da epidemia. No entanto, gostaríamos de saber se mantendo a proporção e variando os valores de λ e γ a epidemia terá o mesmo comportamento e se pode variar de acordo com a estrutura.

Referências Bibliográficas

- [1] BARABÁSI, A.-L. “Spreading Phenomena”. In: *Network Science*, 1 ed., cap. 10, Cambridge University Press, 2016.
- [2] BERNOULLI, D. “Essai d’une nouvelle analyse de la mortalite causee par la petite verole”, *Mem. Math. Phys. Acad. Roy. Sci.*, 1766.
- [3] DIETZ, K., HEESTERBEEK, J. “Daniel Bernoulli’s epidemiological model revisited”, *Mathematical Biosciences*, v. 180, pp. 1–21, dez. 2002.
- [4] KERMACK, W. O., MCKENDRICK, A. G. “A contribution to the mathematical theory of epidemics”, *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, v. 115, n. 772, pp. 700–721, 1927. ISSN: 0950-1207. doi: 10.1098/rspa.1927.0118.
- [5] EUBANK, S., KUMAR, V. S. A., MARATHE, M. V., et al. “Modelling disease outbreaks in realistic urban social networks”, *Nature*, v. 429, pp. 180–184, maio 2004.
- [6] CHAO, D. L., HALLORAN, M. E., OBENCHAIN, V. J., et al. “FluTE, a Publicly Available Stochastic Influenza Epidemic Simulation Model”, *PLOS Computational Biology*, v. 6, n. 1, pp. 1–8, 01 2010. doi: 10.1371/journal.pcbi.1000656.
- [7] PASTOR-SATORRAS, R., VESPIGNANI, A. “Epidemic dynamics and endemic states in complex networks”, *Physical Review*, v. 63, jun 2001.
- [8] MORENO, Y., PASTOR-SATORRAS, R., VESPIGNANI, A. “Epidemic outbreaks in complex heterogeneous networks”, *The European Physical Journal B - Condensed Matter and Complex Systems*, v. 26, pp. 521–529, abr. 2002.
- [9] PASTOR-SATORRAS, R., VESPIGNANI, A. “Epidemic Spreading in Scale-Free Networks”, *Physical Review Letters*, v. 86, n. 14, pp. 3200–3203, abr. 2001.

- [10] DRAIEF, M., GANESH, A. “A random walk model for infection on graphs: spread of epidemics and rumours with mobile agents”, *Discrete Event Dynamic Systems*, v. 21, pp. 41–61, mar 2011.
- [11] ZHOU, J., LIU, Z. “Epidemic spreading in communities with mobile agents”, *Physica A: Statistical Mechanics and its Applications*, v. 388, pp. 1228–1236, apr 2009.
- [12] ZHUGE, H. “Exploring an Epidemic in an E-Science Environment”, *Communications of the ACM*, v. 48, n. 9, pp. 521–529, set. 2005.
- [13] DOER, B., FOUZ, M., FRIEDRICH, T. “Why Rumors Spread So Quickly in Social Networks”, *Communications of the ACM*, v. 55, n. 6, pp. 70–75, jun. 2012.
- [14] BAUCH, C. T., LLOYD-SMITH, J. O., COFFEE, M. P., et al. “Dynamically Modeling SARS and Other Newly Emerging Respiratory Illnesses: Past, Present, and Future”, *Epidemiology*, v. 16, pp. 791–801, nov 2005.
- [15] D., N. “The Metropole, superspreaders, and other mysteries”, *Science*, pp. 1272–1273, mar. 2013.
- [16] BOGUÑÁ, M., PASTOR-SATORRAS, R. “Epidemic spreading in correlated complex networks”, *Phys. Rev. E*, v. 66, pp. 047104, Oct 2002. doi: 10.1103/PhysRevE.66.047104.
- [17] GANESH, A., MASSOULIE, L., TOWSLEY, D. “The effect of network topology on the spread of epidemics”. In: *Proceedings IEEE 24th Annual Joint Conference of the IEEE Computer and Communications Societies.*, v. 2, pp. 1455–1466, March 2005. doi: 10.1109/INFCOM.2005.1498374.
- [18] MARATHE, M., VULLIKANTI, K. S. “Computational Epidemiology”, *Communications of the ACM*, v. 56, n. 7, pp. 88–96, jul. 2013.
- [19] BROECK, W. V. D., GIOANNINI, C., GONÇALVES, B., et al. “The GLE-aMviz computational tool, a publicly available software to explore realistic epidemic spreading scenarios at the global scale”, *BMC Infectious Diseases*, v. 11, n. 1, pp. 37, Feb 2011. ISSN: 1471-2334. doi: 10.1186/1471-2334-11-37.
- [20] BARRETT, C. L., BISSET, K. R., EUBANK, S. G., et al. “Episimdemics: An efficient algorithm for simulating the spread of infectious disease over large realistic social networks”, *Proceedings of the ACM/IEEE Conference on High Performance Computing (SC)*, v. 37, nov 2008.

- [21] NSOESIE, E., MARARTHE, M., BROWNSTEIN, J. “Forecasting Peaks of Seasonal Influenza Epidemics”, *PLOS Currents*, jun 2013.