

**UNIOESTE - Universidade Estadual do Oeste do Paraná**  
**CENTRO DE CIÊNCIAS EXATAS E TECNOLÓGICAS**  
Colegiado de Informática  
*Curso de Bacharelado em Informática*

**Investigação de Modelos Epidemiológicos em Redes Complexas Estruturadas**

*Fabio Gausmann Köerich*

**CASCABEL**  
**2009**

**Fabio Gausmann Köerich**

**Investigação de Modelos Epidemiológicos em Redes Complexas Estruturadas**

Monografia apresentada como requisito parcial para obtenção do grau de Bacharel em Informática, do Centro de Ciências Exatas e Tecnológicas da Universidade Estadual do Oeste do Paraná - Campus de Cascavel

Orientador: Prof. Dr. Reginaldo Aparecido Zara

CASCADEL  
2009

**Fabio Gausmann Köerich**

**Investigação de Modelos Epidemiológicos em Redes Complexas Estruturadas**

Monografia apresentada como requisito parcial para obtenção do Título de Bacharel em Informática, pela Universidade Estadual do Oeste do Paraná, Campus de Cascavel, aprovada pela Comissão formada pelos professores:

---

Prof. Dr. Reginaldo A. Zara (Orientador)  
Colegiado de Informática, UNIOESTE

---

Prof. Dr. Rogério Luiz Rizzi  
Colegiado de Matemática, UNIOESTE

---

Prof. Msc. Guilherme Galante  
Colegiado de Informática, UNIOESTE

Cascavel, 8 de dezembro de 2009

## DEDICATÓRIA

*Dedico este trabalho a meu pai José Rogério Köerich e a minha mãe Sandra Maria Gausmann Köerich pelo apoio incondicional durante esses cinco anos, pois graça a isso, este trabalho pode ser realizado.*

*“Saiba sorrir para a vida a fim de que ela seja a sua própria alegria de viver. A partir daí, a felicidade estará permanentemente ao seu lado.”*

(Iran I. Jacob)

## **AGRADECIMENTOS**

Gostaria de agradecer de modo especial a minha família, minha mãe Sandra, meu pai Rogério, meu irmão Marcos e a minha namorada Patrícia pelo apoio incondicional dado nestes cinco anos de curso, sempre estando ao meu lado, me apoiando nos momentos mais difíceis e me dando forças pra continuar quando tudo parecia sem solução. Graças a esse apoio, hoje posso realizar mais esse feito que a cinco anos parecia inatingível.

Agradeço ao meus amigos(as) Osmar, Alexandre, Gustavo, Yuri, Dani, Eliana e Renata pelo companheirismo durante esses cinco anos, tanto em “festinhas” como as chamávamos, quanto nos trabalhos e tarefas das disciplinas. Além é claro dos que não ainda não chegaram até aqui ou desistiram pelo caminho e contribuíram para a minha formação.

Vale lembrar os formados, como o André Abe que me ajudou muito no projeto de Engenharia de Software, o Cassiano, Rafael, Adelar, Claudir, Marcelo, Thiago, entre outros, pelo apoio e também pelo futebol que jogávamos, pra esquecer um pouco dos trabalhos.

Obrigado também aos professores que auxiliaram na minha formação durante este período de permanência na universidade. O meu agradecimento ao professor Victor pela oportunidade na iniciação de projetos de pesquisa na área de Engenharia de Software.

Agradeço de forma especial ao professor Reginaldo, meu orientador nesse trabalho e com qual também realizei projeto de pesquisa, pelo apoio e compreensão e principalmente pelo seu grande auxílio no desenvolvimento deste trabalho.

O meu obrigado ainda as pessoas que acreditaram e acreditam do meu potencial pois graças a isso hoje consigo alcançar esse objetivo.

# Lista de Figuras

2.1	Representações de grafos valorados. . . . .	4
2.2	Distribuição das conectividades em um grafo aleatório. . . . .	6
2.3	Construção da rede sem-escala segundo o modelo de Barabási-Albert [1] . . . . .	7
2.4	Distribuição de conectividades rede sem-escala . . . . .	8
3.1	Exemplo de rede estruturada em comunidades com 2 comunidades, 3 ligações <i>intra-comunidade</i> e 2 ligações <i>inter-comunidade</i> . . . . .	10
3.2	Construção da rede sem-escala modular segundo Li e Maini [9]. . . . .	13
3.3	$Q$ para valores de $\alpha$ na rede modular segundo o modelo de Li-Maini . . . . .	13
3.4	Matriz de esparsidade para uma rede com $N = 1500$ elementos, $M = 8$ comunidades, $m = 4$ , $n = 7$ e diferentes valores para $\alpha$ . . . . .	14
3.5	Distribuição da Internet ao redor do mundo (a) Densidade mundial dos roteadores com maior nível de esforço e (b) Densidade populacional . . . . .	17
3.6	Construção da rede topológica estruturada em comunidades . . . . .	21
3.7	$Q$ para diferentes valores de $\alpha$ para rede modular espacial . . . . .	22
4.1	Diagrama de estados de diferentes Modelos Compartimentais. Dependendo do “caminho” escolhido, diferentes modelos podem ser formulados [2]. As letras gregas representam as taxas de transição entre os estados. . . . .	29
4.2	Simulação de uma epidemia com um modelo SIR determinístico. Os losangos representam a fração de indivíduos suscetíveis, os quadrados a fração de indivíduos infectados e os círculos a fração de indivíduos recuperados. . . . .	30
5.1	Exemplo de disposição espacial de $M$ comunidades em relação a matriz espacial da rede. . . . .	33

5.2	Exemplo de disposição espacial para $M = 2$ em relação a matriz espacial da rede.	33
5.3	Porção espacial de uma comunidade dividida em 4 faixas . . . . .	35
5.4	Informações sobre a porção espacial de uma comunidade . . . . .	35
5.5	Simulação do Modelo SIR em uma rede modular espacial . . . . .	37
5.6	Dinâmica da epidemia para redes com $N = 6000$ elementos, $M = 2$ comunidades, $m = 4$ , $n = 2$ e diferentes valores para $\alpha$ . . . . .	39
5.7	Espalhamento da epidemia nas faixas da comunidade <i>Infectada</i> para redes com $N = 6000$ elementos, $M = 2$ comunidades, $m = 4$ , $n = 2$ , $L = 100$ , $faixas = 5$ e diferentes valores para $\alpha$ para 10 repetições. O índice da faixa indica a distância em relação ao centro geométrico da comunidade. . . . .	41
5.8	Espalhamento da epidemia nas faixas da comunidade <i>Não Infectada</i> para redes com $N = 6000$ elementos, $M = 2$ comunidades, $m = 4$ , $n = 2$ , $L = 100$ , $faixas = 5$ e diferentes valores para $\alpha$ para 10 repetições. O índice da faixa indica a distância em relação ao centro geométrico da comunidade. . . . .	42
5.9	Espalhamento da epidemia nas faixas da comunidade <i>Infectada</i> para redes com $N = 6000$ elementos, $M = 2$ comunidades, $m = 4$ , $n = 2$ , $L = 100$ , $faixas = 5$ e diferentes valores para $\alpha$ para 1 repetição. O índice da faixa indica a distância em relação ao centro geométrico da comunidade. . . . .	43
5.10	Espalhamento da epidemia nas faixas da comunidade <i>Não Infectada</i> para redes com $N = 6000$ elementos, $M = 2$ comunidades, $m = 4$ , $n = 2$ , $L = 100$ , $faixas = 5$ e diferentes valores para $\alpha$ para 1 repetição. O índice da faixa indica a distância em relação ao centro geométrico da comunidade. . . . .	44



# Lista de Abreviaturas e Siglas

SIS	Suscetível-Infetado-Suscetível
SIR	Suscetível-Infetado-Removido
SIRS	Suscetível-Infetado-Removido-Suscetível
SEIRS	Suscetível-Latente-Infetado-Removido-Suscetível

# Lista de Símbolos

$k_i$	Numero de arestas que se conectam ao vértice $i$
$k_{in}(i)$	Numero de arestas que chegam ao vértice $i$
$k_{out}(i)$	Numero de arestas que partem do vértice $i$
$N$	Número de vértices de um grafo
$Q$	Fator de modularidade
$p$	Probabilidade de inserção de uma aresta entre dois vértices
$P(k)$	Numero de indivíduos que possuem $k$ ligações
$n_0$	Numero inicial de vértice no modelo de Barabási-Albert
$p_i$	Probabilidade de que o vértice $i$ receba uma conexão
$p_j$	Fração de arestas do vértice $j$
$a(t)$	Número de arestas existentes no grafo no passo de tempo $t$
$M$	Número de comunidade na rede
$\alpha$	Probabilidade de existência de ligações inter-comunidades
$m$	Número de arestas intra-comunidade de um vértice ao ser inserido
$n$	Número de arestas inter-comunidade de um vértice ao ser inserido
$(x, y)$	Posição espacial do indivíduo na rede
$d$	Distância Euclidiana entre dois vértices
$P(d)$	Probabilidade de conexão entre dois nós separados por uma distância $d$
$L$	Tamanho da superfície bidimensional linear
$d_{ij}$	Distância entre os vértices $i$ e $j$
$\sigma$	Expoente que rege a vinculação preferencial no modelo de Yook
$\chi$	Expoente que rege a distância nó-a-nó no modelo de Yook
$S$	Suscetível
$I$	Infectado ou Infeccioso
$E$	Latente
$R$	Removido
$\lambda, \gamma, \omega, \delta, \phi$	Taxas de transição entre os compartimentos
$R_0$	Número de reprodução basal
$\lambda_{min}$	Limiar de propagação da epidemia
$\lambda_{ef}$	Taxa de propagação efetiva
$faixas$	Número de faixas espaciais na rede modular espacial
$raio_a$	Raio de abrangência da comunidade $a$
$centro_a$	Centro geométrico da comunidade $a$
$\lambda_I$	Taxa de infecção de indivíduos de uma mesma comunidade
$\lambda_E$	Taxa de infecção de indivíduos de diferentes comunidades
$T_I$	Tempo de permanência de um indivíduo como infectado

# Sumário

<b>Lista de Figuras</b>	<b>vii</b>
<b>Lista de Abreviaturas e Siglas</b>	<b>ix</b>
<b>Lista de Símbolos</b>	<b>x</b>
<b>Sumário</b>	<b>xi</b>
<b>Resumo</b>	<b>xiii</b>
<b>1 Introdução</b>	<b>1</b>
<b>2 Conceitos</b>	<b>3</b>
2.1 Alguns Conceitos sobre Grafos . . . . .	4
2.2 Grafos Aleatórios . . . . .	5
2.3 Redes sem-escala . . . . .	6
<b>3 Redes modulares</b>	<b>9</b>
3.1 Redes estruturadas em comunidades . . . . .	9
3.2 Modelo de Li-Maini . . . . .	11
3.3 Redes modulares espaciais . . . . .	15
3.3.1 O Modelo de Yook . . . . .	15
3.3.2 Modelo de redes modulares espaciais . . . . .	20
<b>4 Modelos Epidemiológicos</b>	<b>24</b>
4.1 Conceitos básicos . . . . .	24
4.1.1 Classificação dos estados . . . . .	26
4.1.2 Transmissão . . . . .	26
4.1.3 O processo infeccioso . . . . .	27
4.1.4 Incidência e prevalência . . . . .	27
4.1.5 Imunidade . . . . .	27

4.1.6	Modelos compartimentais . . . . .	28
4.2	Modelo SIR . . . . .	28
<b>5</b>	<b>Análise do modelo SIR sobre redes modulares espaciais</b>	<b>32</b>
5.1	A disposição espacial dos indivíduos . . . . .	32
5.2	Distribuição espacial dos vértices . . . . .	34
5.3	O modelo SIR aplicado a rede modular espacial . . . . .	36
5.4	Resultados Obtidos . . . . .	38
5.4.1	Influência da Modularidade . . . . .	38
5.4.2	Comportamento em redes espaciais . . . . .	40
<b>6</b>	<b>Considerações Finais e Trabalhos Futuros</b>	<b>45</b>
	<b>Referências Bibliográficas</b>	<b>47</b>

# Resumo

As redes complexas podem representar muitos dos sistemas naturais nos quais o ser humano está inserido, bem como, a rede de contatos dos indivíduos que nela esteja alocados. Com isso, se torna possível a investigação de vários processos dinâmicos que envolvam a população humana, como por exemplo, a propagação de informações e doenças infecciosas entre os indivíduos pertencentes a população em estudo. Neste trabalho, o principal objetivo é investigar a propagação de epidemias em redes complexas sem escala com organização interna estruturada em comunidades, onde, a disposição espacial dos indivíduos é considerada. Estas redes também são denominadas redes modulares espaciais. As simulações foram realizadas sobre redes modulares espaciais construídas a partir da composição dos modelos de Li-Maini e Yook. Foram investigados o comportamento espaço-temporal da epidemia sobre as redes e a influência do fator de modularidade sobre a dinâmica da epidemia. Os resultados revelaram que uma estruturação interna bem definida da rede, ou seja, uma menor interação entre as comunidades, faz com que o processo epidêmico não possa ser avaliado pela análise da população como um todo, já que as diferentes comunidades apresentam picos de infecção distintos. No que diz respeito ao espalhamento espacial da epidemia a partir de uma comunidade inicialmente infectada, mesmo com o início da infecção nos indivíduos localizados na faixa central da comunidade o espalhamento para as outras faixas ocorre independentemente da distância a que o indivíduo esta em relação ao ponto inicial da infecção. Com isso, na comunidade que não estava infectada inicialmente, o comportamento da epidemia é semelhante em todas as porções da rede espacial, revelando que a disposição espacial dos indivíduos, para os casos simulados, não influenciou no espalhamento da epidemia a indivíduos da mesma comunidade e indivíduos de outras comunidades pertencentes a rede.

**Palavras-chave:** Redes complexas, Propagação de Epidemias, Epidemiologia Matemática.

# Capítulo 1

## Introdução

Muitos agrupamentos humanos e seus relacionamentos podem ser simulados por redes de interação as quais podem ser representadas por grafos. Nesses grafos cada um dos vértices simboliza um indivíduo e cada uma de suas arestas constitui uma ligação entre dois indivíduos. Estas ligações podem representar laços de amizade, parentesco, profissionais ou algum relacionamento de qualquer outro tipo [1].

Ao contrário dos grafos aleatórios nos quais as ligações entre os vértices são distribuídas de forma aproximadamente homogênea, em Redes sem-escala, ao se comparar a distribuição de conexões entre os vértices percebe-se que a maioria dos vértices da rede possui uma quantidade pequena de ligações, e uma quantidade pequena de vértices possui uma quantidade grande de ligações, o que a torna uma rede cuja distribuição de ligações ou conectividade é heterogênea [1].

Apesar desta característica de heterogeneidade global, estudos recentes verificaram que algumas redes complexas apresentam níveis de organização interna, com a formação de agrupamentos, onde certos grupos de indivíduos possuem uma grande quantidade de conexões com indivíduos pertencentes a um mesmo grupo, e uma pequena quantidade de ligações com outros indivíduos da rede. Na literatura atual diz-se que estes agrupamentos dos indivíduos formam redes reais com estruturação em comunidades [10][3].

A proposta deste trabalho é identificar os modelos de criação de redes complexas e investigar como a rede está estruturada internamente, para que posteriormente possa-se estudar como essa estruturação pode afetar o comportamento dos indivíduos em relação a fenômenos dinâmicos que possam ocorrer na rede. Os modelos investigados são apresentados nos próximos capítulos.

Foram selecionados dois modelos. O modelo de Li-Maini [9] permite que redes complexas

internamente estruturadas possam ser construídas a partir de um algoritmo fácil de ser implementado. O modelo de Yook [13] pode gerar redes complexas sem escala levando em conta a localização espacial dos indivíduos na rede. Estes dois modelos foram estudados e seus algoritmos foram implementados em linguagem C++.

No capítulo 2 são apresentados alguns conceitos sobre grafos que serão citados no decorrer do trabalho. No capítulo 3 inicialmente são apresentados os modelos utilizados e a caracterização das redes obtidas com base em seu fator de modularidade e posteriormente é apresentado o modelo de Yook generalizado, utilizando as ideias de Li-Maini, a fim de obter redes modulares espaciais. O capítulo 4 descreve como o comportamento de epidemias pode ser representado através dos modelos epidemiológicos, apresentando ainda alguns dos principais modelos atualmente utilizados, bem como, algumas características que os diferenciam. No capítulo 5 é apresentado o modelo epidêmico implementado sobre redes sem-escala organizadas em comunidades espacialmente distribuídos.

Por fim, o capítulo 6 apresenta as considerações finais deste trabalho e sugestões para trabalhos futuros.

# Capítulo 2

## Conceitos

Devido a sua natureza biológica e social, bem como, seu instinto de sobrevivência, os seres humanos vivem em agrupamentos. Esses agrupamentos e os relacionamentos por eles gerados, sejam eles laços de amizade, grau de parentesco ou de qualquer outro tipo, podem ser representados por redes denominadas redes de interação. [3] [1]

As redes de interação podem ser matematicamente representadas por grafos nos quais os vértices representam os indivíduos da rede e as arestas os relacionamentos entre os mesmos.

Nos grafos aleatórios todos os indivíduos da rede possuem uma quantidade similar de ligações, diferente do que ocorre nas chamadas redes complexas sem-escala, onde a maioria dos vértices possui poucas ligações e uma fração de vértices possui muitas ligações.

Essa diferença na distribuição de conectividade das ligações entre vértices faz com que as redes sem-escala sejam consideradas redes heterogêneas, enquanto que os grafos aleatórios são classificados como redes homogêneas pelo fato de seus vértices possuírem um grau de conectividade semelhante.

Além da heterogeneidade na distribuição das conexões, estudos recentes verificaram a existência de níveis de organização interna em algumas redes complexas, fazendo com que ocorra a formação de agrupamentos internos. Nesse agrupamentos um indivíduo pertencente a um grupo possui uma grande quantidade de conexões com indivíduos do mesmo grupo (relativamente ao total de ligações da rede), e uma quantidade pequena de ligações a indivíduos de outros grupos da rede. Esses grupos são similares a comunidades.

Como nesse trabalho serão trabalhados conceitos como conectividade, vértice e arestas, nas próximas seções são introduzidos (de maneira não exaustiva) alguns conceitos sobre grafos que serão citados no decorrer do trabalho.



## 2.1 Alguns Conceitos sobre Grafos

Um grafo  $G$  é definido por um tupla  $\{V, A\}$  onde  $V$  é o conjunto de vértices pertencentes ao grafo  $G$  e representa os indivíduos, objetos ou algo que se deseja representar pertencente a ele. Já o conjunto de arestas (ou ligações)  $A$  representa as relações entre os vértices pertencentes a  $G$ .

O conjunto  $V$  de  $N$  vértices de um grafo  $G$  pode ser representado por  $\{v_1, v_2, v_3, \dots, v_n\}$ . Cada elemento do conjunto de arestas  $A$  representa uma conexão (ou ligação) entre dois vértices, por exemplo, a ligação ou aresta de um vértice  $i$  a qualquer outro vértice  $j$  pode ser representada por  $\{v_i, v_j\}$  indicando a existência da ligação.

Uma aresta entre dois vértices pode ou não possuir uma direção: uma aresta pode partir de um vértice  $i$  e chegar a outro vértice  $j$ , sem que o inverso seja necessariamente verdadeiro. Para que isso ocorra há a necessidade de inserção de outra aresta entre o vértice  $j$  e o vértice  $i$  [5]. Se as arestas possuem direção, os grafos são chamados de dígrafos ou grafos direcionados.

Além da direção também pode ser atribuído um valor a uma aresta, sendo que este valor pode representar o custo para ir de um vértice a outro, e a aresta é representada como  $\{v_i, v_j, custo\}$ . Esse tipo de grafo recebe o nome de grafo valorado (ou multi valorado).

A figura 2.1(a) mostra um exemplo da representação de um grafo direcionado e na figura 2.1(b) não direcionado, ambos valorados, onde os pontos representam os vértices do grafo e as linhas as arestas(ou ligações) entre eles.

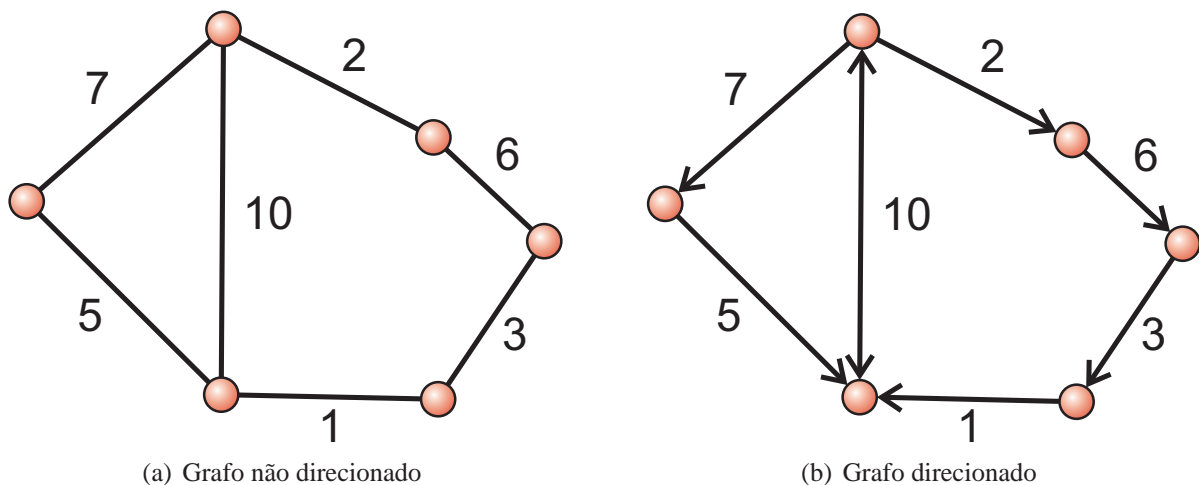


Figura 2.1: Representações de grafos valorados.

Neste trabalho são utilizados grafos não direcionados e não valorados, ou seja, uma aresta conecta dois vértices  $i$  e  $j$  não importando qual é a “origem” e o “destino”, bem como, o custo para o percurso desta aresta.

A ordem de um grafo é dada pela quantidade de vértices no mesmo, e o seu tamanho pelo número de arestas que ele possui. O número de arestas que um vértice possui é chamado de grau (ou conectividade) do vértice  $k_i$  e está limitado ao número total de arestas  $\frac{N(N-1)}{2}$ , no caso de não ser possível a existência de arestas que partam de  $i$  e cheguem nele mesmo (laços).

No caso dos grafos direcionados o grau de um vértice pode ser diferenciado em grau de entrada  $k_{in}(i)$  e grau de saída  $k_{out}(i)$ , assim  $k_i$  é dado pela soma desses dois valores ( $k_i = k_{in}(i) + k_{out}(i)$ ).

## 2.2 Grafos Aleatórios

A distribuição de graus, ou seja, o histograma que representa a distribuição das arestas pelos vértices da rede é um fator de grande importância para determinar a topologia de uma rede de interação [8]. Nesta seção são introduzidos, de forma algorítmica, alguns conceitos importantes para este trabalho relacionados a distribuição de conectividades.

Um grafo aleatório é definido como um conjunto de  $N$  vértices conectados por  $m$  arestas que são escolhidas de forma aleatória entre todas as ligações possíveis [2]. Para criar um grafo aleatório, inicialmente constrói-se uma lista de todos os vértices da rede. Posteriormente, as conexões (ou arestas) são inseridas entre vértices do grafo com uma probabilidade uniforme  $p$  resultando, ao final do processo, em um grafo com número total de arestas  $p \cdot \frac{N(N-1)}{2}$ .

Em um grafo aleatório a maioria dos vértices possui um número de ligações semelhante e igual a número médio de arestas do grafo dada por  $\langle k \rangle = N \cdot p$  para baixos valores de  $p$ . Isso é exemplificado na figura 2.2 que mostra a quantidade de indivíduos ( $P(k)$ ) que possuem certa quantidade de ligações ( $k$ ) para uma rede com 10000 indivíduos e  $p = 0.4$ .

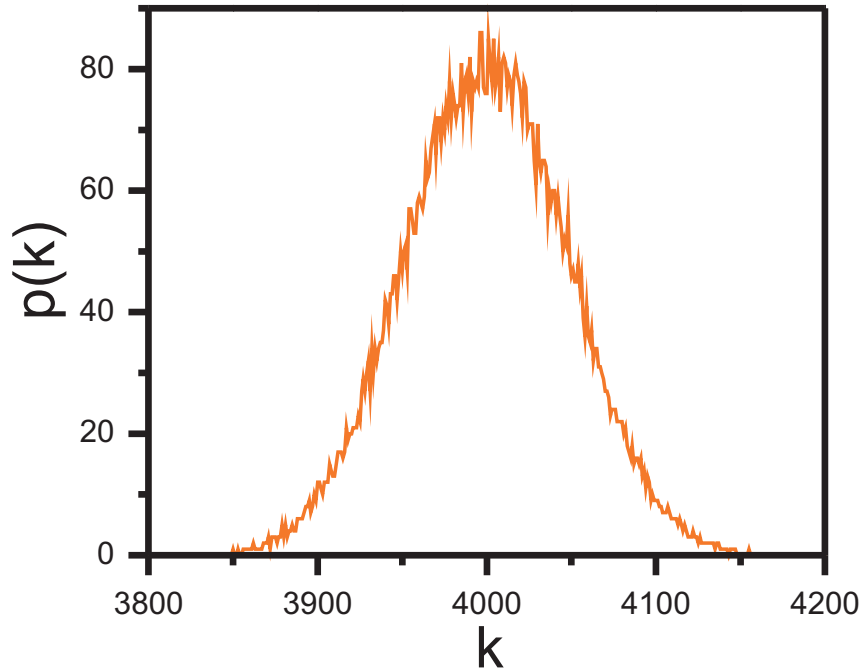


Figura 2.2: Distribuição das conectividades em um grafo aleatório.

## 2.3 Redes sem-escala

As redes sem-escala foram introduzidas por Barabási e Albert [1], para representar redes reais. Além disso, Barabási e Albert [1] apresentaram um modelo de crescimento de rede sem-escala que tem como alicerce dois mecanismos genéricos: crescimento e vinculação preferencial. O seu algoritmo está descrito abaixo:

1. *Crescimento*: Inicia-se com um grafo formado por um número pequeno de nós ( $n_0$ ) totalmente conectados. Em cada passo de tempo, adiciona-se um novo vértice ao gráfico inicial com  $m$  arestas que ligam esse novo nó a  $m$  outros diferentes vértices. Repete-se o procedimento até que um número  $N$  pré-definido de vértices seja inserido na rede.
2. *Vinculação preferencial*: Quando um novo vértice é inserido, a escolha da nova ligação é feita com uma probabilidade  $p$  que depende do grau  $k_i$  do vértice que a receberá. A probabilidade de que o vértice  $i$ , de grau  $k_i$  seja conectado ao novo vértice é  $p_i = \frac{k_i}{a(t)}$ , onde  $a(t)$  é o número de arestas existentes no grafo no passo de tempo  $t$ .

Após  $T$  passos de tempo esse processo resulta em uma rede com  $N = T + n_0$  nós e  $\left[ m \cdot T + \frac{n_0 \cdot (n_0 - 1)}{2} \right]$  arestas.

Em resumo, na construção da rede sem-escala, parte-se primeiramente de um grafo completo de  $(n_0)$  vértices. Também se necessita de um número  $N$  que represente a quantidade de vértices final do grafo. Cada um dos vértices será sequencialmente inserido na rede com  $m$  arestas, onde cada uma delas incidirá sobre um vértice já inserido anteriormente. A escolha desse vértice é feita com base na proporção de arestas que incidem sobre cada vértice em relação ao total de arestas no grafo segundo a probabilidade  $p_i$ . O esquema de implementação utilizado está representado na figura 2.3.

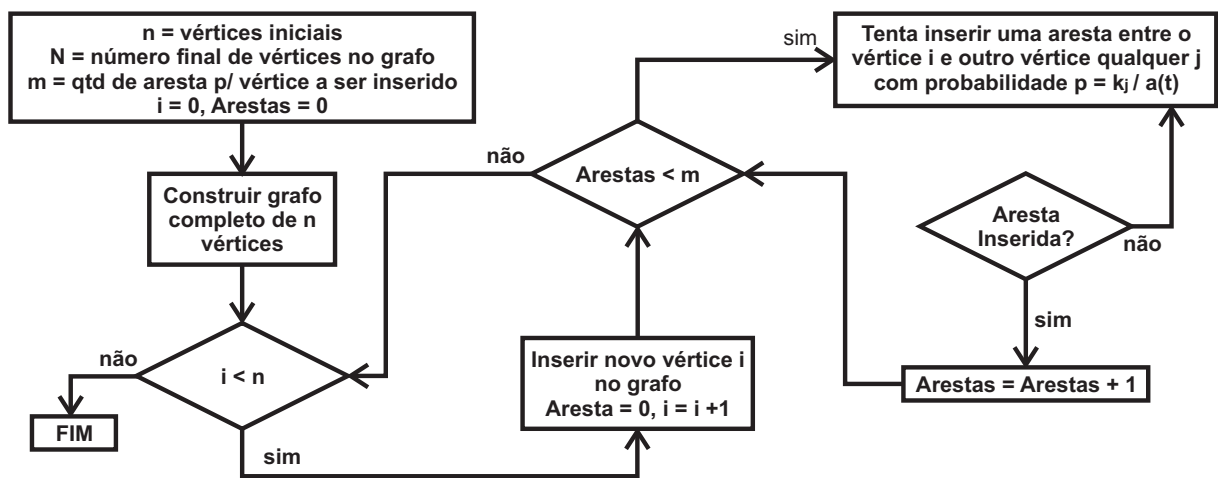


Figura 2.3: Construção da rede sem-escala segundo o modelo de Barabási-Albert [1]

A grande diferença entre as redes sem-escala e os outros tipos de rede é que sua distribuição de graus obedece a uma lei de potências. Na literatura, atribui-se a distribuição de graus em lei de potência ao fato que, no modelo de Barabási-Albert as conexões são feitas com base na conectividade de cada vértice.

Quando um vértice é inserido na rede ele é conectado a  $m$  outros vértices. A escolha desses vértices é feita a partir da quantidade de ligação que possuem em relação a quantidade total de ligações na rede e assim, quanto maior a conectividade de um vértice, maior é a probabilidade de que receba novas ligações, contribuindo para que existam “vértices centrais” que servem como grandes *hubs* de distribuição.

A figura 2.4 mostra o gráfico do comportamento da distribuição de conectividade dos vértices em uma rede sem-escala com 10000 indivíduos e  $m = 2$ . Em escala logarítmica o gráfico é uma reta. As pequenas flutuações para grandes valores de  $k$  são devido ao tamanho do grafo (em física, chamado de “efeito de tamanho infinito”). Esses efeitos podem ser

minimizados aumentando gradativamente o tamanho do sistema e extrapolando para o limite de um sistema infinito.

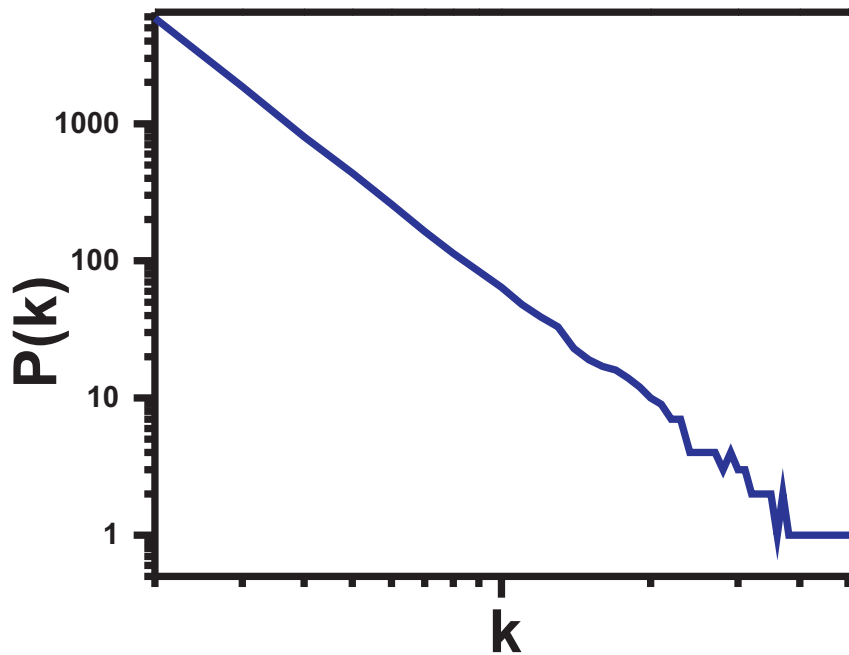


Figura 2.4: Distribuição de conectividades rede sem-escala

Pode-se notar que uma fração muito grande dos indivíduos da rede possui uma quantidade pequena de conexões, enquanto que uma pequena fração deles possui uma quantidade grande de ligações, ou seja a distribuição de conexões é muito heterogênea.

No próximo capítulo são descritos aspectos da estruturação interna destas redes complexas, as quais vão além da caracterização global da rede por sua distribuição de conectividades.

# Capítulo 3

## Redes modulares

### 3.1 Redes estruturadas em comunidades

As redes estruturadas em comunidade são redes complexas que apresentam além das características de heterogeneidade global, certos níveis de organização interna na forma de agrupamentos. Neste tipo de rede a conexão entre os indivíduos de um mesmo agrupamento (ou comunidade) está fortemente estabelecida, mas os indivíduos também possuem ligações com indivíduos de outras comunidades. [9]

Ligações entre indivíduos de uma mesma comunidade são denominadas de ligações *intra-comunidade*, já as ligações entre indivíduos de comunidades diferentes são denominadas conexões *inter-comunidade*. A caracterização destas comunidades pode ser feita intuitivamente com base na diferença da densidade de ligações intra e inter comunidade.

Esta forma de estruturação dos indivíduos pode simular de forma satisfatória a estruturação dos seres humanos em relação a suas redes sociais e de relacionamentos [2]. A figura 3.1 é uma representação compacta de como é a estruturação de uma rede com comunidades. Dois grupos distintos podem ser visualmente identificados. No interior de cada grupo cada indivíduo tem 3 conexões para elementos do mesmo grupo (intra-comunidade) e apenas alguns dos indivíduos conectam-se com elementos de outro grupo (ligações inter-comunidade).

De fato a existência de comunidades é bastante intuitiva, mas, não existe uma única forma de defini-la precisamente. Algumas definições permitem que os vértices possam pertencer a mais de uma comunidade, tornando possível a visão de comunidades sobrepostas e não sobrepostas. O próprio conceito de estruturação em comunidades causa ambiguidade no meio científico já que as comunidades podem ser concebidas como uma partição simples do grafo ou como uma

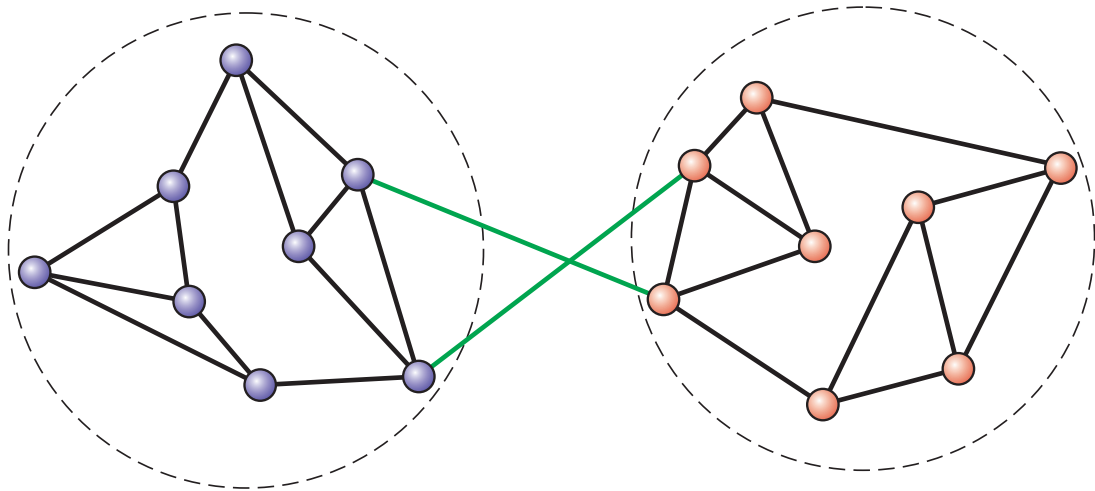


Figura 3.1: Exemplo de rede estruturada em comunidades com 2 comunidades, 3 ligações *intra-comunidade* e 2 ligações *inter-comunidade*

hierarquia de partições. No entanto, é importante frisar que a identificação de aglomerados somente se torna possível em grafos esparsos.

Muitas definições das estruturas em comunidades foram proposta na literatura por físicos e cientistas da computação, podendo estas serem classificadas em três categorias [2]:

1. *Definições locais*: Aqui a atenção se volta a investigação da vizinhança imediata dos vértices, ficando de lado a estruturação geral do grafo.
2. *Definições globais*: Nesta definição se torna possível a identificação das características das comunidades com a análise de subgrafos com relação ao grafo como um todo, uma vez que a comunidades são unidades estruturais do grafo.
3. *Definições baseadas na similaridade dos vértices*: As comunidades são definidas a partir da similaridade dos vértices pertencentes ao grafo. Esta similaridade e baseada em um critério qualitativo local ou global.

Estas medidas não são apenas importantes para a caracterização topológica do grafo, pois também podem dar informações sobre a funcionalidade da rede e a sua constituição. A escolha da estruturação da rede vai depender do estudo que se deseja fazer, e a forma com que ele será dirigido.

Em uma rede com estrutura em comunidades existe geralmente um número grande de possíveis partições. Neste trabalho, uma partição é uma divisão de uma rede em aglomerados,

onde cada um dos vértices pertence a apenas um desses aglomerados. O número de partições de uma rede é necessário para que seja possível o estabelecimento da estruturação em comunidades. A qualidade desta estruturação pode ser avaliada através de uma grandeza conhecida como fator de modularidade [10].

A definição de modularidade mais popular foi proposta por Newman e Girvan [10] e é dada por

$$Q = \frac{1}{2} \sum_{ij} \left( A_{ij} - \frac{k_i k_j}{2a} \right) \delta(C_i, C_j) \quad (3.1)$$

onde soma-se todos os pares de vértices,  $A$  é a matriz adjacência,  $k_i$  é o grau de conectividade do vértice  $i$ , e  $a$  é o número total de arestas do grafo. O valor de  $\delta(C_i, C_j)$  é 1 se  $i$  e  $j$  pertencem a mesma comunidade, e 0 (zero) se  $i$  e  $j$  pertencentes a comunidades diferentes. O elemento  $A_{ij}$  da matriz adjacência será 1 se os vértices  $i$  e  $j$  estão conectados, e 0 (zero) caso contrário.

Segundo Newman [10], a modularidade é a quantidade que mede a qualidade do particionamento de uma rede. Quanto maior o valor da modularidade, a aglomeração em comunidades é mais intensa, sendo que, nos casos ideais, para  $Q = 1$ , a aglomeração é máxima, e se torna mínima quando  $Q = 0$ . Na aplicação de diversos algoritmos de detecção de comunidades em redes reais [3] [7] têm sido encontrados valores entre  $Q = 0.3$  e  $Q = 0.7$  [10].

Como a caracterização da rede é feita por seu fator de modularidade daqui por diante as redes estruturadas em comunidades serão chamadas “redes modulares”.

O interesse neste trabalho está na obtenção de redes modulares não interessando portanto os métodos de particionamento, visto que o objetivo é investigar um processo dinâmico (a propagação de epidemias) sobre redes com essa organização. Por este motivo optou-se por modelos de geração de redes modulares com controle da estrutura em comunidades [8], como será apresentado na próxima seção.

## 3.2 Modelo de Li-Maini

Nesta seção apresenta-se os modelos utilizados e a caracterização das redes obtidas com base em seu fator de modularidade.



O modelo de Li-Maini [8][9] é um gerador de topologia de redes sem-escala modular no qual é possível o controle da intensidade da organização. A intensidade da organização interna é controlada por um parâmetro do modelo e a variação de tal parâmetro reflete-se diretamente sobre o valor de  $Q$  [2]. O número de comunidades ( $M > 2$ ) é um dos parâmetros de inicialização da rede e pode ser ajustado.

O modelo de Li-Maini é baseado no modelo de rede sem-escala de Barabási-Albert[1] e pode ser dividido em três etapas[2]:

1. *Inicialização*: Inicia-se cada uma das  $M$  comunidades de forma totalmente isolada com um pequeno número  $n_0$  ( $n_0 > 1$ ) de vértices totalmente conectados. Posteriormente, são adicionadas conexões entre comunidades (cada conexão é feita escolhendo os vértices aleatoriamente em cada comunidade), de modo que tenha no total  $\frac{M(M-1)}{2}$  conexões inter-comunidades.
2. *Crescimento*: Em cada iteração ou etapa de tempo, um novo vértice é inserido na rede e fazendo parte de uma comunidade escolhida aleatoriamente. Este novo vértice é conectado a  $m$  ( $1 \leq m \leq m_0$ ) vértices da mesma comunidade, caracterizando as conexões intra-comunidade. Este novo vértice também pode se conectar com indivíduos de outras comunidades. A existência de conexões inter-comunidades depende de uma probabilidade  $\alpha$  uniforme. Uma vez determinado se haverá conexões inter-comunidades o vértice se conecta a  $n$  ( $1 \leq n \leq M$ ) vértices de outras comunidades, representando as conexões inter-comunidades. O estabelecimento da conexão inter-comunidades é feita com uma probabilidade de conexão  $\alpha$ . Os valores de  $m$ ,  $n$  e  $\alpha$  são parâmetros do modelo.
3. *Vinculação preferencial*: Cada vértice inserido recebe uma nova conexão com uma probabilidade  $p$ , assim como no modelo de Barabási-Albert. Dessa forma, a rede crescerá até atingir um valor  $N$ , que representa o número de elementos da rede.

O esquema de implementação utilizado está representado na figura 3.2.

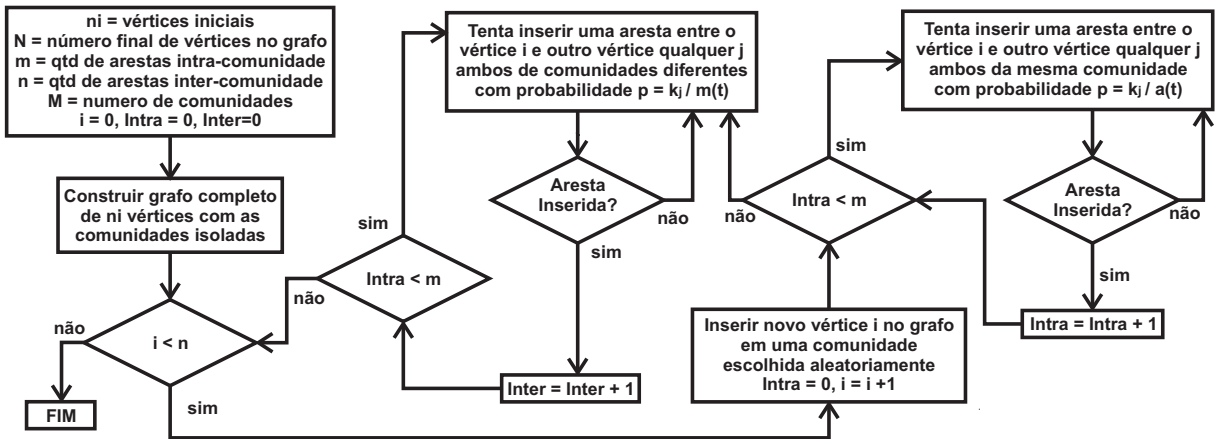


Figura 3.2: Construção da rede sem-escala modular segundo Li e Maini [9].

O modelo foi implementado em linguagem C++ e executado em uma máquina com processador Intel Pentium 4 3.0 GHz e 1 GB de memória RAM. A figura 3.3 apresenta o gráfico do fator de modularidade calculado diretamente na rede modular para diferentes valores de  $\alpha$  e valores de  $m = 4$  e  $n = 2$ .

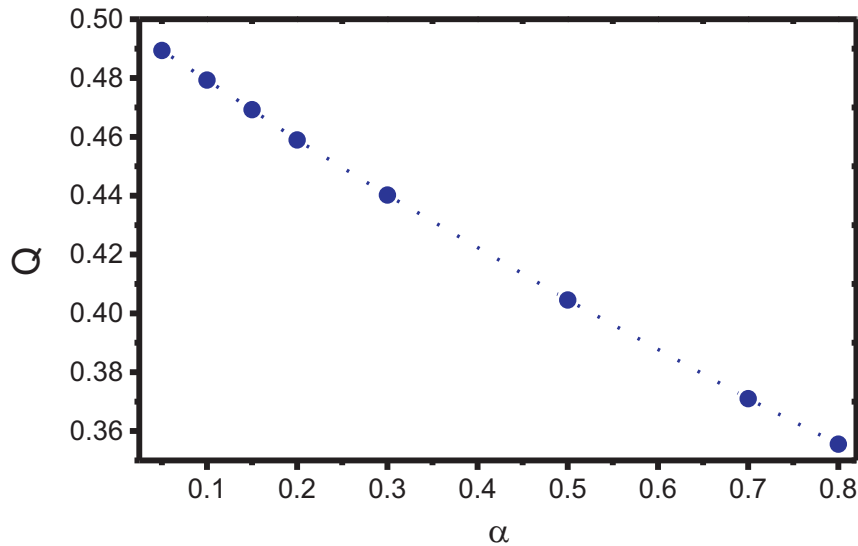


Figura 3.3:  $Q$  para valores de  $\alpha$  na rede modular segundo o modelo de Li-Maini

Como o fator  $\alpha$  controla a probabilidade de que ligações inter-comunidades sejam estabelecidas, quanto maior o valor de  $\alpha$ , maior será o número de ligações inter-comunidades e o índice de modularidade diminui. Por outro lado, se  $\alpha = 0$ , não existirão ligações inter-comunidades e  $Q = 0.5$ , pois existem duas comunidades bem distintas.

O que pode-se notar é que conforme o valor de  $\alpha$  cresce, a aglomeração dos indivíduos se

torna menor, visto que um número maior de ligações são compartilhadas entre as diferentes comunidades e a qualidade da estrutura das comunidades, definida pelo fator de modularidade decai.

A estruturação da rede em comunidades pode ser analisada visualmente através das matrizes de esparsidade. A matriz de esparsidade é uma representação da matriz de adjacências do grafo, construída de forma que vértices pertencentes às mesmas comunidades sejam colocados sequencialmente como as linhas (e colunas) da matriz adjacência.

A Figura 3.4 retirada de [2] mostra a matriz de esparsidade para uma rede com  $N = 1500$  elementos,  $M = 8$  comunidades,  $m = 4$ ,  $n = 7$  e diferentes valores de  $\alpha$ .

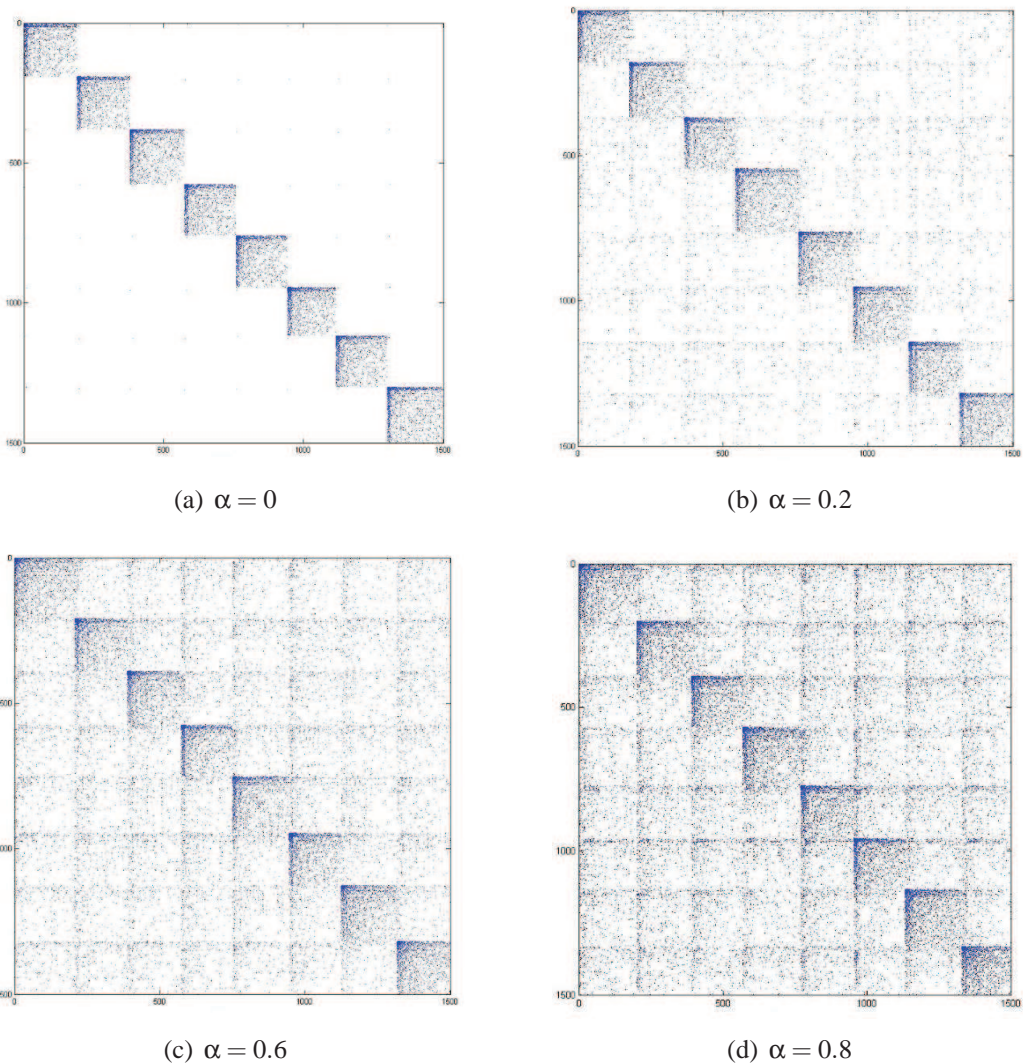


Figura 3.4: Matriz de esparsidade para uma rede com  $N = 1500$  elementos,  $M = 8$  comunidades,  $m = 4$ ,  $n = 7$  e diferentes valores para  $\alpha$

Para melhor verificar a dependência da modularidade em relação à  $\alpha$ , foram utilizados os seguintes valores:  $\alpha = 0$ ,  $\alpha = 0.2$ ,  $\alpha = 0.6$  e  $\alpha = 0.8$ . Quanto menor for o valor de  $\alpha$ , menor será o número de conexões entre elementos de comunidades distintas, e a rede exibe uma estrutura em comunidades bem definida. Quando o valor de  $\alpha$  é aumentado, cresce a probabilidade de que cada vértice realize conexões com elementos de outras comunidades, fazendo aumentar o número de conexões entre as comunidades, diminuindo o índice de modularidade.

### **3.3 Redes modulares espaciais**

Nesta seção é apresentado um modelo de rede modular baseado no modelo de rede de Yook [13]. O modelo de Yook une a distribuição de graus em lei de potência a distribuição espacial dos vértices na rede. Neste trabalho o modelo de Yook foi generalizado, utilizando as ideias de Li-Maini, a fim de obter redes modulares espaciais.

Inicialmente é descrito o modelo de Yook e suas motivações. Os fundamentos, bem como conceitos e definições mencionados neste capítulo foram retirados do artigo Modeling the Internet's large-scale topology, escrito por Yook, Jeong e Barabási [13].

#### **3.3.1 O Modelo de Yook**

O modelo de Yook para redes complexas surgiu durante estudos da estrutura da Internet. O protocolo de Internet é fortemente influenciado pelo desempenho da topologia da rede, logo os geradores de redes são uma condição previa e essencial para a modelagem da Internet. Para controlar de forma eficiente e rotear o tráfego na Internet é importante que os geradores de topologia não só captem a estrutura da rede, mas também permitam um planejamento eficaz e de longo prazo.

O desenvolvimento de bons geradores de topologia é limitado pelo pouco entendimento dos mecanismos básicos que formam a topologia de larga-escala da Internet. Até recentemente, os geradores de topologia da Internet com propósito de pesquisa e desenvolvimento, construídos para gerar redes reais com diferentes parâmetros de entrada, utilizavam modelos de grafos aleatórios.

Com o descobrimento em 1999, por Faloutsos [6], de que a Internet é uma rede sem-escala com uma distribuição de grau em lei de potencia, todos os esforços anteriores acabaram

comprometidos. As pesquisas confirmaram diferenças significativas entre redes sem-escala e aleatórias, frisando que a importância dessas diferenças é muito significativa para ser ignorada. Como exemplos de algumas diferenças entre esses tipos de redes temos:

- Os protocolos desenvolvidos para redes aleatórias obtiveram resultados pobres em topologias sem-escala;
- Uma rede sem-escala apresenta alta tolerância a falhas aleatórias em nós, mas é frágil contra ataques enquanto que redes aleatórias preveem redes robustas a falhas e resistentes à ataques.
- O espalhamento de vírus computacionais em redes sem-escala gera consequências na segurança da rede uma vez que o espalhamento de vírus é muito mais rápido que em redes aleatórias.

Essas diferenças geraram um novo marco no desenvolvimento de geradores de topologia da Internet os quais fornecem topologias sem-escala com melhor concordância com dados empíricos.

### **Layout Físico**

Numa visão mais ampla a Internet é uma rede de roteadores conectados por links. Como cada roteador pertence a alguma autoridade administrativa, ou um sistema autônomo (SA), a Internet é frequentemente considerada com uma rede de SAs interconectados. Para simplificação, e pelo fato de serem estudados simultaneamente o termo nó será usado para representar tanto os roteadores quanto os SA.

Os geradores de topologia de Internet assumem que os nós são distribuídos aleatoriamente em um plano bidimensional. A figura 3.5(a) extraída de [13] mostra um mapa da densidade mundial dos roteadores com maior nível de esforço na Internet. Comparado com a densidade populacional (figura 3.5(b)), nota-se visualmente as correlações entre os roteadores e a densidade populacional nas zonas economicamente desenvolvidas do mundo.

A coincidência entre a densidade populacional e os nós da Internet não era inesperada, já que uma densidade populacional alta implica numa grande demanda por serviços de Internet, resultando numa densidade grande de roteadores e domínios.

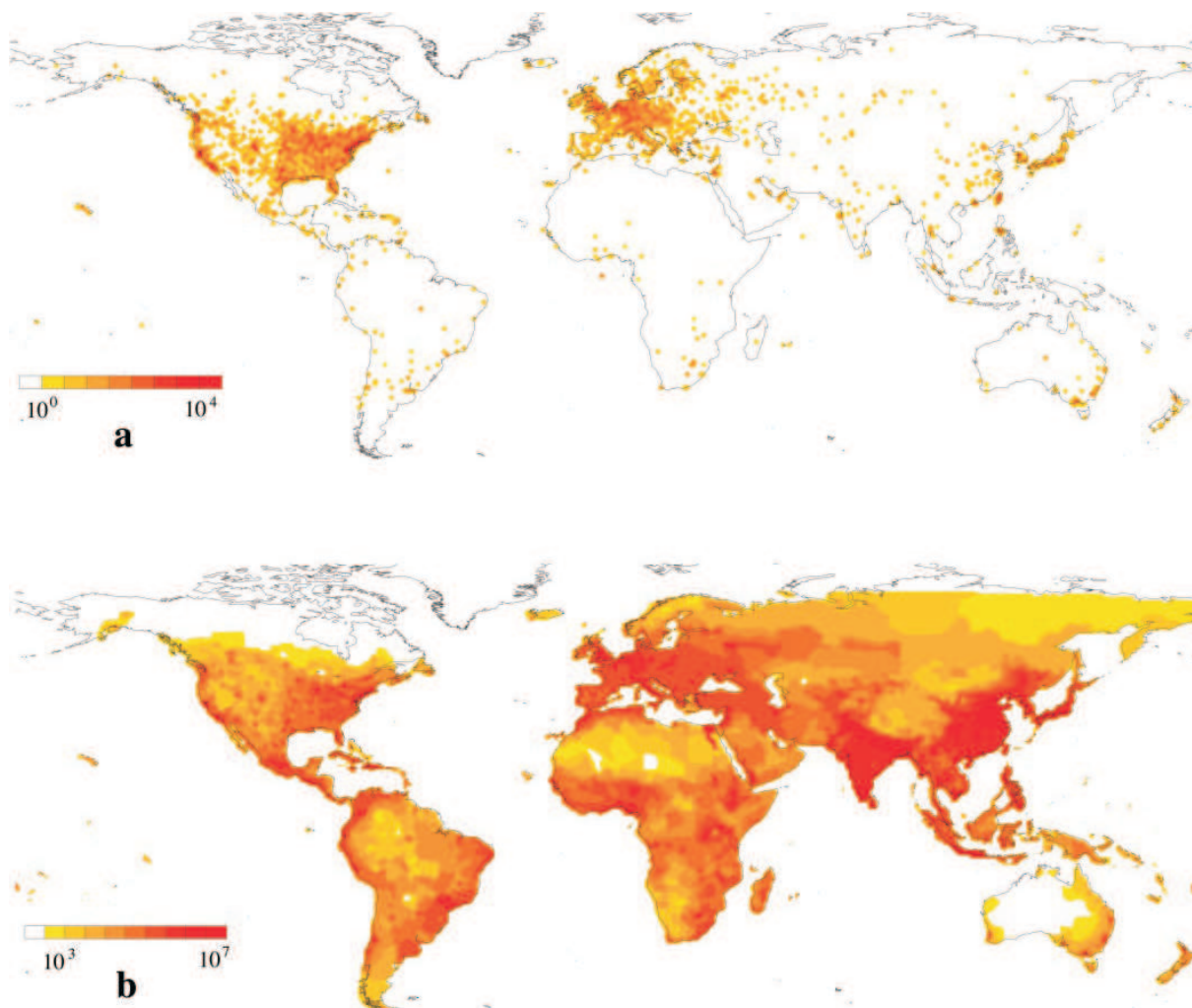


Figura 3.5: Distribuição da Internet ao redor do mundo (a) Densidade mundial dos roteadores com maior nível de esforço e (b) Densidade populacional

Como a conexão entre dois nós na Internet requer recursos extensivos, na criação de uma rede é preferível conectar nós mais próximos, com suficiente largura de banda, um processo que favorece claramente conexões menos extensas. Para desencorajar links longos, os geradores de topologia baseados no modelo de Waxman [12], assumem que a probabilidade de conexão entre dois nós separados por uma distância Euclidiana  $d$  decai com  $P(d) \sim \exp\left(\frac{-d}{d_0}\right)$ , onde  $d_0$  é um parâmetro livre e proporcional ao tamanho do sistema.

Apesar de sua ampla utilização em geradores de topologia de Internet não há nenhuma evidência empírica para essa forma exponencial que proíbe ligações entre nós distantes. A intuição sugere outra coisa: seria de se esperar que a probabilidade de conexão entre dois nós fosse inversamente proporcional à distância entre eles, isto é,  $P(d) \sim \frac{1}{d}$ .

Realmente, o custo da colocação de um link físico entre dois roteadores existentes tem dois componentes: (i) uma conexão física com custos técnicos e administrativos nas duas extremidades do link e (ii) o custo da linha física e sua manutenção. O segundo fator é proporcional ao tamanho da linha. Para grandes distâncias os custos dependentes da distância são dominantes, potencialmente sugerindo uma dependência assintótica  $\frac{1}{d}$ , para a probabilidade de conexão de dois roteadores.

A forma funcional correta de  $P(d)$  é crucial para a modelagem da Internet: simulações indicam que redes desenvolvidas sobre a regra assintótica de Waxman convergem a uma rede com uma distribuição de graus que decai exponencialmente, em contraste com a lei de potências documentada para a Internet.

### **Vinculação Preferencial**

A vinculação preferencial aos nós de maior conectividade é responsável pelo aparecimento da estrutura sem-escala em redes complexas. Assume-se que a probabilidade de que um novo nó se conecte a outro existente com  $k$  ligações depende linearmente de  $k$ , ou seja,  $P(k) = \frac{k}{\sum_i k_i}$ .

Por outro lado, em sistemas reais a vinculação preferencial pode ter uma forma arbitrária não linear. Os cálculos indicam, no entanto, que para  $P(k) \sim k^\sigma$ , com  $\sigma \neq 1$  a distribuição de graus afasta-se de uma lei de potência.

Para que a Internet seja modelada de forma correta,  $P(k)$  precisa ser determinado. O cálculo da mudança  $\Delta k$  em graus de um nó SAs com  $k$  links foi possível através de mapas de SAs da Internet armazenados num intervalo de 6 meses. Nota-se a taxa de aumento do grau de um nó é diretamente proporcional ao número de ligações que o nó tem, oferecendo suporte quantitativo para a presença da vinculação preferencial linear.

Existem quatro mecanismos que agindo de forma independente que contribuem para a topologia de larga-escala da Internet. Primeiro, em contraste com os modelos clássicos de rede a Internet cresce gradativamente, sendo descrita por uma rede que evolui ao invés de um grafo estático. Segundo, os nós não são distribuídos aleatoriamente, mas ambos, roteadores e domínios formam uma escala fractal invariante com uma dimensão fractal  $D_f = 1.5$ . Finalmente, a colocação dos nós é determinada por dois mecanismos concorrentes. Primeiro, a probabilidade de conexão de dois nós cai linearmente com a distância deles, e segundo, a probabilidade de conexão a um nó com  $k$  ligações aumenta linearmente com  $k$ .

O modelo de Yook para formação de tais redes considera um mapa que é uma superfície bidimensional de tamanho linear  $L$ . O mapa é dividido em quadrados de tamanho  $l \times l$  ( $l \ll L$ ), a cada quadrado é atribuída uma densidade populacional  $p(x, y)$ . Em cada passo de interação (tempo) um novo nó  $i$  é adicionado e sua posição é determinada probabilisticamente, de tal forma que a probabilidade de colocação de um nó na posição  $(x, y)$  é diretamente proporcional a  $p(x, y)$ . O novo nó se conecta com  $m$  nós já presentes na rede, sendo que a probabilidade de que o novo nó se conecte a outro um nó  $j$  com  $k_j$  ligações a uma distância  $d_{ij}$  do nó  $i$  conforme [13] é

$$p(k_j, d_{ij}) \sim \frac{k_j^\sigma}{d_{ij}^\chi} \quad (3.2)$$

onde  $\sigma$  e  $\chi$  são expoentes que regem a vinculação preferencial e o custo da conexão a distância nó-a-nó. O aumento de  $\sigma$  favorece nós com grande número de conexões, enquanto que um maior  $\chi$  desencoraja ligações longas.

Segundo Yook [13] é possível identificar de forma inequívoca a posição da Internet dentro do espaço de fase  $\sigma = 1$  e  $\chi = 1$ . Essa separação não deve ser um problema se alguns modelos e a Internet pertencem a uma região do espaço de fase que compartilham os mesmos recursos topológicos universais.

A vantagem deste modelo é a sua flexibilidade: ele oferece um esqueleto universalmente aceitável para potenciais modelos de redes complexas espaciais, nos quais podem ser criadas funcionalidades que podem gerar novas melhorias. Utilizar uma rede que evoluiu para modelar a Internet propicia uma potencial predição do futuro da rede, já que o modelo incorpora apenas mecanismos com tempo invariante que devem continuar conduzindo o desenvolvimento futuro da rede.

Para esse trabalho o modelo de Yook [13] foi generalizado, utilizando as ideias de Li-Maini [9] a fim de obter redes modulares espaciais. A hipótese formulada é que as comunidades sejam estabelecidas em regiões espaciais bem definidas. Elementos tenderiam a se relacionar com elementos da mesma comunidade, espacialmente próxima. Além disso, as conexões de longo alcance representam o estabelecimento dos relacionamentos observados em redes sociais, as quais reconhecidamente podem ser de longo alcance.

O modelo resultante é descrito a seguir.



### 3.3.2 Modelo de redes modulares espaciais

O modelo de redes modulares espaciais é construído a partir do modelo de Yook [13] e das ideias de Li-Maini [8] para representação de aglomerações através da estruturação em comunidades. A construção se baseia no crescimento da rede, dividido em inicialização e crescimento, sendo que posteriormente a vinculação preferencial dos vértices é feita.

Abaixo temos o algoritmo que se divide em três passos e foi utilizado para a criação desse tipo de rede.

1. *Inicialização*: Inicia-se cada uma das  $M$  comunidades de forma totalmente isolada com um pequeno número  $n_0 (n_0 > 1)$  de vértices disposto topologicamente de forma aleatória e totalmente conectados. Posteriormente, são adicionadas conexões entre comunidades (cada conexão é feita escolhendo os vértices aleatoriamente em cada comunidade), de modo que tenha no total  $\frac{M(M-1)}{2}$  conexões inter-comunidades. Para a disposição espacial dos vértices utiliza-se uma matriz quadrada de lado  $L$  ( $L^2$  elementos). Sendo assim, existem  $L \times L$  posições possíveis para a inserção de indivíduos nesta superfície. A matriz é dividida em  $M$  regiões sendo que cada comunidade ocupa uma e apenas uma destas regiões. A posição de um elemento no espaço coincide com sua localização na matriz.
2. *Crescimento* : Em cada etapa de tempo, um novo vértice é inserido a uma comunidade escolhida aleatoriamente e é adicionado em uma posição  $(x, y)$  aleatória ( $0 \leq x < L$ ) ( $0 \leq y < L$ ) ( $x, y \in \mathbb{Z}$ ). Este novo vértice será conectado a  $m$  ( $1 \leq m \leq m_0$ ) vértices da mesma comunidade, caracterizando as conexões intra-comunidade e a  $n$  ( $1 \leq n \leq M$ ) vértices de outras comunidades, representando as conexões inter-comunidades. A conexão inter-comunidade será feita obedecendo a uma probabilidade de conexão  $\alpha$  conforme descrito no modelo de Li-Maini. Os valores de  $m$ ,  $n$  e  $\alpha$  são definidos de antemão.
3. *Vinculação preferencial* : Cada vértice inserido recebe uma nova conexão com uma probabilidade  $p_{ij} = \frac{p_j^\sigma}{d_{ij}^\chi}$ , sendo que  $p_j$  assim como no modelo de Barabási-Albert é a fração de arestas do vértice  $j$  elevada a um expoente de controle  $\sigma$  e  $d_{ij}$  a distância entre dois vértices  $i$  e  $j$  elevada a um expoente  $\chi$ . Dessa forma, a rede crescerá até atingir um valor  $N$ , que representa o número de elementos da rede.

Se na vinculação preferencial  $\chi = 0$  teremos que a distância entre os vértices não influenciará na probabilidade de conexão entre eles e teremos uma rede sem-escala convencional. Vale ainda ressaltar que quanto menor o valor de  $\chi$  menor a influência da distância na vinculação dos vértices.

Em resumo, a vinculação preferencial se baseia na fração entre a proporção de conexões dos vértices já pertencentes a rede em relação ao todo e a distância entre um vértice a ser inserido e um vértice qualquer ao qual deseja-se conectar.

Uma das diferenças entre o Modelo de Yook [13] e o modelo de Barabási-Albert [1] se dá pelo fato de que no de Yook, no passo da Vinculação Preferencial, a probabilidade de um vértice se conectar a outro não se baseia apenas na conectividade do vértice ao qual o novo vértice será conectado, mas na fração entre esse valor e a distância entre os dois. Nesta implementação é comparada também a existência de comunidades espacialmente localizadas.

O esquema de implementação que utilizamos está representado na figura 3.6 e foi implementado em linguagem C++ e executado em uma máquina com processador Intel Pentium 4 3.0 GHz e 1 GB de memória RAM.

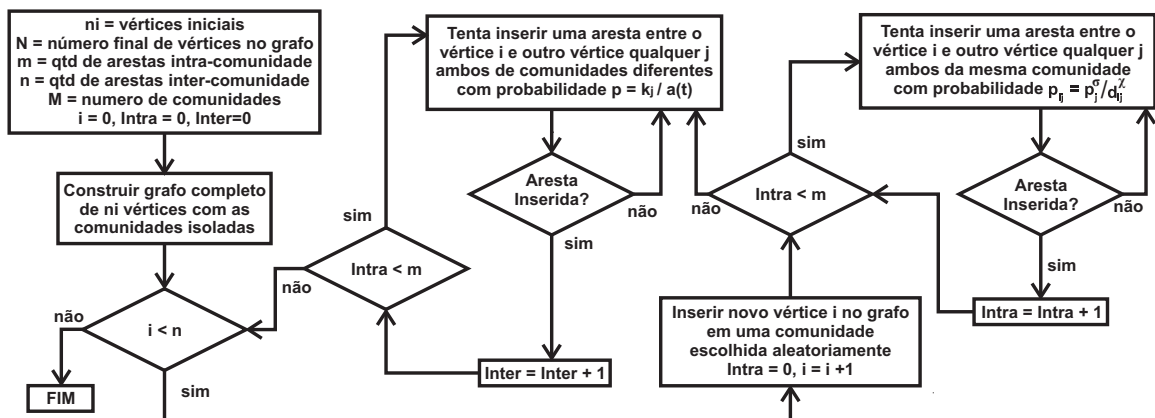


Figura 3.6: Construção da rede topológica estruturada em comunidades

Para iniciar a construção da rede inicia-se a superfície de tamanho  $L \times L$  vazia, ou seja, não existe nenhum indivíduo na rede.

Após isso são inseridos  $n_0$  indivíduos na rede. Cada uma das  $M$  comunidades da rede será iniciada com aproximadamente  $\frac{n_0}{M}$  indivíduos, sendo a probabilidade de que um vértice inserido pertencerá a uma comunidade igual a  $\frac{1}{M}$ .

Um novo elemento inserido na rede é adicionado a uma comunidade escolhida

aleatoriamente, respeitando a disposição espacial da comunidade.

Após a escolha da comunidade o indivíduo é disposto em uma posição  $(x,y)$  escolhida aleatoriamente na região daquela comunidade. A posição onde o vértice será posicionado não pode estar sendo ocupada por outro indivíduo inserido anteriormente na rede.

Com uma posição definida na superfície o próximo passo consiste em conectar o vértice a  $m$  indivíduos que pertençam à mesma comunidade que ele. O vértice pode ainda, com uma probabilidade  $\alpha$ , se conectar a outros  $n$  indivíduos de comunidades diferentes da sua, seguindo o modelo de Li-Maini.

A escolha de um vértice  $j$ , pertencente à rede, ao qual o vértice  $i$ , que está sendo inserido, irá se conectar é baseada na razão entre a fração de ligações de  $j$  em relação ao número total de arestas do grafo  $\left(\frac{k_j}{a}\right)$  e a distância Euclidiana entre os vértices  $i$  e  $j$  ( $d_{ij}$ ) calculada conforme a Equação 3.2.

Assim como no modelo de Li-Maini o processo de inserção de novos vértices termina após  $T$  passos de repetição. Quando essas inserções findarem a rede terá  $n_0 + N$  indivíduos dispostos na superfície.

Foram simuladas redes com  $N = 2000$  indivíduos e dimensão espacial  $L = 100$ . A figura 3.7 apresenta o gráfico do fator de modularidade calculado diretamente na rede modular espacial em relação a valores de  $\alpha$  e valores de  $m = 4$  e  $n = 2$ . O valor dos expoentes  $\sigma$  e  $\chi$  foram fixos em  $\sigma = \chi = 1$

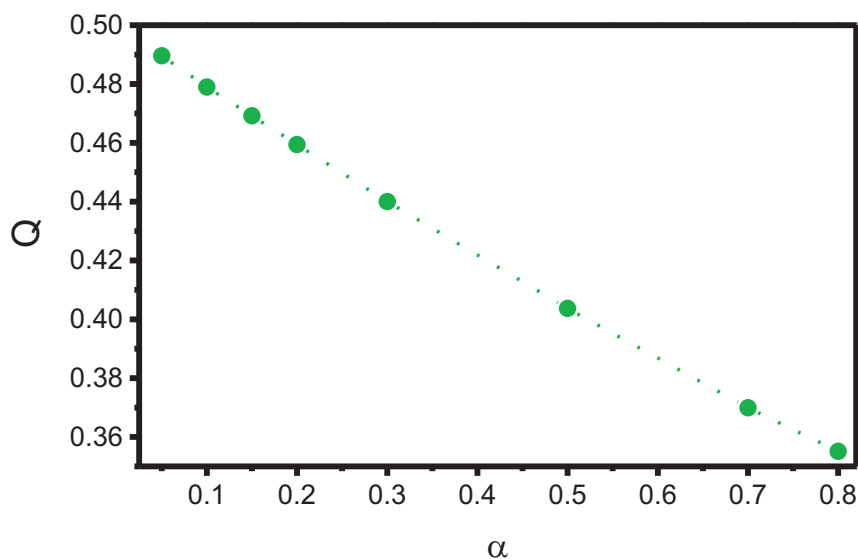


Figura 3.7:  $Q$  para diferentes valores de  $\alpha$  para rede modular espacial

Percebe-se, assim como nas redes modulares de Li-Maini [9] que o crescimento do valor de  $\alpha$  torna a aglomeração dos indivíduos maior, fazendo com que a qualidade da estrutura das comunidades, definida pelo fator de modularidade, decaia.

Comparando os filtros de modularidade em ambos os tipos de rede (com e sem distribuição espacial) verifica-se que não existem diferenças significativas, tanto nos valores obtidos quanto no comportamento em relação a  $\alpha$ .

Como o valor de  $\alpha$  não depende da distribuição espacial dos indivíduos a modularidade é semelhante, tanto para redes modulares, quanto para redes modulares espaciais. Com isso a epidemia se comporta de maneira similar em ambas as redes.

O próximo capítulo descreve como o comportamento de epidemias pode ser representado através dos modelos epidemiológicos, apresentando ainda alguns dos principais modelos atualmente utilizados e algumas características que os diferenciam.

# Capítulo 4

## Modelos Epidemiológicos

Este capítulo descreve como o comportamento de epidemias pode ser representado através dos modelos epidemiológicos. São apresentados ainda alguns dos principais modelos atualmente utilizados, bem como, algumas características que os diferenciam.

Na seção 4.1 são introduzidas informações para auxiliar a compreensão do processo epidêmico utilizado. A seção 4.2 descreve o modelo SIR que serviu de base para o desenvolvimento desse trabalho.

### 4.1 Conceitos básicos

Uma epidemia pode ser caracterizada pela existência de um número além do esperado de casos de uma doença em uma comunidade ou grupo, em um curto período de tempo. Com isso, conceitua-se epidemiologia como o estudo dos efeitos de uma doença sobre uma determinada população, visando, através de análises dos fatores determinantes para o surgimento da enfermidade, propor medidas de prevenção, controle ou erradicação [2]. Ressalta-se que a preocupação com os efeitos da epidemia não está focada apenas em relação aos casos fatais de doenças, mas também a outras consequências causadas pelas doenças em estudo.

Em uma comunidade o comportamento de uma epidemia varia de acordo com o tempo, ou seja, está em continua mudança. Está mudança condiciona o comportamento da doença e está relacionada com as características da epidemia [11].

A epidemiologia matemática é uma das áreas da epidemiologia que vem se fortalecendo devido a interação interdisciplinar entre biólogos, cientistas da computação, epidemiologistas, físicos, matemáticos, entre outros. Sua fundamentação se consolida a partir de quantificações

de fenômenos biológicos, transcritos através de hipóteses matemáticas que quantificam os aspectos do fenômeno. Os modelos matemáticos desenvolvidos dessa forma, em condições ideais, oferecem evidências suficientes para a criação e implantação de medidas de controle e prevenção de doenças. Devido a sua importância, modelar matematicamente doenças infecciosas tem sido objeto de vários trabalhos [11].

Uma das abordagens da modelagem matemática consiste em aplicar modelos determinísticos onde, dadas as condições iniciais, se torne possível, através de equações diferenciais e métodos computacionais que as possam solucionar, determinar de forma precisa o futuro da epidemia, elucidar a dinâmica das interações e facilitar a detecção de doenças (re)emergentes [11].

Os métodos probabilísticos, em particular os processos estocásticos, apesar de não poderem prever exatamente o estado futuro, mas apenas as probabilidades associadas a ocorrência de cada um dos estados, desempenha um papel importante na modelagem epidemiológica. Devido às técnicas utilizadas para suas soluções, em geral esses modelos contém mais detalhes do processo que um modelo determinístico [2].

Como os modelos estocásticos envolvem maiores detalhes sobre as propriedades do sistema, e podem ser elaborados com base nas características dos indivíduos e na sua estrutura de contatos, sua descrição do sistema é acompanhada pelo aumento da complexidade matemática necessária para a solução. Devido a isso, são necessárias técnicas computacionais de investigação.

A população pode ser representada matematicamente como um grafo e as interações entre indivíduos como redes de contatos. Estas redes podem ser formadas por populações humanas, animais, vegetais, redes de computadores e a própria web (para o caso da propagação de vírus de computador) [1].

Com o desenvolvimento dos geradores de topologia para redes de interesse, tornou-se possível a investigação do comportamento de modelos para processos físicos, sejam eles de motivação biológica ou social. Pesquisas sobre os processos dinâmicos sobre redes complexas [1][10], tem progredido menos do que os estudos sobre suas características estruturais. Isso ocorre devido ao fato de que o entendimento completo e simultâneo da estrutura e seus efeitos na dinâmica de um sistema é difícil. Assim, é comum optar por dividir o problema em partes: investigar inicialmente a estrutura da rede e depois os processos dinâmicos sobre elas. Com

isso, avanços importantes, em particular no estudos de falhas da redes e processos epidêmicos sobre elas têm sido alcançados [1][10].

As redes de contato, após sua construção, podem ser analisadas afim de obter o comportamento dos processos físicos nelas imersos. Nestas redes os indivíduos são associados aos vértices (ou nós) da rede e as arestas(ou ligações) aos contatos existentes entre os indivíduos. Torna-se também possível a investigação dos mecanismos de propagação dos processos físicos, biológicos ou sociais embebidos nas redes.

As simulações computacionais auxiliam os modelos matemáticos na representação dos fenômenos epidemiológicos e na construção e validação de teorias sobre sistemas biológicos complexos que envolvam a propagação de doenças, já que experimentos de campo não podem ser conduzidos em epidemiologia [2]. Com isso, neste trabalho os processos epidêmicos epidêmicos são tratados no ponto de vista computacional.

#### **4.1.1 Classificação dos estados**

Quando se deseja estudar doenças infecciosas, os indivíduos em uma população hospedeira podem ser divididos em algumas classes baseados em seu estado em relação a doença. Entre as classes comumente utilizadas destacam-se:

- Suscetível (S): Indivíduo saudável que não possui a doença, mas que pode contraí-la de um indivíduo infectado.
- Infectado (I): Indivíduo que esta doente e pode ou não transmitir a infecção.
- Removidos (R): Indivíduo removido seja por morte, isolamento, ou recuperado que adquiriram imunidade à doença.
- Latente (E): Indivíduo que está infectado, mas não exibe sintomas de infecção e não transmite a infecção para outros indivíduos.

#### **4.1.2 Transmissão**

De maneira geral, as doenças infecciosas pode-se ser divididas quanto a transmissão em direta (contágio) ou indireta. A transmissão direta é a transmissão cuja disseminação ocorre diretamente, através do meio físico, entre os indivíduos infectados e suscetíveis se propagando

a uma distância mediana. Já a transmissão indireta necessita de um agente transmissor ou contato físico íntimo, podendo assim se propagar a grandes distâncias [11]. Neste trabalho é investigado o processo de transmissão direta.

### **4.1.3 O processo infeccioso**

No caso de infecções com transmissão direta, o processo infeccioso é iniciado com o contato de um indivíduo suscetível com um outro indivíduo infectado portador do agente infeccioso. O indivíduo antes suscetível passa a ser designado latente, ou seja, indivíduo infectado que não é transmissor da doença, devido ao fato do vírus ou bactéria estar aumentando sua concentração do organismo do hospedeiro.

Após o período de latência ocorre o período infeccioso, no qual, o indivíduo passa a ser denominado infectante ou infeccioso, já que teve contato com o agente infeccioso e conseguiu sobreviver durante o período de latência do mesmo, passando a eliminar tal agente para o ambiente. Posteriormente, a produção de anticorpos específicos em combate ao agente invasor ocorre, o que resulta na eliminação do mesmo de seu organismo. Por fim, tem-se o período de imunidade ou imune, o qual, em geral, é permanente para aquele agente infeccioso específico e os indivíduos passam a ser denominados recuperados ou imunes [2][11].

### **4.1.4 Incidência e prevalência**

Conceitos também importantes em epidemiologia são a incidência e a prevalência da doença em relação a população. A incidência pode ser definida como a taxa de aparecimento de novos casos da doença na população em estudo, enquanto que a prevalência é a proporção infectada da população hospedeira, ou com algum marcador da infecção passada ou presente em um determinado período de tempo.

### **4.1.5 Imunidade**

Quando um hospedeiro suscetível não é encontrado pelo agente transmissor da doença o processo infeccioso pode ser interrompido. Isso pode ocorrer quando existe uma elevada proporção de indivíduos imunes ao agente na população.

A imunidade coletiva é definida com a resistência de uma população ou grupo à introdução



e espalhamento de um agente infeccioso e é baseada em uma proporção elevada de indivíduos imunes entre os membros desse grupo ou população e na distribuição uniforme desse indivíduos imunes a doença [2].

#### **4.1.6 Modelos compartimentais**

Os modelos compartimentais dividem a população em compartimentos, onde cada uma das classes de indivíduos é um compartimento. A cada instante de tempo, um indivíduo deve pertencer a um dado compartimento.

Em um modelo dinâmico, especificando os estados e as regras de transição possíveis entre os estados, os indivíduos podem mudar de compartimento conforme o novo estado adquirido. Entre os modelos compartimentais mais conhecidos formulados a partir das classes suscetíveis (S), infecciosos (I), latentes (E), e removidos (por imunização ou morte) (R) são o Suscetível-Infeccioso-Suscetível (SIS), Suscetível-Infeccioso-Removido (SIR) e o Suscetível-Latente-Infeccioso-Removido (SEIR), mostrados de forma esquemática na Figura 4.1. Existem ainda outros modelos mais sofisticados, mas que podem ser interpretados como variantes dos modelos anteriores [2].

## **4.2 Modelo SIR**

O modelo SIR é um dos modelos epidemiológicos mais simples de propagação de doenças em uma população. Neste modelo compartimental os indivíduos são divididos nas classes:

- **Suscetíveis (S):** quando o indivíduo não possui a doença, mas pode se contaminar caso entre em contato com um indivíduo com a doença. Quando isso ocorre, o indivíduo se torna infectado, ou seja, não existe um período de latência da doença no organismo do mesmo.
- **Infectados (I):** quando o indivíduo possui a doença e pode transmiti-lá a outros indivíduos com os quais entre em contato. Quando isso ocorre desde o momento em que foi contaminado, o indivíduo infectado é também chamado de infeccioso.
- **Removidos ou Recuperados (R):** quando o indivíduo se recuperou da doença e possui imunidade permanente, ou seja, não pode ficar infectado novamente, ou, quando o

indivíduo morreu após a infecção.

No modelo matemático para o modelo SIR, um indivíduo suscetível possui uma probabilidade  $\lambda$  de contrair a doença por meio de um indivíduo infectado a cada etapa de tempo. Já um indivíduo infectado se recupera e se torna imune com uma taxa  $\gamma$  constante. Assim no modelo SIR, as quantidades indicadas na figura 4.1 são  $\varphi = \omega = \delta = 0$  e  $\lambda \neq 0$  e  $\gamma \neq 0$ .

Em geral, as taxas  $\lambda$ ,  $\gamma$ ,  $\omega$ ,  $\delta$  e  $\varphi$  marcadas na figura 4.1 são tratadas como constantes, porem, modelos mais gerais podem ser construídos com taxas variáveis. Alguns outros modelos, diferentes do modelo SIR, levam em conta que um indivíduo removido torna-se suscetível a uma taxa  $\delta$ . Já, no modelo SIS, por exemplo, um indivíduo infectado volta a ficar suscetível seguindo uma taxa  $\omega$ . A Figura 4.1 extraída de Antoniazzi [2] ilustra um Diagrama de estados genérico a partir do qual diferentes modelos compartimentais podem ser formulados.

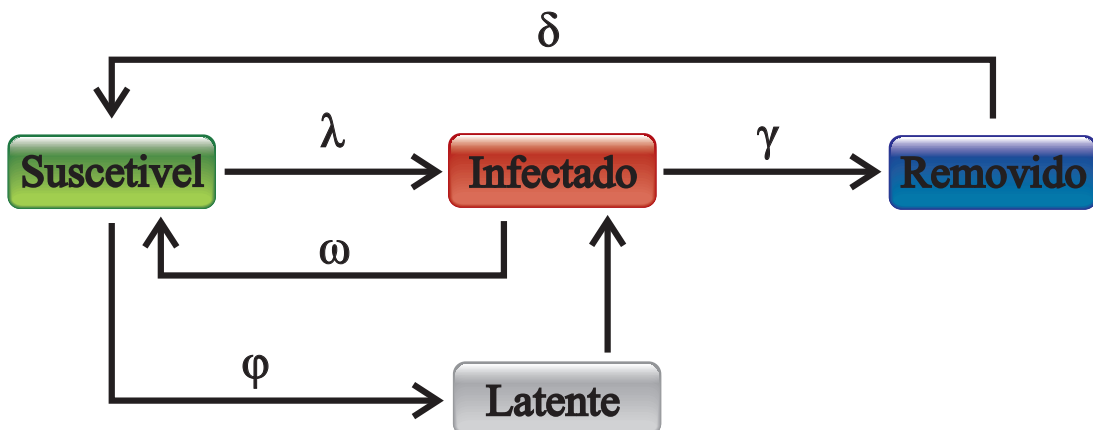


Figura 4.1: Diagrama de estados de diferentes Modelos Compartimentais. Dependendo do “caminho” escolhido, diferentes modelos podem ser formulados [2]. As letras gregas representam as taxas de transição entre os estados.

O modelo SIR relaciona-se com doenças onde os indivíduos infectados podem se recuperar adquirindo imunidade permanente. Doenças como rubéola, varicela, sarampo e caxumba que ocorrem com maior frequência na infância podem ser modeladas através de modelos SIR [11]. Partindo da introdução de uma doença em uma população suscetível, a investigação do modelo SIR consiste no acompanhamento das transições dos indivíduos entre os estágios (ou estados) da doença que se espalha seguindo as taxas  $\lambda$  e  $\gamma$  acima descritas.

A Figura 4.2 demonstra o comportamento de uma epidemia hipotética descrita pelo modelo SIR. No eixo das ordenadas temos a fração de indivíduos e no eixo das abscissas o número de passos de tempo em que a simulação ocorreu (75 iterações). Os indivíduos da população em estudo são inicialmente suscetíveis a doença. Em seguida uma pequena fração é infectada e o sistema evolui conforme a descrição do modelo SIR. O número de indivíduos suscetíveis decai rapidamente com o espalhamento da epidemia que faz com que o número de indivíduos infectados cresça rapidamente. Porém, conforme os indivíduos infectados vão se recuperando, percebe-se que a fração de indivíduos não suscetíveis (fração de indivíduos infectados + fração de indivíduos removidos) faz com que a epidemia perca força, fazendo com que ela desapareça quando não existem mais indivíduos suscetíveis na população. Um ponto importante a ser notado é que neste caso a epidemia infectou todos os indivíduos da população e por isso todos eles, ao final da simulação, se tornaram removidos.

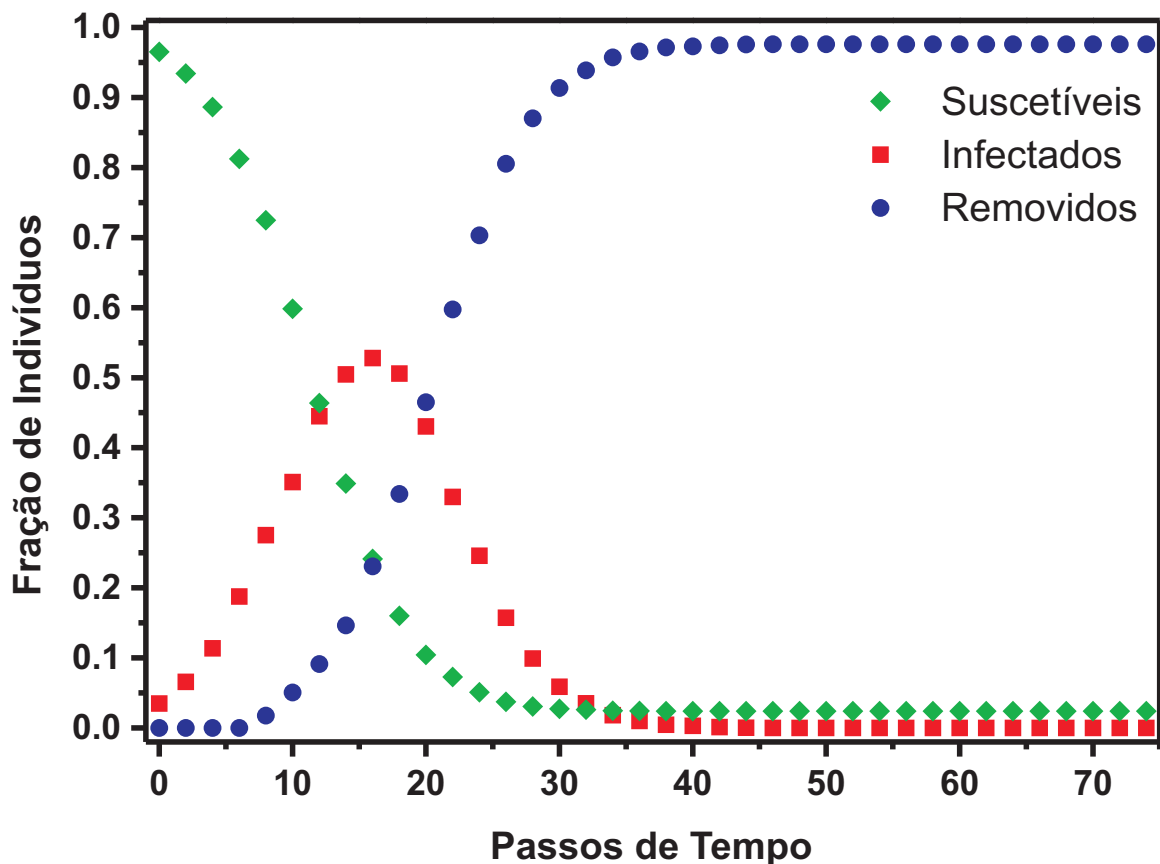


Figura 4.2: Simulação de uma epidemia com um modelo SIR determinístico. Os losangos representam a fração de indivíduos suscetíveis, os quadrados a fração de indivíduos infectados e os círculos a fração de indivíduos recuperados.

No exemplo acima, quando não existem mais indivíduos suscetíveis na população a epidemia desaparece e o estado final de qualquer indivíduo é removido, ou seja, todos os indivíduos foram infectados. Existem situações onde a epidemia desaparece mesmo existindo indivíduos suscetíveis, efeito esse ocasionado pela diminuição na quantidade de indivíduos infectados. Para que a epidemia possa se sustentar é necessária uma fração mínima de indivíduos infectados na população, fração essa chamada de número de reprodução basal  $R_0$ . Em modelos determinísticos,  $R_0$  pode ser expresso como a razão entre as taxas de transmissão e recuperação  $R_0 = \frac{\lambda}{\gamma}$ . Como a quantidade  $\frac{\lambda}{\gamma}$  determina o espalhamento da doença, é comum trabalhar com uma taxa efetiva de transmissão dada por  $\lambda_{ef} = \frac{\lambda}{\gamma}$  [2].

Quando se utiliza sistemas representados por grafos aleatórios, nota-se a existência de um limiar de propagação da epidemia  $\lambda_{min}$ . Caso a taxa de propagação efetiva  $\lambda_{ef}$  seja maior que o limiar  $\lambda_{min}$ , a infecção se espalha pela população tornando-se persistente. No entanto, se  $\lambda_{ef}$  for menor que  $\lambda_{min}$ , a infecção desaparece rapidamente e a epidemia não é caracterizada.

Analisando a propagação de epidemias em redes sem-escala, observa-se a não existência do limiar de epidemia, ou seja,  $\lambda_{min} \rightarrow 0$ . Sem o limiar, a doença persiste a qualquer taxa de espalhamento, fazendo com que redes sem-escala se tornem propensas a propagação de epidemias [2].

No próximo capítulo é descrito o modelo específico utilizado para as simulações em redes modulares espaciais. O modelo específico foi construído através das características gerais do modelo SIR padrão, além de serem inseridas características relacionadas a disposição espacial dos vértices a fim de verificar como a distribuição espacial influencia no espalhamento da epidemia na população representada por uma rede sem-escala.

# Capítulo 5

## Análise do modelo SIR sobre redes modulares espaciais

Neste capítulo é apresentado o modelo epidêmico implementado sobre redes sem escala organizadas em comunidades espacialmente distribuídos.

A seção 5.1 relembra alguns conceitos e informações que serão úteis no decorrer do capítulo enquanto na seção 5.2 e 5.3 são apresentadas as informações dos detalhes da simulação e análise dos dados.

### 5.1 A disposição espacial dos indivíduos

Para explicar como a disposição espacial dos vértices foi feita, considera-se uma população de  $N$  indivíduos separados em  $M$  comunidades. Esta população será representada por um grafo introduzido em uma matriz quadrada de lado  $L$ , sendo  $N$ ,  $M$  e  $L$  parâmetros de inicialização do modelo implementado.

O processo de criação da rede modular espacial inicia-se com uma pequena quantidade de vértices ( $n_0$ ) cada comunidade contendo  $\frac{n_0}{M}$  indivíduos. Posteriormente, é feita a inserção de outros indivíduos até que um total de  $N$  vértices seja atingido, conforme descrito no capítulo 3.

Quando um indivíduo é inserido na rede, é feita inicialmente, a seleção da comunidade a qual ele irá pertencer. No modelo implementado a escolha da comunidade a qual o vértice pertencerá influenciará na sua disposição espacial. Isso ocorre pois considera-se que cada comunidade estará disposta em uma porção espacial quadrada da matriz espacial da rede.

A figura 5.1 mostra a disposição das porções para  $M$  comunidades em relação a toda a rede.

Assim cada comunidade ocupa uma área de  $\left(\frac{L}{M}\right)^2$

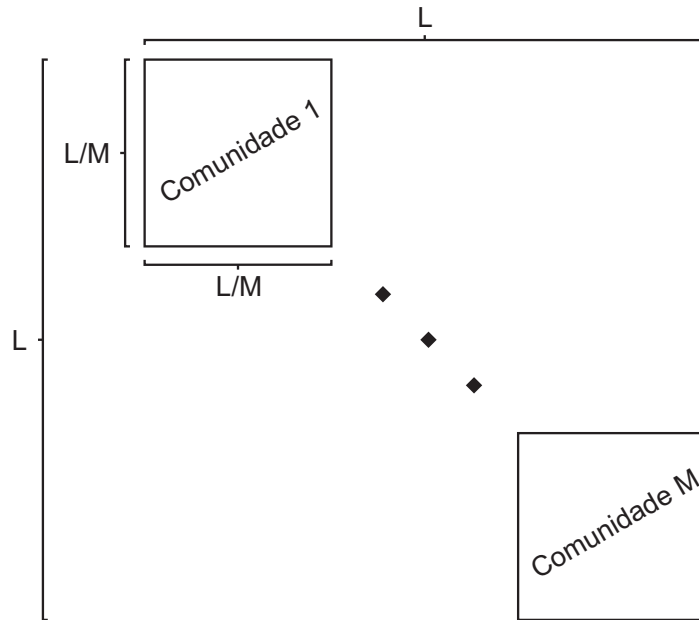


Figura 5.1: Exemplo de disposição espacial de  $M$  comunidades em relação a matriz espacial da rede.

A figura 5.2 mostra a disposição das porções espaciais de cada uma das comunidades para  $M = 2$ , o qual foi o caso investigado neste trabalho.

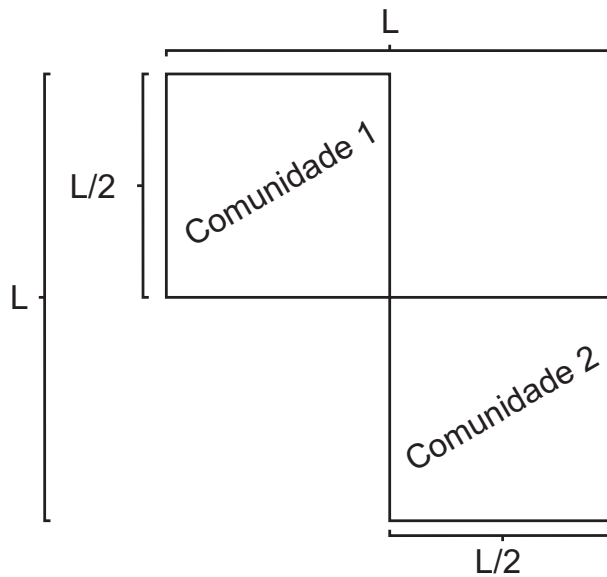


Figura 5.2: Exemplo de disposição espacial para  $M = 2$  em relação a matriz espacial da rede.

Durante a fase de construção da rede quando um vértice é inserido a sua posição espacial dependerá da localização da porção da comunidade a que ele pertença. Considerando que o indivíduo pertença a comunidade  $a$  de modo que  $(1 \leq a \leq M)$  tem-se que ele estará disposto na posição  $(x, y)$  para  $((a - 1) \cdot L/M \leq x, y \leq a \cdot L/M)$  ou seja, dentro da porção espacial da comunidade  $a$ .

O vértice é conectado a  $m(1 \leq m \leq m_0)$  vértices da mesma comunidade, caracterizando as conexões intra-comunidade e também pode se conectar com indivíduos de outras comunidades. A existência de conexões inter-comunidades depende de uma probabilidade  $\alpha$  uniforme. Caso existam as conexões inter-comunidades o vértice se conecta a  $n(1 \leq n \leq M)$  vértices de outras comunidades.

Uma nova conexão é recebida pelo vértice  $i$  com uma probabilidade  $p_{ij} = \frac{p_j^\sigma}{d_{ij}^\chi}$ , sendo  $p_j$  a fração de arestas do vértice  $j$  elevada a um expoente de controle  $\sigma$  e  $d_{ij}$  a distância entre dois vértices  $i$  e  $j$  elevada a um expoente  $\chi$ , conforme descrito no capítulo 3.

## 5.2 Distribuição espacial dos vértices

A fim de verificar como a doença se espalha espacialmente dentro de cada comunidade e entre as comunidades optou-se por classificar os vértices conforme a distância ocupada em relação ao centro geométrico da comunidade. Cada comunidade foi dividida em faixas de largura  $r$  e o estado de cada indivíduo é mostrado com função do tempo e da posição ocupada. O número de faixas utilizadas (*faixas*) é um dos parâmetro da simulação e pode ser controlado. A figura 5.3 mostra a separação da porção de uma comunidade para *faixas* = 4.

O cálculo da faixa a qual o individuo pertence é feita a partir do raio de abrangência da comunidade  $raio_a$  o qual fornece uma medida da dimensão linear da comunidade. Qualquer elemento da comunidade  $a$  estará em uma posição  $centro_a < r < raio_a$ .

Com o auxilio da figura 5.4 obtêm-se a equação 5.1:

$$\begin{aligned} raio_a^2 &= b^2 + c^2 = (L/2M)^2 + (L/2M)^2 = 2 \cdot (L/2M)^2 \\ raio_a &= \sqrt{2 \cdot (L/2M)^2} \\ raio_a &= \sqrt{2} \cdot L/2M \end{aligned} \tag{5.1}$$

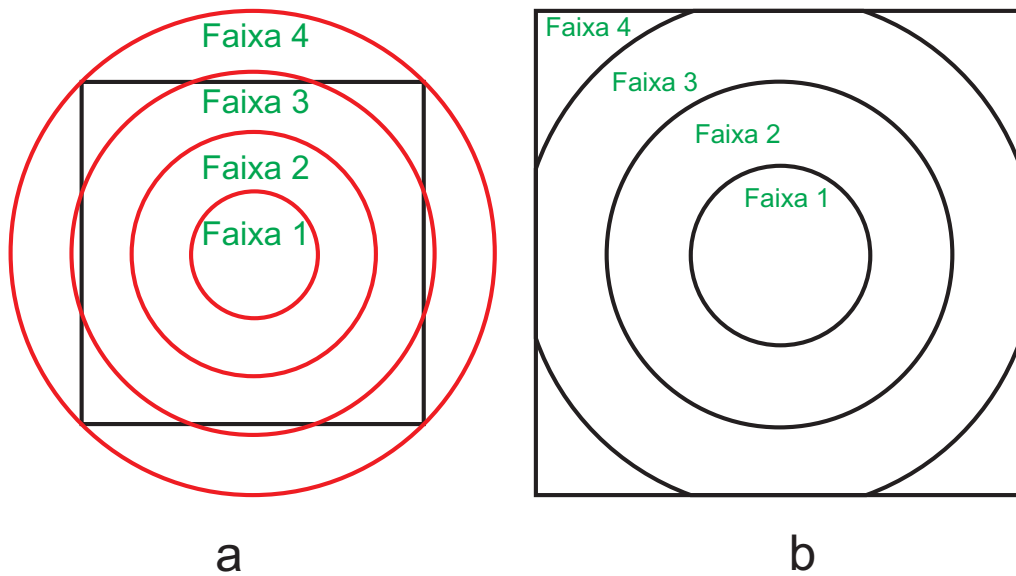


Figura 5.3: Porção espacial de uma comunidade dividida em 4 faixas

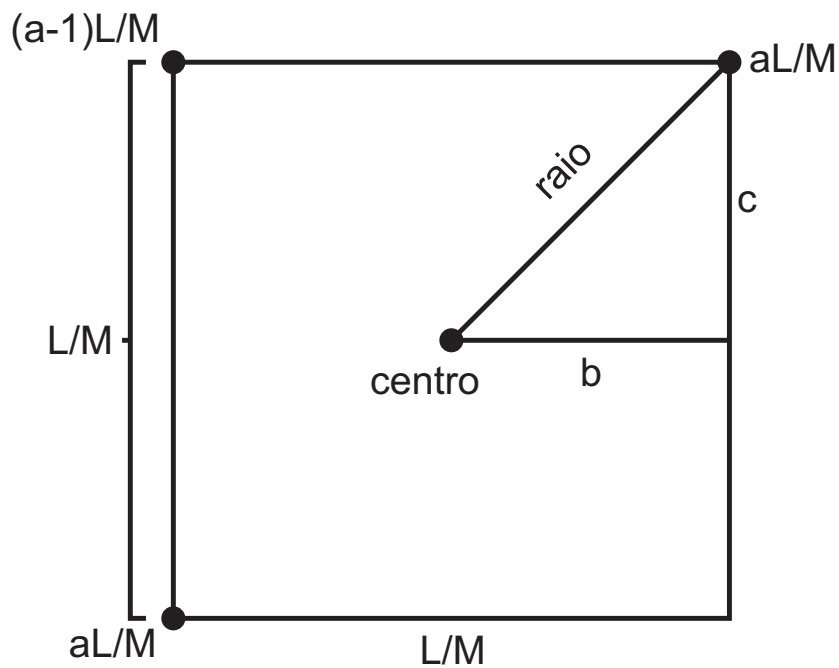


Figura 5.4: Informações sobre a porção espacial de uma comunidade

Para saber a qual faixa espacial um vértice  $i$  pertence basta obter o resultado da divisão inteira entre a distância do vértice  $i$  ao centro da porção espacial de sua comunidade e a fração pertinente a cada faixa. Novamente com o auxílio da figura 5.4 temos que o centro de uma



comunidade  $a$  pode ser calculado por:

$$\begin{aligned}
2 \cdot \text{centro}_a^2 &= b^2 + c^2 = \left( \frac{(a-1) \cdot L/M + a \cdot L/M}{2} \right)^2 + \left( \frac{(a-1) \cdot L/M + a \cdot L/M}{2} \right)^2 \\
&= \left( \frac{L/M \cdot (2a-1)}{2} \right)^2 + \left( \frac{L/M \cdot (2a-1)}{2} \right)^2 = 2 \cdot \left( \frac{L/M \cdot (2a-1)}{2} \right)^2 \\
\text{centro}_a^2 &= \left( \frac{L/M \cdot (2a-1)}{2} \right)^2 \\
\text{centro}_a &= \sqrt{\left( \frac{L/M \cdot (2a-1)}{2} \right)^2} = \frac{L/M \cdot (2a-1)}{2} \\
\text{centro}_a &= \frac{L(2a-1)}{2M} \tag{5.2}
\end{aligned}$$

Com o centro da comunidade pode-se agora calcular a distância entre um vértice  $i$  qualquer, pertencente a comunidade, e o centro da comunidade. Dados os ponto  $x_1$  e  $y_1$  pertencentes ao vértice  $i$  e os pontos  $x_2$  e  $y_2$ , tem-se a distância entre esses dois pontos  $d$  como:

$$\begin{aligned}
d^2 &= (x_1 - x_2)^2 + (y_1 - y_2)^2 \\
d &= \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \tag{5.3}
\end{aligned}$$

Devido ao fato de que a captura dessa informações por faixas possa se tornar muito custosa dependendo da quantidade de indivíduos na população, nem sempre será calculada essa informação em cada passo de tempo da simulação. Com isso, o monitoramento do estado de cada indivíduo pertencente a cada faixa é feito em intervalos de tempo  $\Delta t$ .

### 5.3 O modelo SIR aplicado a rede modular espacial

A figura 5.5 apresenta o esquema da simulação do modelo SIR aplicado a rede modular espacial, e como a obtenção das informações sobre as faixas é feita durante todo o processo de simulação.

Após a construção da rede modular espacial, todos os indivíduos da rede são inicializados como suscetíveis a doença, ou seja, podem ser infectados caso haja um contato com algum vizinho que esteja infectado. Após essa inicialização é feito o cálculo da quantidade de indivíduos em cada uma das faixas espaciais.

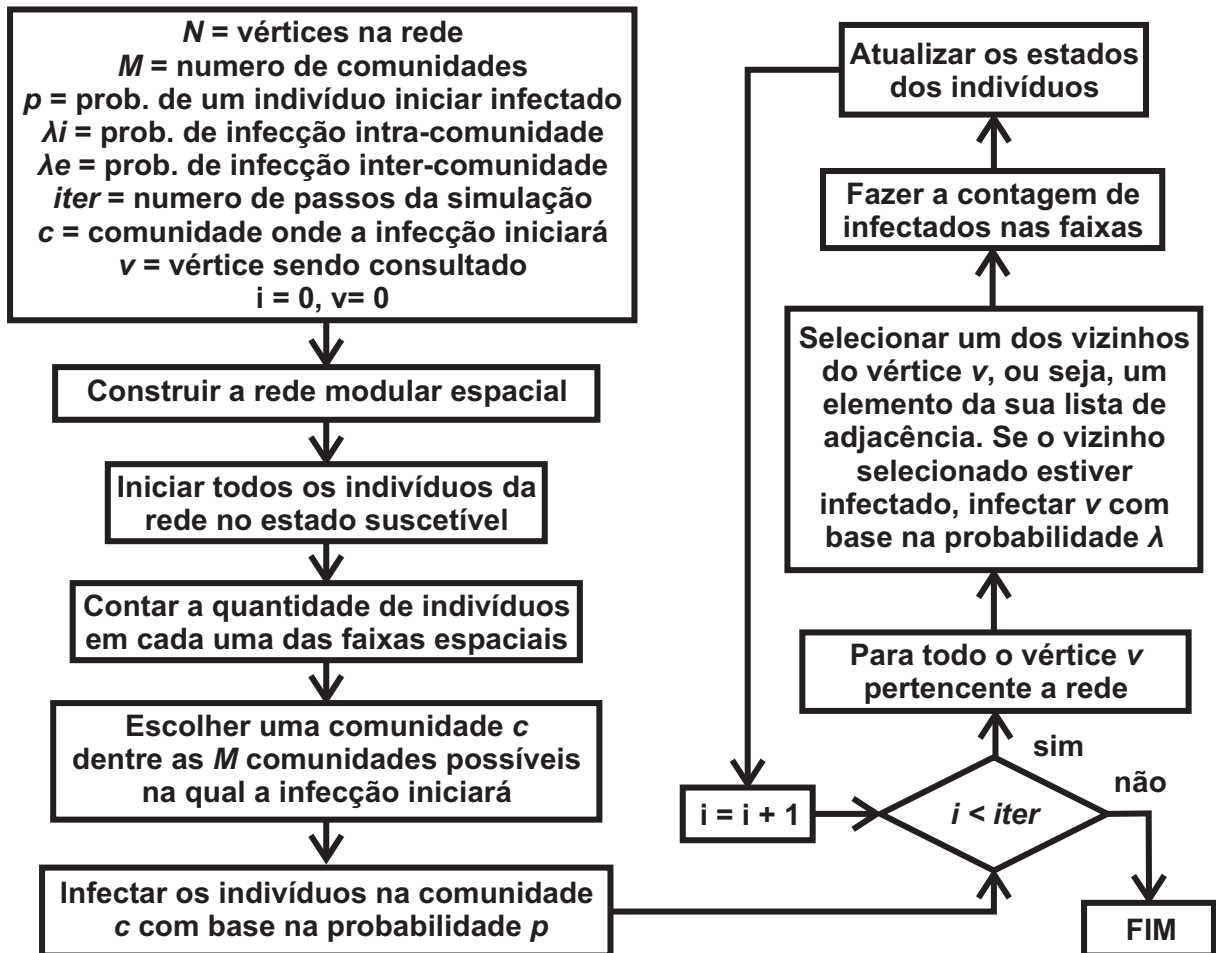


Figura 5.5: Simulação do Modelo SIR em uma rede modular espacial

O processo de infecção dos indivíduos inicia-se sempre em uma comunidade  $c$  qualquer, escolhida aleatoriamente dentre as  $M$  comunidades da rede. A distribuição dos vértices inicialmente infectados pode ser aleatória ou centralizada em uma das faixas espaciais. No caso da infecção aleatória os indivíduos da comunidade  $c$  são infectados com probabilidade  $p$  que é um dos parâmetros do modelo. Já quando a infecção é centralizada em uma das faixas espaciais, somente os indivíduos dessa faixa podem ser infectados com probabilidade  $p$ .

Após esse processo inicial de infecção dos indivíduos, a simulação irá acontecer por um determinado período de tempo (*iterações*). Em cada uma dessas iterações cada um dos indivíduos será consultado a fim de que possa ser definido qual será o seu estado na próxima iteração da simulação.

Nessa consulta um dos seus vizinhos, ou seja, um elemento da sua lista de adjacências, é escolhido aleatoriamente. No caso desse vizinho estar infectado então o vértice que está sendo

consultado poderá ser infectado com probabilidade  $\lambda$ . Essa infecção do indivíduo consultado ocorre com base na taxa de infecção de indivíduos de uma mesma comunidade  $\lambda_I$  (infecção intra-comunidade) ou na taxa de infecção de indivíduos de comunidades diferentes  $\lambda_E$  (infecção inter-comunidade). No modelo SIR o indivíduos permanecem no estado infectado por um determinado tempo denominado  $T_I$  e no estado removido por um tempo indeterminado.

Após a consulta a todos os vértice da rede, ocorre o cálculo do número de indivíduos infectados em cada uma das faixas espaciais definidas na seção anterior. Esse cálculo da quantidade de indivíduos infectados em cada um das faixas espaciais, para cada comunidade, consiste em passar por todos os indivíduos da comunidade, e caso o indivíduo esteja infectado, definir a qual das faixas espaciais ele pertence atualizando a quantidade naquela faixa.

Após o cálculo da quantidade de indivíduos infectados por faixa espacial, o último passo de uma iteração consiste em atualizar os estados de todos os vértices da rede. Esse processo é repetido até que o número de iterações da simulação se esgote.

## 5.4 Resultados Obtidos

Para verificar como a modularidade influencia o espalhamento da doença na rede e como isso ocorre espacialmente, simulações foram feitas com redes contendo  $N = 6000$  indivíduos, divididos em  $M = 2$  comunidades e distribuídos em um espaço de tamanho  $L = 100$ .

Nas próximas seções são discutidos o espalhamento da epidemia como função do tempo e a influência da modularidade sobre o processo, bem como o espalhamento espacial da epidemia intra e inter comunidade. A seção 5.4.1 mostra os resultados obtidos com a variação do valor de  $\alpha$  que influencia fortemente na separação das comunidades. Já a seção 5.4.2 mostra como o espalhamento da epidemia ocorre espacialmente através das faixas em uma única rede e no caso de médias entre simulações.

### 5.4.1 Influência da Modularidade

A modularidade influencia no processo de espalhamento da epidemia uma vez que fornece uma medida da interação entre as diferentes comunidades. Antoniazzi [2] investigou o comportamento da epidemia nas redes modulares que não levam em consideração a posição (ou disposição) espacial dos indivíduos. Com base nisso o propósito desse trabalho consiste em

avaliar esse comportamento nas redes modulares espaciais, comparando os estados.

Conforme mostrado no capítulo 3 o fator de modularidade não é afetado pela disposição espacial das comunidades. Desta forma os resultados globais da epidemia não se diferenciam daqueles observados por Antoniazzi [2], conforme mostrado abaixo.

No processo de simulação da epidemia a doença inicia em uma das comunidades da rede, enquanto que a outra comunidade permanece totalmente suscetível. Como as simulações foram feitas com  $M = 2$ , a comunidade onde a infecção iniciará é denominada *Infectada*, enquanto que a outra comunidade é denominada *Não Infectada*. A nomenclatura *Média* indica a fração de indivíduos em um estado  $k$  na rede toda.

A figura 5.6 apresenta a fração de indivíduos infectados na rede em relação ao tempo para simulações com diferentes valores de  $\alpha$ .

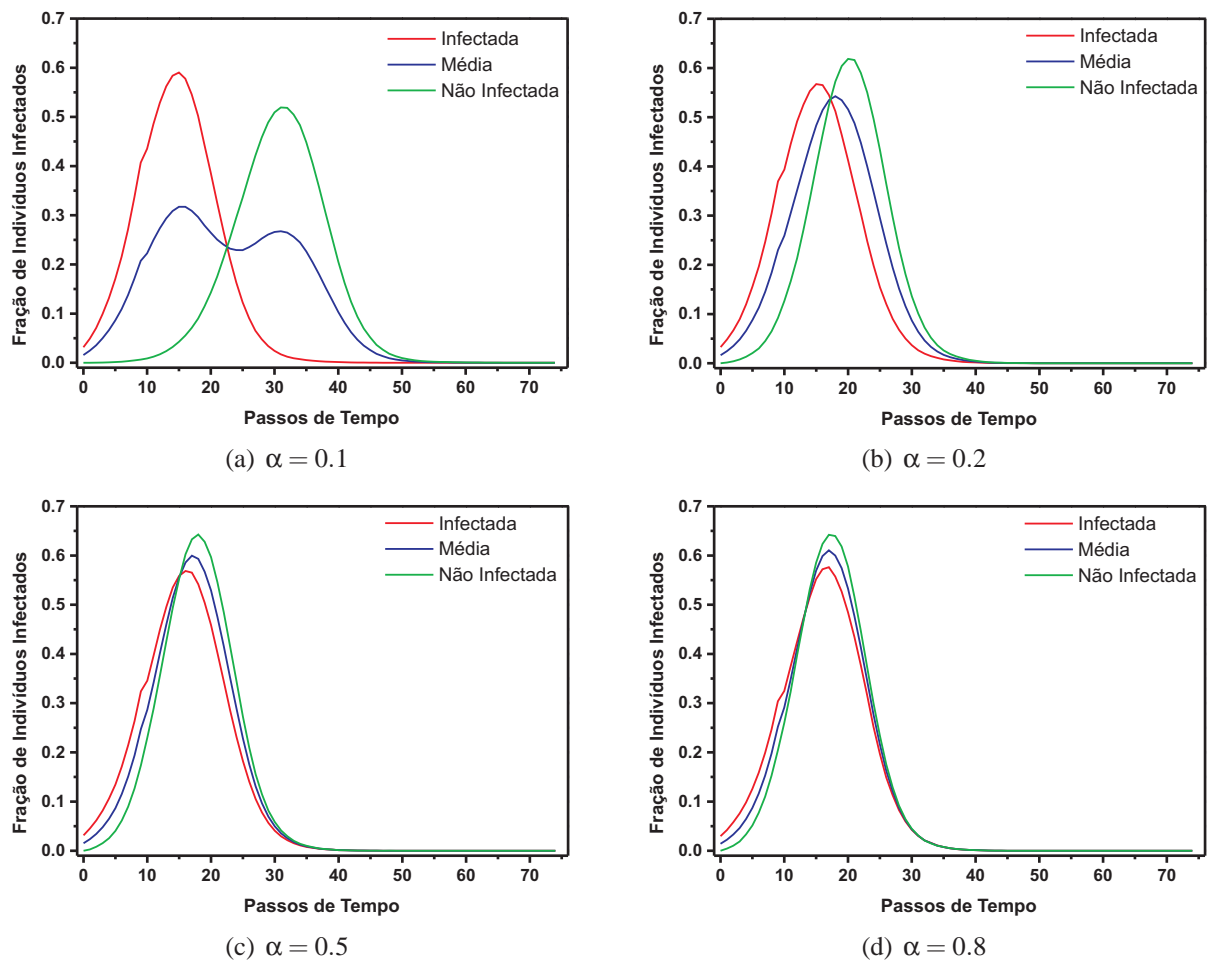


Figura 5.6: Dinâmica da epidemia para redes com  $N = 6000$  elementos,  $M = 2$  comunidades,  $m = 4$ ,  $n = 2$  e diferentes valores para  $\alpha$

Nota-se que o valor de  $\alpha$  influencia fortemente no espalhamento da epidemia de uma comunidade para outra. Isso ocorre devido ao fato de que quanto maior o valor de  $\alpha$  maior será a interação entre as comunidades (maior valor de  $Q$ ), ou seja, uma maior quantidade de indivíduos possuirá ligações (ou conexões) com indivíduos de comunidades diferentes da sua. Logo, percebe-se que essa quantidade de ligações *inter-comunidade* define em qual momento as comunidades atingirão o seu pico de infecção.

Quando o valor de  $\alpha$  é pequeno as comunidade estão, de certa forma, bem separadas, ou seja, são pouco interagentes. Nota-se na figura 5.6(a) que os picos de indivíduos infectados nas comunidades *Infectada* e *Não Infectada* estão deslocados. Com isso, observa-se na *Média* dos infectados que existem dois picos, os quais são resultantes da sobreposição das epidemias que estão ocorrendo em cada comunidade.

Na figura 5.6(b) os picos das comunidades estão mais próximos, ou seja, devido ao fato de uma maior interação (diminuição de  $Q$ ) a doença atingiu a comunidade *Não Infectada* mais rapidamente que para  $\alpha = 0, 1$ . Isso pode ser notado também nas figuras 5.6(c) e 5.6(d), onde as comunidades estão mais entrelaçadas e a epidemia se espalha quase que semelhantemente em ambas as comunidades.

## 5.4.2 Comportamento em redes espaciais

Nesta seção são apresentados os resultados obtidos em simulações com médias sobre 10 repetições, com o propósito de verificar o comportamento da epidemia espacialmente. Vale ressaltar que nos resultados aqui apresentados a epidemia inicia-se sempre na faixa espacial central (faixa 1) de uma das comunidades, já que pretende-se investigar o espalhamento espacial da epidemia.

A figura 5.7 mostra o comportamento médio da fração de vértices infectados na comunidade *Infectada* para simulações com diferentes valores de  $\alpha$ .

Percebe-se facilmente que pelo fato da doença estar inicialmente localizada na faixa 1 ela atinge seu pico antes das outras. É interessante notar que nas faixas espaciais restantes a fração de indivíduos infectados é semelhante durante toda a simulação, ou seja, para esse caso o espalhamento da epidemia nessa faixas é similar.

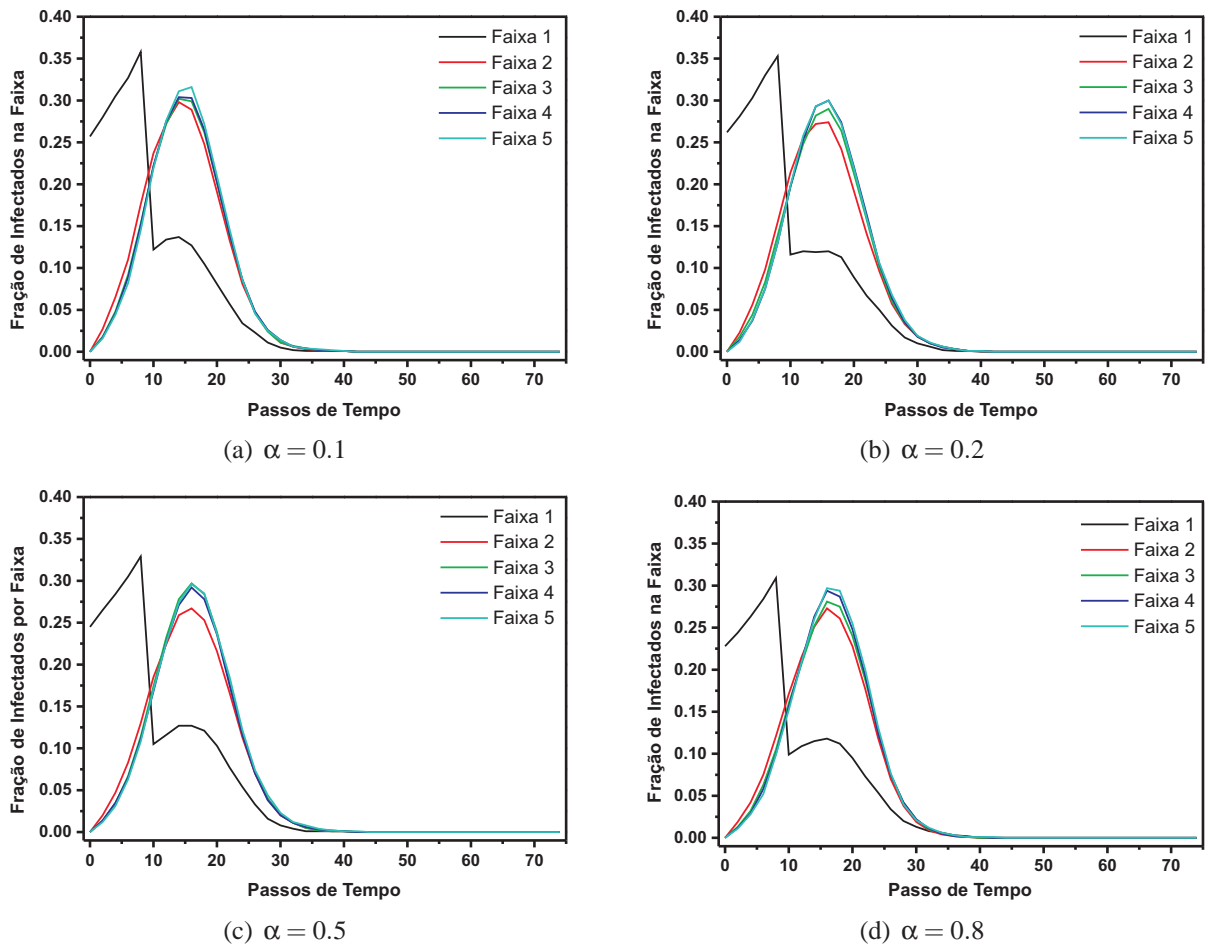


Figura 5.7: Espalhamento da epidemia nas faixas da comunidade *Infectada* para redes com  $N = 6000$  elementos,  $M = 2$  comunidades,  $m = 4$ ,  $n = 2$ ,  $L = 100$ ,  $faixas = 5$  e diferentes valores para  $\alpha$  para 10 repetições. O índice da faixa indica a distância em relação ao centro geométrico da comunidade.

Devido ao fato de que a epidemia inicia na faixa mais central da comunidade, esperava-se que quanto maior a distância da faixa em relação ao centro geométrico da comunidade, mais tempo levaria para que a infecção atingisse indivíduos nela dispostos. Porém como visto na figura 5.7 isso não ocorre, ou seja, a um espalhamento similar. Isso ocorre devido a existência das ligações de longo alcance dentro da comunidade, assim, como os indivíduos da faixa central podem estar conectados a outros indivíduos das faixas mais externas, a infecção acaba atingindo os indivíduos dessas faixas em tempos próximos.

Na figura 5.8 observa-se o comportamento na comunidade *Não Infectada* inicialmente para simulações com diferentes valores de  $\alpha$ .

Neste caso nota-se que, quanto mais interagentes são as comunidades, mais cedo a

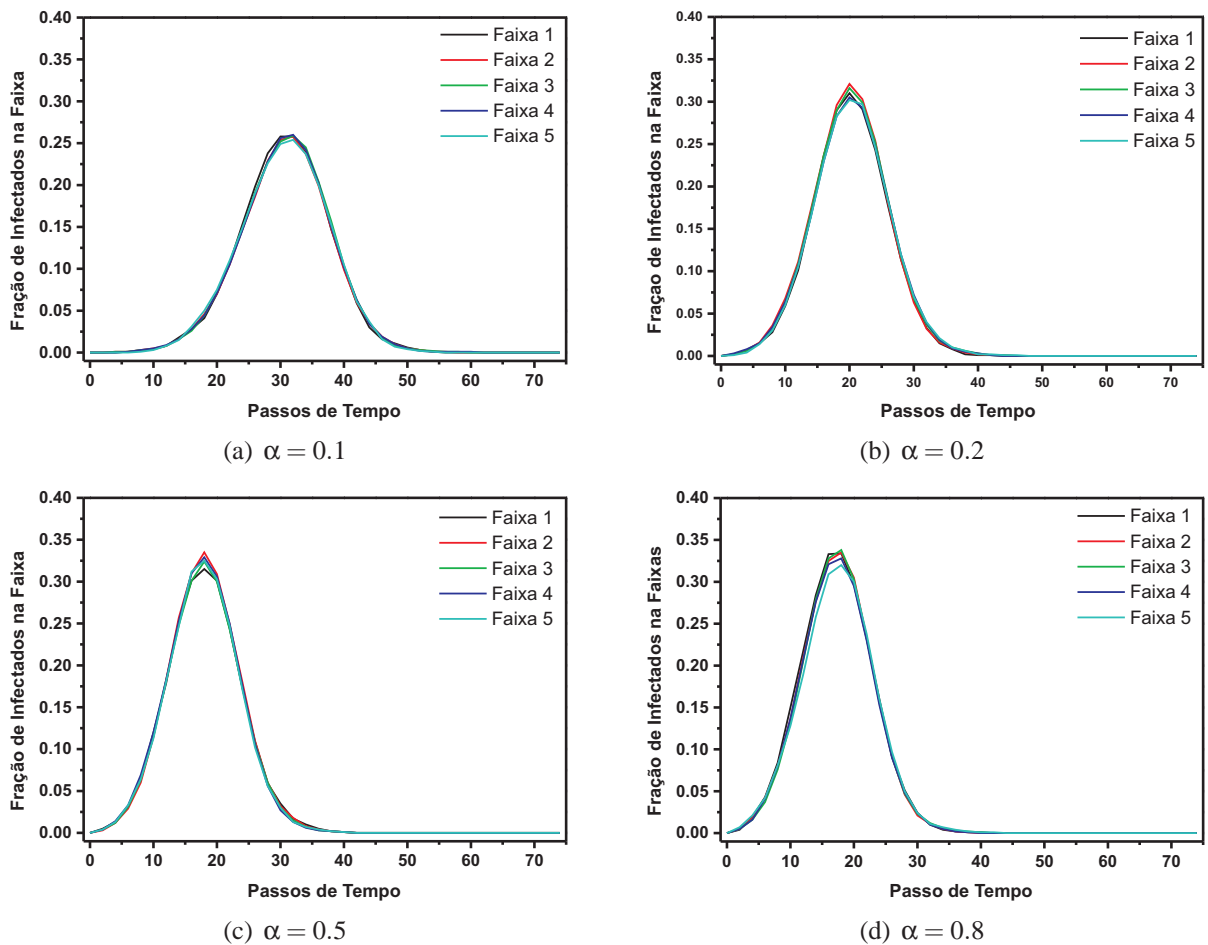


Figura 5.8: Espalhamento da epidemia nas faixas da comunidade *Não Infectada* para redes com  $N = 6000$  elementos,  $M = 2$  comunidades,  $m = 4$ ,  $n = 2$ ,  $L = 100$ ,  $faixas = 5$  e diferentes valores para  $\alpha$  para 10 repetições. O índice da faixa indica a distância em relação ao centro geométrico da comunidade.

comunidade *Não Infectada* é atingida pela epidemia. Para  $\alpha = 0.1$  o pico nas faixas está próximo do passo 30, quando  $\alpha = 0.2$  o pico é atingido em 20 passos, e para  $\alpha = 0.5$  e  $\alpha = 0.8$  isso ocorre em torno do passo 15.

Nos resultados discutidos acima foram feitas simulação com médias sobre 10 repetições. Porém o comportamento da epidemia pode ter sido influenciado pelas médias, já que a epidemia pode alcançar a comunidade inicialmente não infectada em faixas espaciais diferentes, em cada uma das repetições. Com isso, foram feitas simulação com uma única repetição a fim de verificar se o comportamento da epidemia se assemelha ao caso médio.

Vale ressaltar que, assim como no caso médio, nos resultados aqui apresentados a epidemia inicia-se sempre na faixa espacial central (faixa 1) de uma das comunidades, já que pretende-se

investigar o espalhamento espacial da epidemia.

A figura 5.9 mostra o comportamento na comunidade *Infectada* para simulações com diferentes valores de  $\alpha$ .

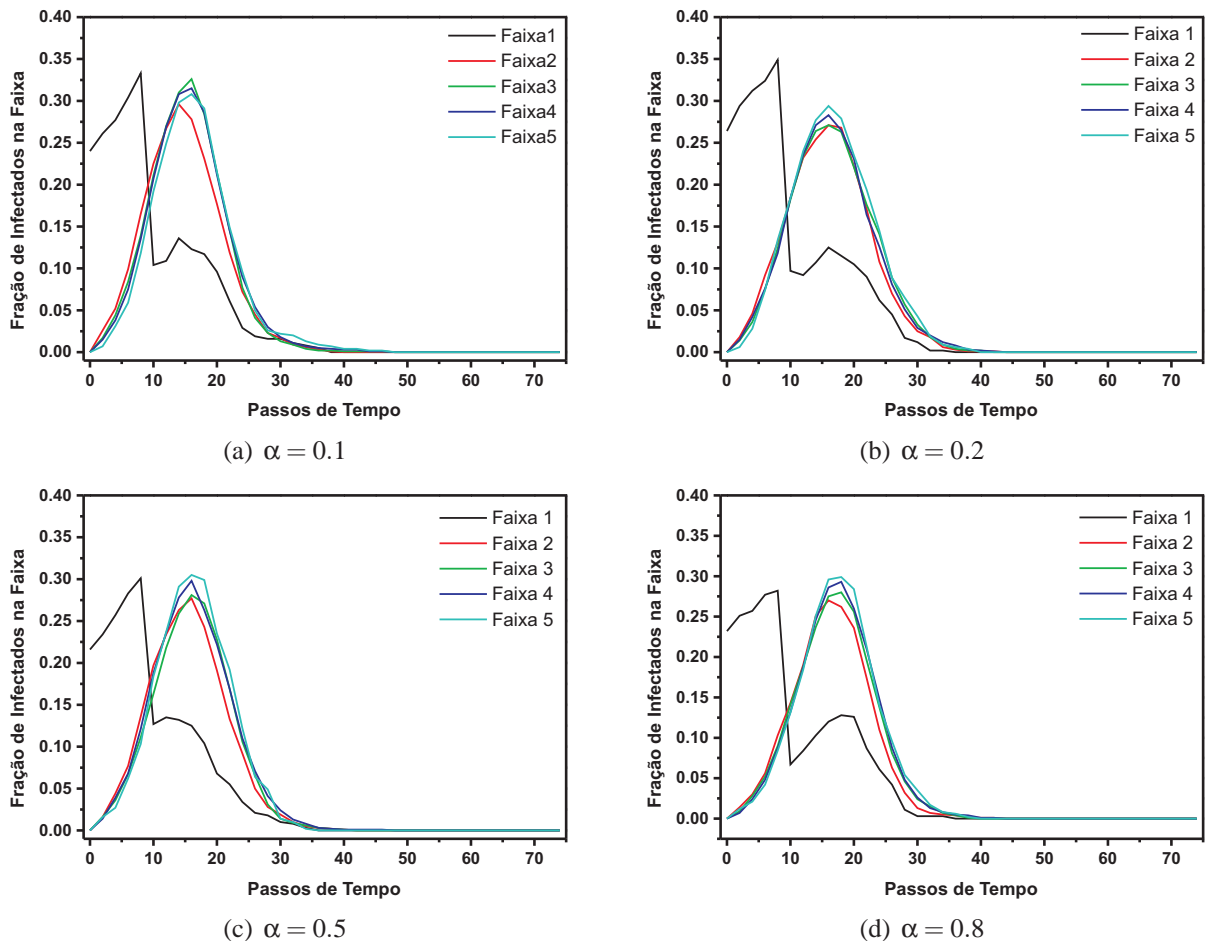


Figura 5.9: Espalhamento da epidemia nas faixas da comunidade *Infectada* para redes com  $N = 6000$  elementos,  $M = 2$  comunidades,  $m = 4$ ,  $n = 2$ ,  $L = 100$ ,  $faixas = 5$  e diferentes valores para  $\alpha$  para 1 repetição. O índice da faixa indica a distância em relação ao centro geométrico da comunidade.

O espalhamento espacial da doença entra as faixas na rede para uma única repetição, apesar das curvas menos suaves, é similar ao espalhamento no caso médio, ou seja, as médias sobre as repetições não afetaram os resultados acima mostrados.

Na figura 5.10 tem-se o comportamento na comunidade *Não Infectada* para simulações com diferentes valores de  $\alpha$ .



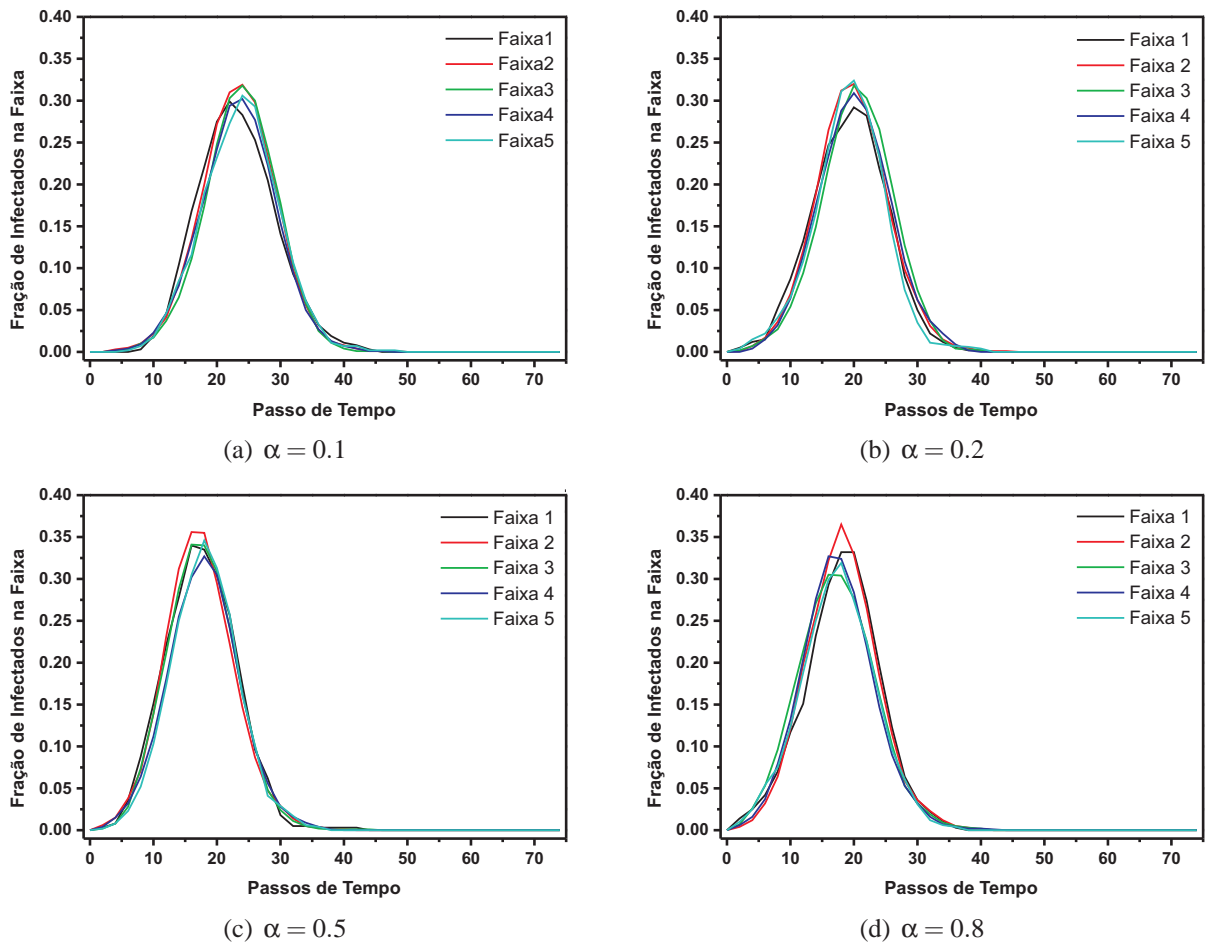


Figura 5.10: Espalhamento da epidemia nas faixas da comunidade *Não Infectada* para redes com  $N = 6000$  elementos,  $M = 2$  comunidades,  $m = 4$ ,  $n = 2$ ,  $L = 100$ ,  $faixas = 5$  e diferentes valores para  $\alpha$  para 1 repetição. O índice da faixa indica a distância em relação ao centro geométrico da comunidade.

Neste caso nota-se que, da mesma forma que para o caso médio, quanto mais próximas as comunidades estão, mais cedo a comunidade *Não infectada* é atingida pela epidemia, ou seja, o comportamento da a comunidade *Não Infectada* também não é afetado pelas médias.

## Capítulo 6

# Considerações Finais e Trabalhos Futuros

As redes complexas podem representar muitos dos sistemas naturais nos quais o ser humano está inserido, bem como, a rede de contatos dos indivíduos que nela esteja alocados. Com isso, se torna possível a investigação de vários processos dinâmicos que envolvam a população humana, como por exemplo, a propagação de informações e doenças infecciosas entre os indivíduos pertencentes a rede em estudo. Tais estudos tornam-se cada vez mais importante a investigação dos processos contidos nos agrupamentos humanos, já que essa investigação pode auxiliar na elaboração de políticas de prevenção e controle de doenças, por exemplo.

O principal objetivo desse trabalho foi investigar a propagação de epidemias em redes modulares espaciais e como isso ocorre espacialmente. Para esse fim, foi implementado um gerador de topologia de rede e, logo após, foi desenvolvido um modelo epidemiológico com o fim de simular os efeitos das redes criadas sobre a dinâmica do modelo epidêmico.

Para a implementação do gerador de redes modulares espaciais, foi criado um modelo com base no modelo de Li-Maini [9], que propicia o controle sobre o fator de modularidade, e no modelo de Yook [13] que define a disposição espacial dos indivíduos.

Para as simulações da epidemia em redes modulares espaciais, uma variação do modelo epidemiológico Suscetível-Infectado-Removido (SIR) foi desenvolvida. Isso ocorreu já que o modelo teve que ser modificado para que o mesmo pudesse ser definido sobre a rede modular espacial.

Assim como em Antoniazzi [2] os resultados revelaram que, quando uma rede possui um estrutura interna bem definida, ou seja, há uma menor interação entre as comunidades, o processo epidêmico não pode ser avaliado pela análise da população como um todo. Isto ocorre já que neste caso, as diferentes comunidades apresentam picos de infecção distintos,

logo, enquanto que em uma comunidade a epidemia está declinando, em outras pode estar se expandindo. Logo, quanto maior a interação entre as comunidades, mais semelhante será o espalhamento da epidemia em diferentes comunidades.

No que diz respeito ao espalhamento espacial da epidemia a partir de uma comunidade inicialmente infectada, percebe-se que mesmo a infecção iniciando nos indivíduos localizados na faixa central da comunidade o espalhamento para as outras faixas ocorre independentemente da distância a que o indivíduo está em relação ao ponto inicial da infecção. Com isso, na comunidade que não estava infectada inicialmente, o comportamento da epidemia é semelhante em todas as porções da rede espacial.

O estudo revelou ainda que a disposição espacial dos indivíduos, para os casos simulados, não influencia no espalhamento da epidemia a indivíduos da mesma comunidade e indivíduos de outras comunidades, ou seja, se assemelha ao caso onde a disposição espacial não é considerada e conclui-se que isto deve-se a existência de ligações de longo alcance entre indivíduos dispostos geometricamente distantes.

Entre os estudos em andamento nessa linha destaca-se “Epidemic spreading in weighted scale-free networks with community structure” [4] que baseado no modelo de doenças SIS, investigou o espalhamento da epidemia em redes ponderadas sem escala e com estrutura em comunidade.

Como trabalhos futuros sugere-se manipular a influência da quantidade de conexões de um vértices ou da distância durante o processo de criação das conexões entre os indivíduos através de seus expoentes, a fim de verificar se esses fatores poderão influenciar no espalhamento interno e externo da epidemia entre as comunidades.

# Referências Bibliográficas

- [1] ALBERT, R.; BARABÁSI, A. L. Statistical mechanics of complex networks. **Reviews of Modern Physics**, [S.l.], v.741, p.47, 2002.
- [2] ANTONIAZI, D. **Propagação de Epidemias em Redes Complexas Estruturadas**. Cascavel - PR: Universidade Estadual do Oeste do Paraná - UNIOESTE, 2008. Trabalho de conclusão de curso.
- [3] CALDARELLI, G. **Scale-Free Networks: Complex Webs in Natural, Technological and Social Sciences**. OXFORD University Press, 2006.
- [4] CHU, X. et al. Epidemic spreading in weighted scale-free networks with community structure. **Journal of Statistical Mechanics: Theory and Experiment**, [S.l.], v.07, p.7043, 2009.
- [5] DE ANGELIS, A. F. **Redes Complexas**. Tutorial. Universidade de São Paulo - Instituto de Física de São Carlos.
- [6] FALOUTSOS, M.; FALOUTSOS, P.; FALOUTSOS, C. On power-law relationship of the internet topology. **Computer Communication Review**, New York, v.29, p.251–262, 1999.
- [7] FORTUNATO, S.; CASTELLANO, C. Community structure in graphs. **[physics.soc-ph]**, [S.l.], v.arXiv:0712.2716v1, 2007.
- [8] HUANG, W.; LI, C. Epidemic spreading in scale-free networks with community structure. **Journal of Statistical Mechanics: Theory and Experiment**, [S.l.], v.P01014, 2007.

- [9] LI, C.; MAINI, P. K. An evolving network model with community structure. **Journal of Physics A: Mathematical and General**, [S.l.], v.38, p.9741, 2005.
- [10] NEWMAN, M. E. J.; GIRVAN, M. Finding and evaluating community structure in networks. **Physical Review E**, [S.l.], v.69, p.026113, 2004.
- [11] PAN, C. Z. **Epidemiologia Matemática: conceitos, processos e modelos**. Cascavel - PR: Universidade Estadual do Oeste do Paraná - UNIOESTE, 2007. Monografia.
- [12] WAXMAN, B. M. Routing of multiple connections. **Journal on selected areas in Communications**, [S.l.], v.6, n.9, p.1617, 1988.
- [13] YOOK, S. H.; JEONG, H.; BARABÁSI, A. L. Modeling the internet's large-scale topology. **PNAS**, [S.l.], v.99, n.21, Outubro, 2002.