

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
UFRGS

Dissertação de Mestrado

Dinâmica de Epidemias numa Rede Social Real

Alexsandro Marian Carvalho

Dissertação apresentada ao Programa de Pós-Graduação em Física da Universidade Federal do Rio Grande do Sul, sob orientação do Prof. Dr. Sebastián Gonçalves, como preenchimento parcial dos pré-requisitos para a obtenção do título de Mestre em Física.

Porto Alegre, Junho de 2008

Resumo

Fenômenos como a rápida propagação de epidemias em escala mundial têm chamado a atenção para a importância da estrutura da rede de contatos, através da qual os membros de uma comunidade interagem entre si. O efeito da topologia das redes de contatos sociais na dinâmica de epidemias tem sido estudado recentemente de maneira geral em redes paradigmáticas (mundo pequeno, livre de escala, etc), mostrando efeitos que dão outra dimensão aos resultados típicos de campo médio, i.e, ausência de limiar de epidemia e limiar de epidemia dependente da topologia. Aproveitamos o primeiro levantamento completo de uma rede real de contatos amorosos, denominada por seus autores de romântica, para estudar a propagação hipotética de uma doença sexualmente transmissível com dinâmica SIR (suscetível, infectado e removido). Quando as interações entre os indivíduos permanecem inalteradas no tempo (rede estática), comparamos as propriedades epidêmicas entre as topologias real e paradigmáticas. Além disso, incluímos nesta comparação a variabilidade na frequência de interações entre os indivíduos. Por fim, consideramos a dinâmica real de formação de pares na rede romântica e estudamos a propagação da doença. Confrontamos as propriedades epidêmicas da rede romântica dinâmica com a versão estática. Os resultados obtidos indicam por uma lado a substancial diferença entre a rede real e a rede livre de escala, em todos os cenários. Por outra parte, sugerem que o espalhamento da epidemia é bem menos grave na rede romântica dinâmica quando comparado a versão estática dessa rede.

Abstract

Phenomenons as the fast propagation of epidemics the world scale has been getting the public's attention in general for the importance of the structure of the network of contacts through which a community's members interact amongst themselves. The effect of the topology of the networks of social contacts in the dynamics of epidemics has been studied recently in a general way in networks pragmatics (small-world, scale-free, etc), showing effects that give other dimension to the typical results of medium field, i.e, absence of threshold epidemic and threshold of epidemic dependent of the topology. We utilize the first complete rising of a real network of loving contacts, denominated by their authors of romantic, to study the hypothetical propagation of a disease sexually transmissible with dynamics SIR (susceptible, infected and removable). In the context in that the interactions among the individuals stay unaffected in the time (static), we compared the epidemic properties between the topologies real and pragmatics. Besides, we use this comparison the variability in the frequency of interactions among the individuals. Finally, we consider the real dynamics of formation of pairs in the romantic network and we study the propagation of the disease. We confront the epidemic properties in the networks romantic dynamics and static. The results obtained indicate not only the substantial difference between the real network and the scale-free network in all of the sceneries as well as they suggest that the dispersal of the epidemic is less serious in the dynamic formation of the pairs when compared the static situation in the romantic network.

Agradecimentos

Agradeço inicialmente a meu orientador prof. Sebastián Gonçalves, cujos conhecimentos e habilidades não só enriqueceram esta dissertação como também me motivaram a execução da mesma. Quero destacar minha gratidão a minha esposa, Viviane, cuja paciência e dedicação se prolongou por muito mais tempo do que exigia a redação destas não muitas linhas. Agradeço também a meu amigo e colega Vilarbo cujos incentivos em minhas jornadas de estudos estende-se desde da época de graduação. Finalizo agradecendo a meus pais, Cleci e Mauri, por todo o carinho.

Sumário

1	Introdução	6
2	Introdução às Redes Complexas	8
2.1	Conceitos Básicos	9
2.2	Medidas	12
2.3	Estruturas e Modelos	16
2.4	Redes no Mundo Real	21
3	Elementos de Epidemiologia Matemática	24
3.1	Breve Digressão Histórica	24
3.2	Modelo SIR Padrão	26
3.3	Epidemias em Redes Complexas	31
4	Dinâmica SIR nas Redes Paradigmáticas e Romântica	36
4.1	Rede Romântica	37
4.2	Redes Estáticas	38
4.3	Propriedades Estruturais das Redes Estáticas	38
4.4	Simulação Computacional	41
4.5	Resultados	43
5	Conclusões	54
	Referências Bibliográficas	57

Capítulo 1

Introdução

A partir do avanço do conhecimento médico sobre os agentes causadores das infecções, no século passado, foi possível sistematizar a dinâmica de transmissão de uma doença infecciosa em termos de modelos matemáticos. O primeiro, e provavelmente mais usado desses modelos, denomina-se de modelo SIR, nomenclatura esta que se deve à dinâmica de transmissão entre as classes (suscetível, infectado e removido) de indivíduos dentro da população. Praticamente todos os estudos realizados com esta dinâmica, desde o início do século passado, consideram as classes distribuídas homogeneamente na população, sem levar em conta a estrutura de interação entre os indivíduos. Entretanto, na última década, percebeu-se que a transmissão de uma doença dá-se segundo uma rede de contatos que depende da estrutura de interação entre os indivíduos, a qual pode depender do tipo de doença infecciosa. A estrutura das interações pode ser representada por certas redes paradigmáticas (mundo pequeno, livre de escala, etc). A introdução da heterogeneidade espacial levou a uma conclusão distinta acerca do número crítico de infectados necessários para levar a população a uma epidemia, sugerindo um deste valor ou, em algumas redes livres de escala, mesmo um valor nulo. Representar as interações intrínsecas de uma população e seus elementos e as possíveis regras de transmissão para que o espalhamento de uma epidemia ocorra de acordo com o que se observa na natureza em um modelo, passa por um estudo de situações epidêmicas hipotéticas em redes de contatos reais.

Dentro deste contexto, este trabalho tem por objetivo implementar e analisar simulações

computacionais que descrevam a dinâmica de transmissão de uma hipotética doença sexualmente transmissível dentro de topologias paradigmáticas e real. A partir destas simulações é possível avaliar o impacto do tamanho de epidemia em tais redes tanto no que se refere a efeitos da estrutura bem como a aspectos ligados a regras de transmissão. Desta maneira, implementamos as redes paradigmáticas com o mesmo tamanho da componente gigante da rede romântica. Assim, é possível fornecer subsídios para realizar um paralelo entre os comportamentos já explorados na literatura das redes paradigmáticas e esta topologia real.

Em termos de estrutura, este trabalho está dividido em cinco capítulos, dos quais está introdução é o primeiro. No capítulo dois, apresentamos os rudimentos sobre as redes complexas com o intuito de introduzir ao leitor os conceitos e resultados relevantes que serão utilizados frequentemente nos capítulos seguintes, bem como fixar a notação empregada. Iniciamos o capítulo três com uma importante contextualização dos avanços nos modelos epidêmicos. Nesse capítulo nos concentramos na discussão do modelo SIR tradicional e a descrição da propagação de doenças em redes complexas. No capítulo quatro, implementamos e analisamos resultados de simulações computacionais para a propagação de uma doença na topologia real (estática e dinâmica) e paradigmáticas apresentando as diferenças no tamanho final da epidemia, diferenças estas vinculadas às estruturas de contatos, regras de transmissão e dinâmica dos contatos. No último capítulo, capítulo 5, apresentamos as conclusões e perspectivas dos resultados obtidos no capítulo anterior.

Capítulo 2

Introdução às Redes Complexas

As redes, basicamente descritas como conjuntos de itens conectados entre si, são observadas em inúmeras situações, desde o nível subatômico até as mais complicadas estruturas sociais ou materiais concebidas pela humanidade. Há variáveis importantes inerentes às ligações dos constituintes da rede e às formas de construção de tais relações. Propriedades essenciais das redes estão na própria topologia, na descrição física ou geométrica das mesmas. Todos estes elementos se perdem quando o foco deixa de ser a rede e passa a incidir apenas nos ítems. Com efeito, examinar um único neurônio não é o suficiente para descrever o cérebro, da mesma forma que analisar um indivíduo ou um computador não permite conclusões amplas a respeito, respectivamente, das sociedades ou da Internet.

Este capítulo introduz as redes complexas, apresentando os conceitos básicos. A primeira seção destina-se a base conceitual, no qual indicamos o que são e como podem ser definidas os diferentes tipos de redes. Na segunda seção, veremos algumas propriedades estatísticas cujo o exame se faz necessário para adequada caracterização da topologia. A terceira seção por sua vez apresenta as principais estruturas (tipos de redes) e os modelos responsáveis por gera-las. Finalizamos este capítulo, quarta seção, apresentando exemplos sistemas naturais de estrutura complexa.

2.1 Conceitos Básicos

O estudo de redes complexas é um território da teoria de grafos (Albert e Barabási [1]). A teoria dos grafos, particularmente, começou com o trabalho de Leonhard Euler para resolver o problema das sete pontes de Königsberg, cidade localizada na Prússia do século XVIII (atual Kaliningrado, Rússia), onde haviam duas grandes ilhas com terra firme e conectadas entre elas por sete pontes. Discutia-se nas ruas da cidade a possibilidade de alguém atravessar todas as pontes sem repetir nenhuma. A chance de tal façanha havia se tornado uma lenda popular quando Leonhard Euler, em 1736, provou que não existia tal caminho. Ele modelou o problema como um grafo, veja figura 2.1, transformando as pontes em linhas (arestas) que ligam círculos (vértices), representando as ilhas e margens do rio, criando assim o que pode ser o primeiro grafo da história. Contudo, um

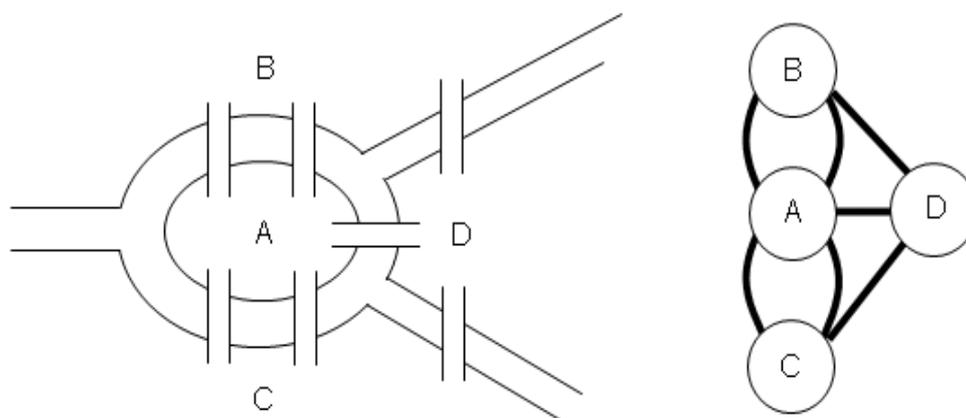


Figura 2.1: Representação do problema das sete pontes de Königsberg sugerido por Euler. Na linguagem dos grafos tal figura é constituída por quatro vértices e sete arestas.

conjunto de vértices ligados por arestas é somente um tipo simples de rede, tal como o grafo sugerido no problema das sete pontes de Königsberg. Segundo Newman [2], pode existir mais do que um tipo de vértice numa rede, ou mais do que uma aresta. Os vértices e arestas ainda podem exibir propriedades distintas entre eles. Por exemplo, numa rede formadas por pessoas, os vértices podem representar homens ou mulheres, pessoas de diferentes nacionalidades, idades, etc. Arestas por sua vez podem representar relações de amizades, relações românticas, proximidade geométrica, etc. Estas relações podem ser simétricas como no problema das pontes ou não, como numa relação hierárquica. Destacamos na figura 2.2 alguns tipos de redes.

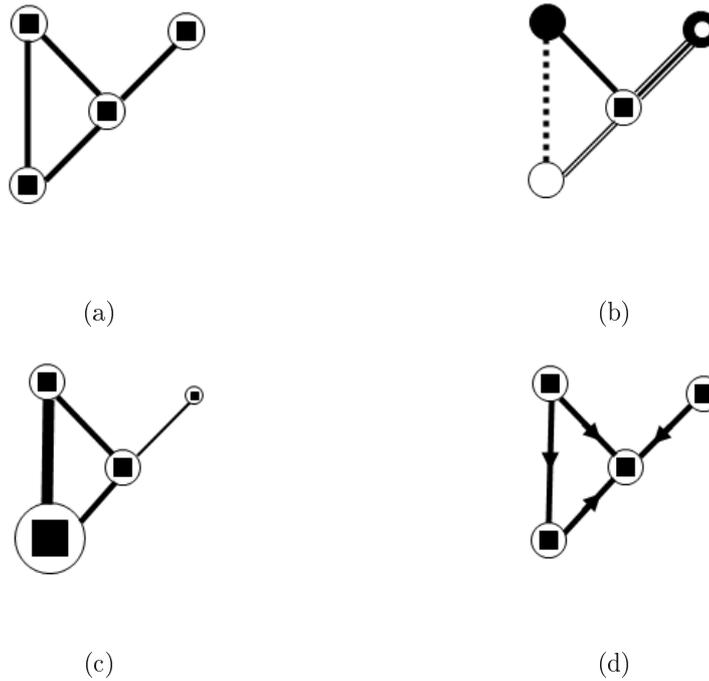


Figura 2.2: Exemplos de alguns tipos de redes: (a) rede não-dirigida com um tipo de vértice e um tipo de aresta; (b) rede com um número discreto de tipos de vértices e arestas; (c) rede com os pesos dos vértices e das arestas variando; (d) rede dirigida em que cada aresta tem uma direção.

As redes podem ser estáticas, quando não há variação no número de vértices, arestas ou mesmo na configuração das ligações; ou dinâmicas, sendo que, neste caso, é possível modelar o seu crescimento pela análise da variação de sua estrutura no tempo. Embora as redes reais sejam dinâmicas, elas podem ser analisadas como estáticas dentro de um intervalo de tempo em que as variações são inexistentes ou pouco importantes (Costa et al [3]).

Por vezes, é útil representar uma rede por uma matriz, que recebe o nome de matriz de adjacência. Para tanto, considerando num primeiro momento apenas redes não-dirigidas, devemos observar que a construção desta matriz dependerá, basicamente, se a rede é do tipo binária (rede sem pesos) ou não binária (rede pesada). Assim, se a rede é do tipo binária denotaremos a matriz de adjacência por \mathcal{A} (com elementos a_{ij}), enquanto a notação utilizada para os grafos não binários será \mathcal{W} (com elementos w_{ij}). O elemento a_{ij} é igual a unidade se existe uma conexão entre os vértices i e j , caso contrário seu valor é nulo. Já o elemento w_{ij} é igualmente nulo quando não há ligação entre os vértices i e j e igual a intensidade da conexão, peso, quando os vértices i e j estão ligados. Uma outra forma de representação das redes é dada por intermédio

de listas. De maneira análoga às matrizes de adjacência temos que fazer distinção entre as redes binárias e não binárias. As listas para redes binárias são construídas de tal forma que apenas os vértices (i, j) que possuem ligações são armazenados, enquanto que para as redes não binárias a lista tem um terceiro elemento associado a intensidade das ligações, (i, j, w_{ij}) . Em termos computacionais, cada estrutura tem sua vantagem. O uso das listas permite maior economia de memória (quando as redes são esparsas) do que o uso das matrizes de adjacência, embora o acesso às ligações seja mais complexo, porque são necessárias buscas na lista. Todas as afirmações realizadas neste parágrafo continuam válidas para redes dirigidas; contudo, devemos levar em conta agora a orientação dos vértices para a construção da matriz de adjacência ou lista, ou seja, o elemento armazenado tanto na matriz de adjacência quanto na lista será aquele que além de ligar os vértices i e j possuir a direção de ligação entre os mesmos. Neste trabalho nos restringiremos às matrizes de adjacência. Como exemplos, as redes na figura 2.3 são representadas por matrizes

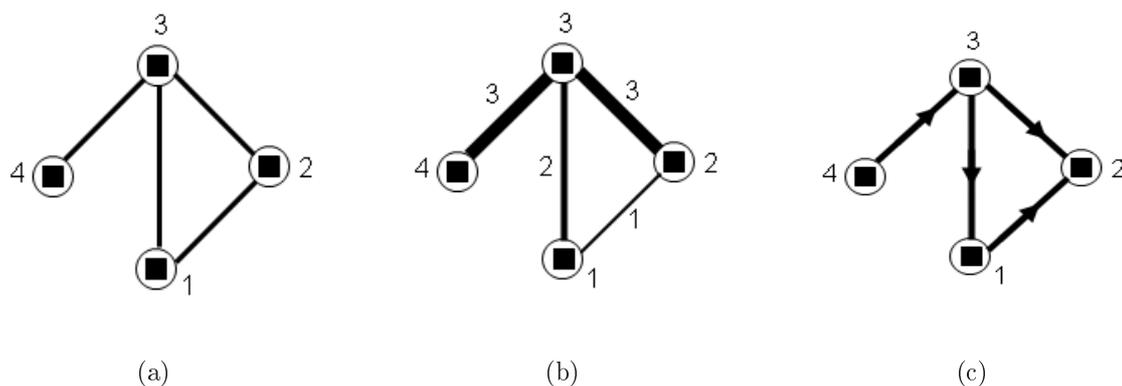


Figura 2.3: (a), Rede binária não-dirigida; (b), Rede não binária não-dirigida; (c), Rede binária dirigida

de adjacência, respectivamente, como

$$\mathcal{A}_a = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \quad \mathcal{W}_b = \begin{pmatrix} 0 & 1 & 2 & 0 \\ 1 & 0 & 3 & 0 \\ 2 & 3 & 0 & 3 \\ 0 & 0 & 3 & 0 \end{pmatrix} \quad \mathcal{A}_c = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

Notamos que as redes não-dirigidas fornecem uma matriz de adjacência simétrica, $a_{ij} = a_{ji}$ e, por outro lado, as redes dirigidas levam a matrizes de adjacência não simétricas.

2.2 Medidas

Redes complexas podem apresentar diferentes topologias, dependendo dos mecanismos que determinam sua formação e/ou evolução. Para se quantificar a estrutura das ligações, diversas medidas têm sido desenvolvidas. Apresentamos a seguir algumas medidas básicas, para um estudo mais detalhado sugerimos os trabalhos de Albert e Barabási [1], Newman [2] e Costa et al [3].

A característica fundamental de um vértice é o seu grau (conectividade). Denotada por k_i , ela indica o número de ligações entre o vértice i e os demais. Em redes não-dirigidas o grau está associado a matriz de adjacência da seguinte forma:

$$k_i = \sum_j^N a_{ij} \quad (2.1)$$

onde N refere-se ao número total de vértices, e a_{ij} é o elemento (i, j) da matriz de adjacência definida na seção anterior. A partir da medida do grau podemos determinar o grau médio, dado

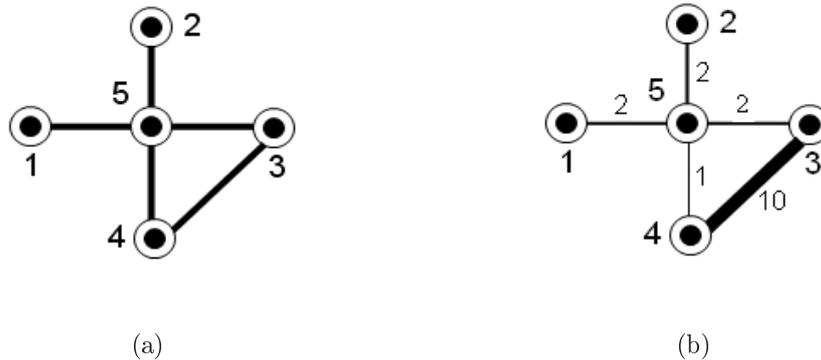


Figura 2.4: Exemplo de duas redes formadas por 5 vértices e 5 arestas. Na rede sem peso (a) identificamos o grau de cada vértice como $k_1 = k_2 = 1$, $k_3 = k_4 = 2$ e $k_5 = 4$, sendo assim o vértice 5 é o hub porque concentra grande parte das conexões da rede que possui grau médio $\langle k \rangle = 2$. A rede com peso (b) possui as forças dos vértices s_1, s_2, s_3, s_4 e s_5 , respectivamente, como 2, 2, 12, 11 e 7, com força média $\langle s \rangle = 6.8$ e, notamos ainda, que o vértice 3 é um hub por possuir ligações com maior intensidade.

por

$$\langle k \rangle = \frac{1}{N} \sum_{i=1}^N k_i. \quad (2.2)$$

Se as conexões na rede possuem pesos é utilizada a força do vértice, s_i , ao invés do grau, que é

calculada por

$$s_i = \sum_j^N w_{ij} \quad (2.3)$$

A respectiva média global é a força média da rede

$$\langle s \rangle = \frac{1}{N} \sum_{i=1}^N s_i. \quad (2.4)$$

A figura 2.4 ilustra as definições mencionadas. Estas medidas, apesar de simples, podem ser utilizadas para a identificação dos hubs (vértices com maior número de conexões) e para quantificar a densidade de conexões. Os hubs têm importância fundamental na formação da estrutura das redes complexas, já que a sua remoção pode causar a fragmentação da rede, resultando em componentes não conectados (vértices isolados, ou grupos de vértices conectados entre si, porém desconectados da maior parte da rede). Esse fenômeno está relacionado às falhas aleatórias e ataques em redes, sendo que, dependendo da aplicação, pode-se desejar redes robustas ou vulneráveis.

Para ter uma caracterização global dos grau dos vértices da rede definimos uma quantidade chamada distribuição de grau, $P(k)$; dois exemplos são dados na figura 2.5. Esta quantidade representa a probabilidade de um vértice escolhido aleatoriamente ter grau k . Da mesma forma é definida a distribuição da força, $P(s)$. Tais medidas são importantes, por exemplo, para determinar se as conexões entre os vértices são uniformes, servindo como uma medida básica de classificação.

Referindo-nos novamente aos vértices, cada um possui uma quantidade associada que mede o nível de conexão entre seus vizinhos (vértices conectados a ele). Em redes sociais, essa quantidade, chamada aglomeração (*clustering*) ou transitividade, indica a probabilidade de dois amigos quaisquer A e B terem um amigo C em comum. Para medir a fração de tais subgrafos usa-se o coeficiente de clusterização (ou coeficiente de aglomeração), que mede a razão entre o número de arestas entre os vizinhos de um dado vértice i , denotado por e_i , e o número máximo possível de arestas entre esses vizinhos, que é dado por $k_i(k_i - 1)/2$. A relação que define o coeficiente de

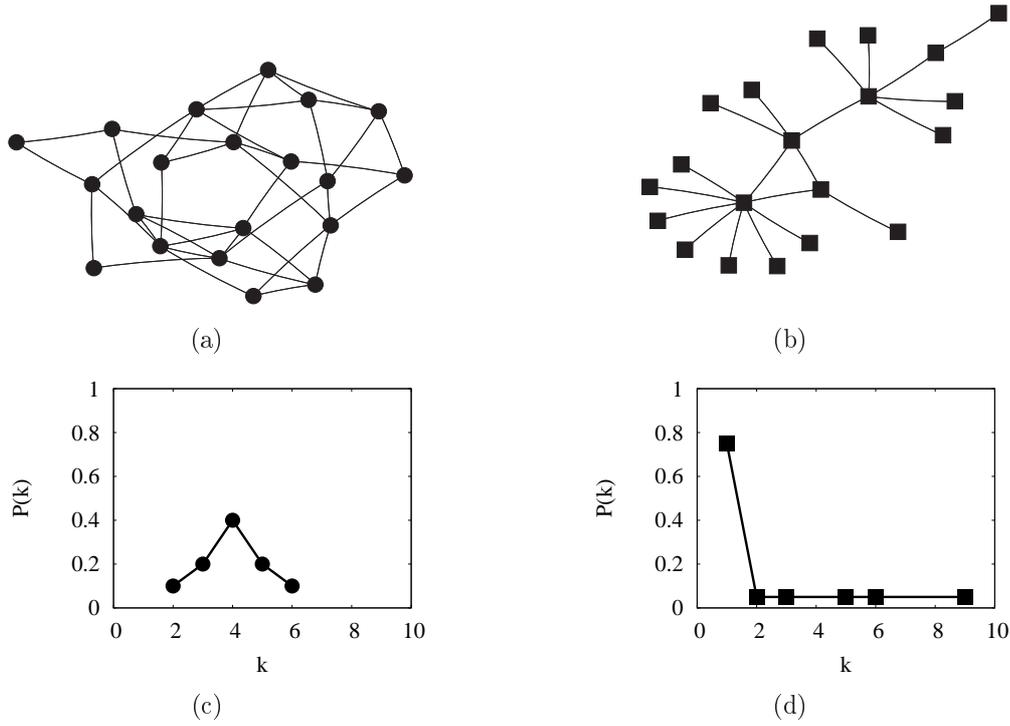


Figura 2.5: Dois exemplos de redes, (a) e (b), e suas respectivas distribuições de grau, (c) e (d).

clusterização resulta:

$$C_i \equiv \frac{2e_i}{k_i(k_i - 1)} = \frac{\sum_{j=1}^N a_{ij} \sum_{m=1}^N a_{jm} a_{mi}}{k_i(k_i - 1)} = \frac{\sum_{j=1}^N a_{ij} \mathcal{A}_{ji}^2}{k_i(k_i - 1)}, \quad (2.5)$$

sendo que o último termo permite a obtenção de C_i baseado na matriz de adjacência \mathcal{A} . Como medida de caracterização global da rede, é calculada a média do coeficiente de aglomeração entre todos os vértices da rede,

$$C = \frac{1}{N} \sum_{i=1}^N C_i. \quad (2.6)$$

Na figura 2.6 podem ser vistos alguns exemplos.

Até o momento investigamos medidas associadas às conexões dos vértices. Uma outra medida relevante está associada com a distância entre os vértices numa dada rede. O comprimento do caminho que conecta dois vértices i e j é dado pelo número de arestas ao longo deste caminho. O menor caminho entre dois vértices i e j , d_{ij} , é dado pelo menor caminho entre todos os caminhos que conectam estes vértices. Sua determinação é importante para caracterização da estrutura interna das redes e na investigação de efeitos dinâmicos relativos ao transporte e à comunicação.

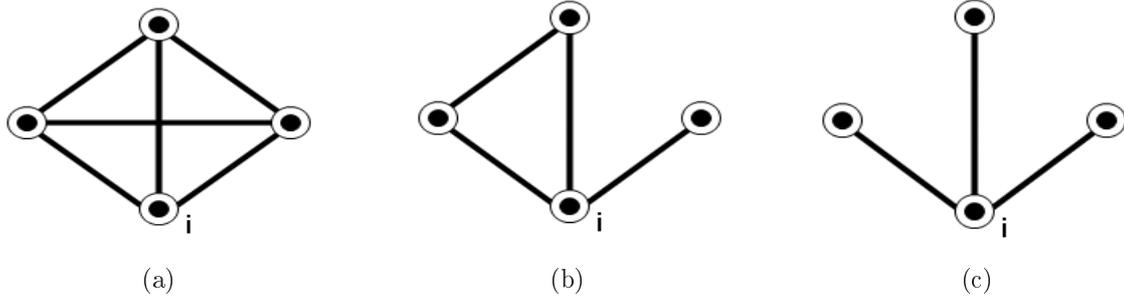


Figura 2.6: Exemplos de redes com diferentes coeficientes de clusterização. Em (a) todos os vértices estão conectados entre si, $C_i = 1$, (b) $C_i = 1/3$ e em (c) os vizinhos do vértice i não possuem conexão entre eles, $C_i = 0$.

Os menores caminhos entre todos os vértices em uma rede podem ser representados através de uma matriz de distâncias Ω , cujos elementos d_{ij} expressam o valor do menor caminho entre os vértices i e j . O valor $d_{max} = \max d_{ij}$ é chamado diâmetro da rede, D . A média entre os valores na matriz Ω exprime o caminho característico da rede (menor caminho médio), sendo calculada por

$$\ell = \frac{1}{N(N-1)} \sum_{i \neq j} d_{ij} \quad (2.7)$$

Caso i e j não pertençam a um mesmo componente conectado, $d_{ij} = \infty$, o que é um problema na definição de ℓ . Uma possibilidade para sair desta problemática é utilizar a chamada eficiência global, definida como

$$E = \frac{1}{N(N-1)} \sum_{i \neq j} \frac{1}{d_{ij}} \quad (2.8)$$

isto é, a média dos inversos dos menores caminhos. Tomamos a rede na figura 2.7, para ilustrar os aspectos abordados neste parágrafo. Observando a figura, podemos construir a matriz de

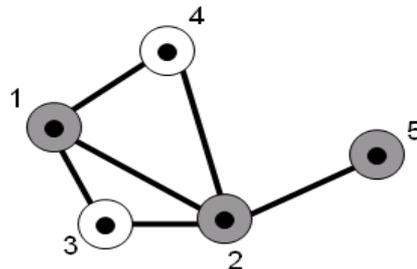


Figura 2.7: Exemplo de caminho na rede. Em cinza destacamos os vértices que pertencem ao menor caminho que separa os vértices 1 e 5.

distância

$$\Omega = \begin{pmatrix} 0 & 1 & 1 & 1 & 2 \\ 1 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 2 & 2 \\ 1 & 1 & 2 & 0 & 2 \\ 2 & 1 & 2 & 2 & 0 \end{pmatrix} \quad (2.9)$$

A partir da matriz Ω verificamos que o diâmetro é $D = 2$, e fazendo a média das distâncias, equação (2.7), obtemos $\ell = 1.4$.

2.3 Estruturas e Modelos

O trabalho de Euler no século XVIII, descrito no início deste capítulo, foi apenas um passo para o desenvolvimento da teoria de grafos. Contudo, Euler investigou apenas uma classe de pequenos grafos com alto grau de regularidade.

Uma fonte particularmente rica de idéias foi o estudo de grafos aleatórios, grafos em que as arestas são distribuídas aleatoriamente. Redes com uma topologia complexa freqüentemente parecem aleatórias, permitindo assim o emprego da teoria de grafos aleatórios no estudo delas. A teoria de grafos aleatórios foi introduzida por Erdős e Rényi [4]. Um grafo aleatório é definido como N vértices conectados por M arestas, dadas aleatoriamente dentro das $N(N-1)/2$ possíveis. No total existem $C_M^{N(N-1)/2}$ grafos com N vértices e M arestas. Uma alternativa equivalente é o modelo binomial, onde, do conjunto de N vértices inicialmente desconectados, todos os pares de vértices são percorridos, apenas uma vez, e conectados com probabilidade p . Assim sendo, todas as ligações possuem a mesma probabilidade de ocorrerem, ou seja, a rede gerada tem uma estrutura altamente homogênea. Conseqüentemente, o número de arestas é uma variável aleatória com valor esperado $E(M) = p[N(N-1)/2]$. A distribuição da conectividade para essas redes, quando N é grande e a conectividade média é mantida constante, tende à distribuição de Poisson. Além disso, o caminho mínimo médio é pequeno nessas redes, caindo com o logaritmo do tamanho da rede, $\ell \sim \log N / \log \langle k \rangle$, sendo $\langle k \rangle = 2M/N = p(N-1)$ o número médio de conexões na rede. Erdős e Rényi estavam interessados apenas na riqueza matemática das redes aleatórias e não em

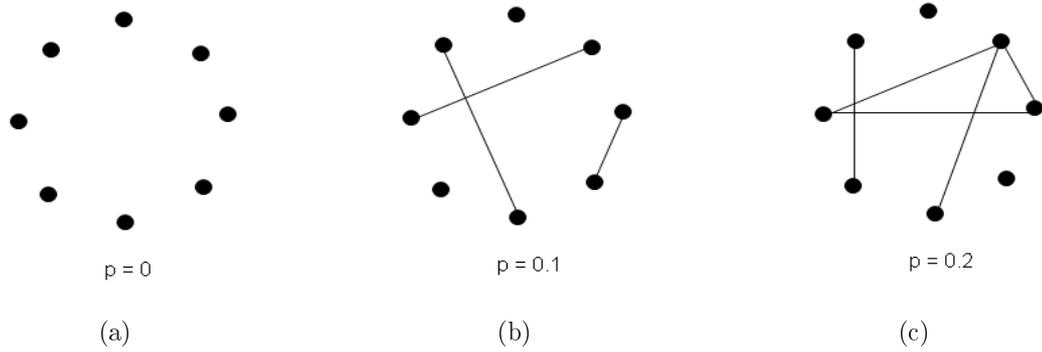


Figura 2.8: Ilustração do processo de evolução do grafo para o modelo de Erdős-Rényi. Iniciamos com $N = 8$ vértices isolados. (a),(b) e (c): Conectamos cada par de vértices com probabilidade p . Cada grafo corresponde a uma única realização.

aplicações práticas, tendo apenas mencionado em seu artigo de 1959 que a evolução dos grafos poderia ser considerada como um modelo muito simplificado de certas redes de comunicação. Mesmo assim, o modelo de Erdős e Rényi suscitou questões relacionadas à estrutura das conexões em redes reais, como por exemplo, se estas conexões poderiam ser representadas pelo modelo aleatório. Entretanto, essa resposta só foi obtida no final da década de 90, quando novas bases de dados surgiram e o poder computacional aumentou.

O primeiro passo nessa descoberta foi dado em 1998, quando Watts e Strogatz observaram que em algumas redes reais, tais como a rede de neurônios do *Caenorhabditis elegans* e a rede de distribuição de energia dos Estados Unidos, a presença de *loops* (caminhos fechados) é muito maior do que nas redes aleatórias com mesmo número de vértices e arestas (Watts e Strogatz [5]). Esse foi o primeiro indício de que as redes reais não são completamente aleatórias, mas possuem uma determinada lei de formação. Baseados nesta descoberta, Watts e Strogatz sugeriram um modelo alternativo aos grafos aleatórios, chamado modelo *small world* de Watts-Strogatz (em analogia ao fenômeno descoberto por Stanley Milgram [6]), que apresenta o efeito *small world*. Neste modelo, eles assumem que as redes presentes na natureza não são completamente regulares e nem mesmo aleatórias, todavia se situam entre esses dois extremos. O algoritmo deste consiste: (i) Ordem: Inicia-se com um anel regular com N vértices em que cada vértice é conectado a κ primeiros vizinhos ($\kappa/2$ de cada lado), considere $N \gg \kappa \gg \log N \gg 1$. (ii) Randomizar: Aleatoriamente reescreva cada aresta da rede regular com probabilidade p de modo que auto-

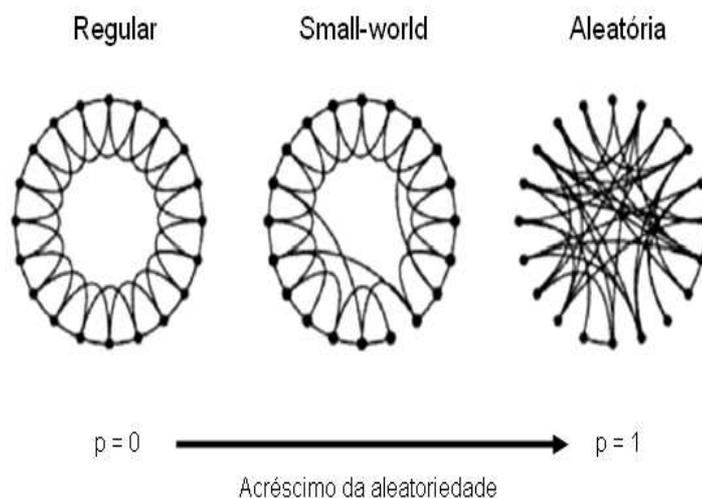


Figura 2.9: O processo aleatório de re-conexão das arestas no modelo de Watts e Strogatz, que interpola entre um anel regular e uma rede aleatória sem alterar o número de vértices e arestas. Iniciamos com $N = 20$ vértices cada um conectado aos seus primeiros quatro vizinhos. Para $p = 0$ o anel original não é alterado, com o p aumentando a rede torna-se crescentemente desordenada até que para $p = 1$ todas as arestas são reescritas aleatoriamente (Watts e Strogatz [5]).

conexões e conexões duplas são excluídas. Este processo introduz $pN\kappa/2$ arestas cujas conexões não estão restritas a seus primeiros vizinhos (arestas de longo alcance). Conseqüentemente, variando p transitamos entre a regularidade e a aleatoriedade, este processo é ilustrado na figura 2.9. Na figura 2.10 mostramos o comportamento do coeficiente de clusterização médio, C , e a distância média, L , como função da probabilidade de redirecionamento das arestas, p . Destacamos três regiões com características topológicas diferentes. A primeira corresponde aos baixos valores p em que C é praticamente invariante e L com um grande valor, quando comparado as outras regiões. Esta região corresponde a rede regular. Para grande valores de p , C indo a zero e L pequeno, logo associamos esta região a completa desorganização das redes aleatórias. Entre as duas regiões, identificamos as redes de *small-world*. É importante destacar que o limite entre estas regiões não é bem definido na literatura. O trabalho de Watts e Strogatz formalizou um antigo problema de sociologia proposto por Granovetter [7], que estudou as relações entre pessoas na sociedade e mostrou que as ligações mais fracas de amizade são extremamente importantes. Granovetter demonstrou que os laços afetivos entre familiares e entre amigos íntimos não oferecem uma diversidade de conhecimento tão grande como as relações entre pessoas conhecidas e amigos distantes. No modelo de Watts e Strogatz, as ligações mais fracas sugeridas por Granovetter

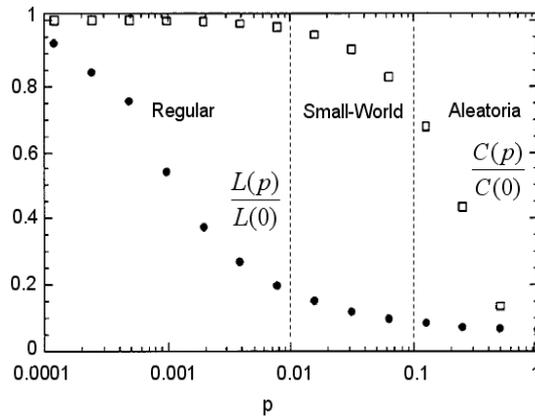


Figura 2.10: Coeficiente de clusterização médio, $C(p)$, e distância média $L(p)$ para o modelo de Watts-Strogatz. Os dados são normalizados por seus respectivos valores na rede regular, $C(p = 0)$ e $L(p = 0)$. Destacamos as regiões características das topologias regular, *small-world* e aleatória (Watts e Strogatz [5]).

são aquelas estabelecidas pela reconexão de arestas. Estas ligações têm um papel fundamental na conexão entre os membros de grupos fechados (formado por familiares e amigos íntimos) e o mundo externo. É através delas que as redes sociais se tornam *small world*. O trabalho de Watts e Strogatz trouxe o fenômeno *small world* da sociologia para as comunidades formadas por físicos e matemáticos, inspirando novas investigações em redes complexas.

Em 1999 Barabási e Albert [8] decidiram verificar se o fenômeno *small world* estava presente na *World Wide Web* (WWW). Para isso, mapearam a topologia das conexões entre as páginas da WWW e descobriram que além dela apresentar o fenômeno *small world* ($\ell \sim 11$), a distribuição de conexões não é aleatória, mas do tipo livre de escala, que é da forma $P(k) \sim k^{-\gamma}$. Inspirados por estas descobertas, eles propuseram um modelo de crescimento que gera redes livres de escala em base a dois processos: (i) Crescimento: Iniciando-se com um pequeno número de vértices N_0 , a cada passo é adicionado um vértice com m ($m \leq N_0$) arestas que se conectam com vértices já presentes na rede. (ii) Ligação preferencial: O novo vértice, que vai ser adicionado à rede, tende a se conectar com os vértices mais conectados, ou seja, a probabilidade P de um novo nó se conectar ao vértice i depende do grau k_i do vértice i , de modo que

$$P(k_i) = \frac{k_i}{\sum_j k_j} \quad (2.10)$$

Depois de t passos de tempo este processo resulta numa rede com $N = t + N_0$ vértices e mt arestas.

Interessante notar que estes dois mecanismos de construção das redes livres de escala não estão

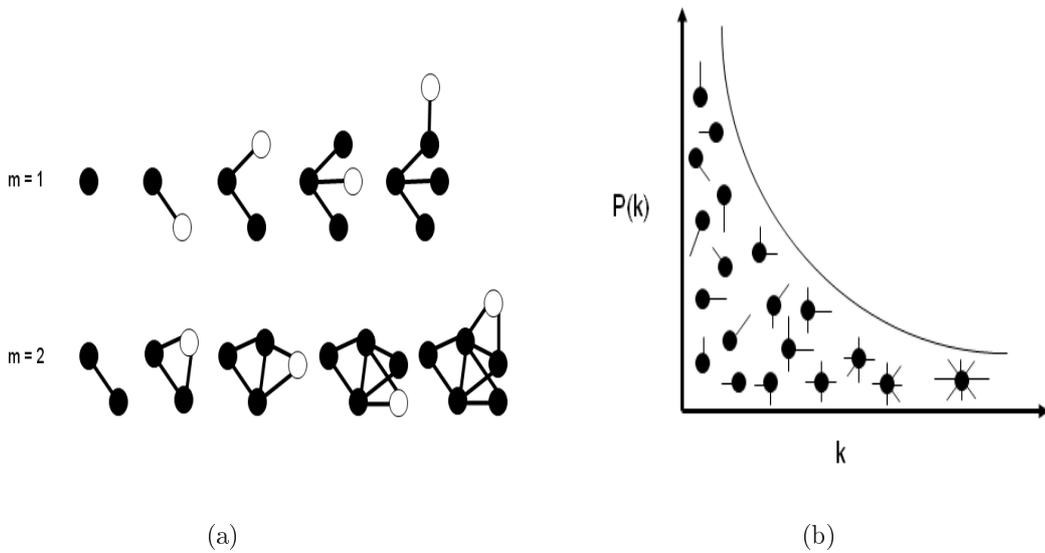


Figura 2.11: (a) Dinâmica de anexação preferencial do modelo de Albert-Barabasi [8] com parâmetros $m = 1$ e $m = 2$. O vértice branco representa o novo vértice adicionado em cada passo de tempo. (b) Representação esquemática para a distribuição de grau de uma hipotética lei de potência. Podemos observar que muitos vértices possuem poucos *links* enquanto que poucos vértices possuem muitos *links*.

presentes no modelo aleatório de Erdős e Rényi e no modelo *small world* de Watts e Strogatz, pois nestes não há adição de novos vértices e as conexões são estabelecidas de forma homogênea, havendo uma conectividade característica (conectividade média) na rede. Por outro lado, nas redes geradas pelo modelo livre de escala, os vértices mais conectados tendem a receber mais conexões, paradigma conhecido como “*O rico fica mais rico*”. Estas redes são formadas por um reduzido número de vértices altamente conectados, denominados hubs, e por uma grande quantidade de vértices pouco conectados, o que define a distribuição livre de escala ¹ (figura 2.11(b)).

Algumas medidas básicas das redes geradas pelos modelos citados nesta seção possuem expressões analíticas (Newman [2]). Além dos modelos discutidos anteriormente, muitos outros foram sugeridos de modo a permitirem uma modelagem mais precisa de diversas redes reais. Para maiores detalhes sobre tais modelos, recomendamos a revisão apresentada por Boccaletti et al. [9].

¹A distribuição de grau resultante do mecanismo de formação das redes de Albert-Barabasi é $P(k) = 2m^2/k^3$.

2.4 Redes no Mundo Real

Sistemas compostos por muitas unidades semelhantes que interagem entre si são muito comuns, e, em geral, a rede de interações não é completamente aleatória e sequer completamente regular, como já observamos na seção anterior. Por isso, encontramos redes complexas nos mais diversos contextos. Nesta seção apresentaremos algumas estruturas de redes, de diferentes tipos, presentes na natureza ou construída pelo homem, como as redes sociais, biológicas, de informação e tecnológicas.

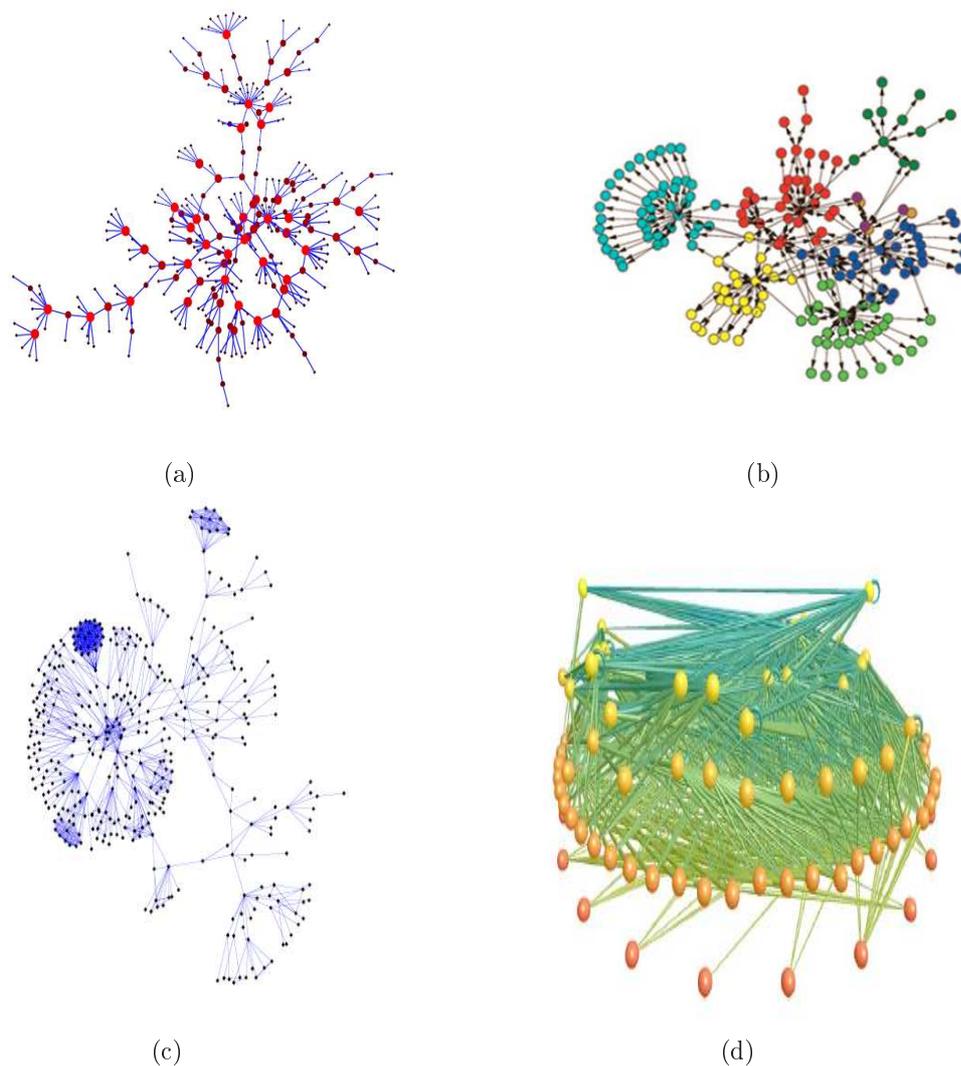


Figura 2.12: Exemplos de redes no mundo real: (a) Rede social de contatos sexuais (Potterat et al [10]). (b) Rede de hipertextos num sítio da *Web*, exemplo de rede de informação (Newman e Girvan [11]). (c) Redes de roteadores da internet, exemplo de rede tecnológica. (d) Rede da cadeia alimentar predador-presa de espécies de lago de água doce (Martinez [12]).

Uma rede social não é mais que um conjunto de pessoas ou grupos de pessoas, ligados entre

si por relações que podem ser profissionais, familiares ou outras. A introdução de modelos matemáticos foi feita por Rapoport [13], o primeiro autor a realçar a importância da distribuição do número de ligações. Outro estudo importante nesta área foi o de Milgram [6], através das suas experiências que levaram ao conceito de *small world*. Estas experiências estão na base do conceito dos seis graus de separação. A reconstrução de redes sociais é difícil, dada a subjetividade envolvida e a limitação dos métodos de entrevistas e aplicação de questionários. Em alguns casos, há registros que auxiliam nesta tarefa, como, por exemplo, documentos de composição de diretorias e associações, contratos trabalhistas, co-autoria em publicações, registros de comunicações de correios, telefonia, mensagens eletrônicas, etc.

Redes de informação ou de conhecimento são obtidas a partir de bases de conhecimento formal ou banco de dados, como as citações de artigos científicos, a World Wide Web, os registros de patentes, a estrutura das linguagens, etc. O exemplo clássico de uma rede de informação é a rede das citações entre artigos científicos. A maioria dos artigos publicados cita outros trabalhos anteriores relacionados ou importantes para este. Estas citações geram uma rede dirigida em que os vértices são os artigos científicos e as arestas as citações para artigos anteriores. A estrutura da rede de citações reflete assim a estrutura da informação armazenada nos seus vértices: daí a designação rede de informação. Estas redes têm certas particularidades, que por vezes não surgem noutras. Isso prende-se com o fato de, após a publicação de um artigo que cita outros anteriores, as citações não serem mais alteradas. Outros exemplos são as redes bipartidas de preferências, onde os dois tipos de nós representam os indivíduos e os objetos de sua preferência (programas de TV, bens de consumo, etc.). As arestas podem ser pesadas, indicando maior ou menor afinidade com o objeto.

Redes tecnológicas são feitas pelo ser humano, projetadas tipicamente para distribuição de facilidades ou recursos, tais como eletricidade, água, etc. Figuram nesta categoria as rotas dos meios de transporte, de correios e serviços de distribuição, circuitos eletrônicos e a própria Internet.

As redes complexas são identificadas em muitos sistemas biológicos. Há ocorrências muito claras, como no sistema nervoso ou vascular, cada qual consistindo naturalmente numa rede. Por outro lado, uma análise mais profunda permite a percepção das redes no metabolismo dos seres

vivos: substratos e produtos são vértices de uma rede cujas arestas representam a existência de uma reação metabólica que, a partir do substrato, gere o correspondente produto. A mecânica de interação física das proteínas também corresponde a uma rede complexa de interação. Redes reguladoras são importantes na biologia: a expressão de um gene, vista como a produção da proteína que este codifica é controlada pela presença ou ausência de outras proteínas, ativadoras ou inibidoras. O genoma em si forma uma rede com os vértices representando as proteínas e arestas representando as dependências da produção de proteínas. Cadeias alimentares são um bom exemplo de redes complexas, de significativo interesse nos estudos de ecossistemas.

Capítulo 3

Elementos de Epidemiologia Matemática

As epidemias não são oriundas do contexto contemporâneo. Dentro da história é possível contemplar inúmeros eventos epidêmicos responsáveis por ter dizimado grande parte da população, tal como a peste bubônica no século XIV. As descobertas de vacinas e a implementação de planos de vacinação em massa juntamente com outros avanços tecnológicos permitiram o controle de uma parcela das doenças contagiosas. Contudo, o risco de propagação epidêmica sempre esteve latente devido à pluralidade das doenças. Tal risco motivou a elaboração de modelos para descrever os fatores responsáveis na propagação de uma doença.

Destinamos este capítulo a um estudo introdutório destes modelos enquadrados dentro da epidemiologia matemática. Na primeira seção exploramos breves aspectos históricos que fundamentaram e permitiram avanços no referido campo. Seguimos para a segunda seção, na qual apresentamos em detalhe o modelo SIR tradicional. Na terceira seção, apresentamos as principais contribuições dos estudos teóricos de epidemias em redes complexas.

3.1 Breve Digressão Histórica

A aplicação da matemática ao estudo de doenças infecciosas teve início em 1760 por Daniel Bernoulli, ao usar um método matemático para avaliar a efetividade de técnicas de controle da varíola, com o objetivo de influenciar as políticas de saúde pública (Bernoulli [14]). O desenvolvimento de modelos matemáticos aplicados a fenômenos epidemiológicos mais amplos ficou limitado pelas

restrições intrínsecas à ausência de conhecimento médico sobre os agentes causadores das infecções. Somente a partir do século XIX, com o nascimento da bacteriologia, com Louis Pasteur e Robert Koch, e da descoberta dos vírus, foi possível identificar as causas das doenças infecciosas e, conseqüentemente, aplicar à epidemiologia modelos matemáticos bem fundamentados.

Baseado nas descobertas citadas, no início do século XX, Hamer postulou que o curso de uma epidemia depende da taxa de contato entre indivíduos suscetíveis e infecciosos (Hamer [15]). Esse conceito é um dos mais importantes em epidemiologia matemática, sendo chamado princípio de ação das massas, no qual a taxa líquida de espalhamento da infecção é assumida como sendo proporcional ao produto entre a densidade de pessoas suscetíveis e a densidade de pessoas infecciosas. O princípio foi originalmente formulado em um modelo em tempo discreto, mas em 1908 Ronald Ross transferiu o problema para um modelo em tempo contínuo, em seu trabalho pioneiro sobre a dinâmica da malária (Ross [16]). As idéias de Hamer e Ross foram estendidas e exploradas em mais detalhes por Soper, que deduziu um mecanismo fundamental responsável pela periodicidade de epidemias, freqüentemente observada (Soper [17]), e por Kermack e McKendrick, que estabeleceram a famosa Teoria do Limiar (Kermack e McKendrick [18]). Essa teoria, que afirma que a introdução de poucos indivíduos infecciosos em uma comunidade de suscetíveis não gera uma epidemia, a menos que a densidade ou número de suscetíveis esteja acima de um certo valor crítico, é, juntamente com o princípio de ação das massas, um alicerce da epidemiologia teórica moderna. Estes trabalhos, desenvolvidos nos primeiros trinta anos do século XX, resultam no que hoje conhecemos como modelo SIR determinístico, o qual discutiremos com mais detalhes na próxima seção.

A partir destes estudos, e com o grande avanço obtido pelo conhecimento biológico durante as décadas subseqüentes, a epidemiologia matemática desenvolveu-se e cresceu rapidamente, generalizando os modelos determinísticos iniciais e propondo novos modelos estocásticos que, com o aumento da capacidade de computação, ganham cada vez mais generalidade. Nos últimos dez anos, fundamentado nos trabalhos de Watts e Strogatz [5] e Albert e Barabási [8], percebeu-se que a transmissão de uma doença dá-se segundo uma rede de contatos que depende da estrutura de interação entre os indivíduos (Pastor-Satorras e Vespignani [19], Kuperman e Abramson [20],

Newman [21], Jones e Handcock [22] e Shirley e Rushton [23]). Com efeito, o modelo SIR, que toma as classes suscetível (S), infectado (I) e removível (R) homogeneamente distribuídas na população, não leva em conta aspectos dependentes da topologia. Dentro deste contexto, as redes complexas são o suporte natural para o estudo da propagação de doenças infecciosas. O conceito de redes complexas levantou uma nova questão referente à dinâmica de epidemia: como representar as interações intrínsecas de uma população e seus elementos em um modelo, e quais seriam as possíveis regras de transmissão a serem adotadas, para que o espalhamento de uma epidemia ocorra efetivamente de acordo com o que se observa na natureza? Particularmente, visto que a heterogeneidade espacial é um aspecto de grande importância, estudos de simulações em redes realistas, com o intuito de estabelecer um paralelo com os resultados em topologias conhecidas, (redes paradigmáticas) representam um campo aberto de pesquisas, explorado em parte por Bansal et al. [24] e Ward [25].

3.2 Modelo SIR Padrão

Vamos discutir nesta seção um exemplo da formulação compartimental dinâmica. Nesse tipo de abordagem a estratégia é classificar a população em categorias (ou compartimentos) disjuntas de indivíduos.

Dentro dessa abordagem determinística o modelo SIR é um dos mais comuns na literatura, já que com ele é possível introduzir os principais conceitos do processo epidêmico de maneira simples, servindo de base, como veremos, para a formulação de procedimentos nos estudos mais gerais em modelagem de epidemias. Segundo esse modelo, a população pode ser dividida da seguinte forma: S , a categoria dos suscetíveis, ou seja, dos indivíduos ou agentes que não estão infectados, mas podem ser infectados; I , a categoria dos infectados, ou seja, daqueles indivíduos que são capazes de transmitir a doença aos suscetíveis; R , a categoria dos recuperados (ou removidos/imunes), ou seja, daqueles que já foram infectados pela doença e morreram (ou se isolaram), ou ficaram permanentemente imunes à doença. Assim, em um instante de tempo t , uma população é caracterizada pelas quantidades $S(t)$, $I(t)$ e $R(t)$ que representam os números de indivíduos em cada

categoria, onde

$$N(t) = S(t) + I(t) + R(t) \quad (3.1)$$

é o número total de indivíduos da população a cada instante (tomado como constante no que segue). Usando essa lógica compartimental, se quisermos um maior realismo na descrição do processo epidêmico, devemos aumentar o número de compartimentos (ou estados) em que os indivíduos dessa população estão classificados. É importante chamar a atenção aqui que essa classificação não significa em uma segregação da população. Nessa classe de modelos não há dimensão espacial, e qualquer indivíduo pode estar em qualquer lugar a cada instante; somente os atributos são diferentes para cada indivíduo. Isto quer dizer que este tipo de modelo assume que os indivíduos de qualquer classe estão homogeneamente misturados (aleatoriamente misturados), de maneira que qualquer indivíduo de uma classe pode estar em contato com qualquer outro indivíduo de outra classe em qualquer instante, a figura 3.1 representa tal situação.

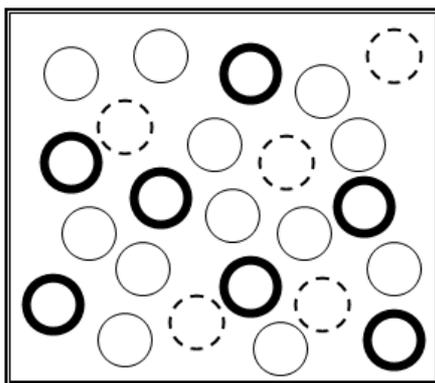


Figura 3.1: Esquema de uma população isolada, classificada de acordo com o modelo SIR; note que as classes estão misturadas aleatoriamente.

Com isso em mente passamos a definir a dinâmica do modelo (representada esquematicamente na figura 3.2), fazendo as três hipóteses básicas que definem as transições entre os estados:

1. O número de agentes na classe de infectados aumenta a uma taxa proporcional ao número de agentes na classe infectante e ao número de agentes na classe suscetível. Matematicamente, isso é representado pelo termo $\beta S(t)I(t)$, onde β é uma constante chamada de taxa de infecção. Obviamente, essa taxa representa também a velocidade com a qual os suscetíveis saem do seu compartimento. Esta hipótese é herdada do estudo de cinética química, ao



Figura 3.2: Representação da dinâmica das classes do modelo SIR.

considerar que uma reação química se dá quando dois componentes se chocam de maneira efetiva, conhecida como lei de ação das massas.

2. A taxa com a qual os indivíduos infectados são transferidos para a classe dos removidos (ou recuperados) é proporcional ao número de infectados. Esse fato é modelado pelo termo $\gamma I(t)$, onde $\gamma = 1/\tau$ também é uma constante, a taxa de recuperação, modelada como o inverso do tempo médio em que um indivíduo infectado permanece nesse estado até se recuperar.
3. Nesse modelo não é considerado o (possível) tempo de incubação; assim os indivíduos suscetíveis passam diretamente para o estado infectado.

Esses mecanismos são descritos formalmente pelas seguintes equações diferenciais:

$$\frac{dS}{dt} = -\beta SI \quad (3.2)$$

$$\frac{dI}{dt} = \beta SI - \gamma I \quad (3.3)$$

$$\frac{dR}{dt} = \gamma I \quad (3.4)$$

Na figura 3.3 resolvemos numericamente este sistema de equações diferenciais para uma situação hipotética. Observe que as curvas para a dinâmica dos estados estão normalizadas por N . Com efeito, $s = S/N$, $i = I/N$ e $r = R/N$. As figuras 3.4(a) e 3.4(b) indicam, respectivamente, os efeitos da variação dos parâmetros β e γ sobre o número de infectados. Destas curvas podemos verificar, por exemplo, que a medida que diminuimos o tempo que um indivíduo permanece infectado (aumentado γ), diminuimos o dano na população.

Podemos determinar quais as condições para uma epidemia se estabelecer, isto é, $dI/dt > 0$.

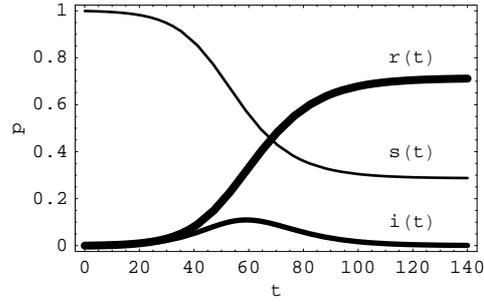


Figura 3.3: Dinâmica dos estados do modelo SIR para $\beta = 1/4$ e $\gamma = 1/7$ dadas as condições iniciais $s(0) = 0.999$, $i(0) = 0.001$ e $r(0) = 0$ em que t é o tempo e p refere-se a fração da população.

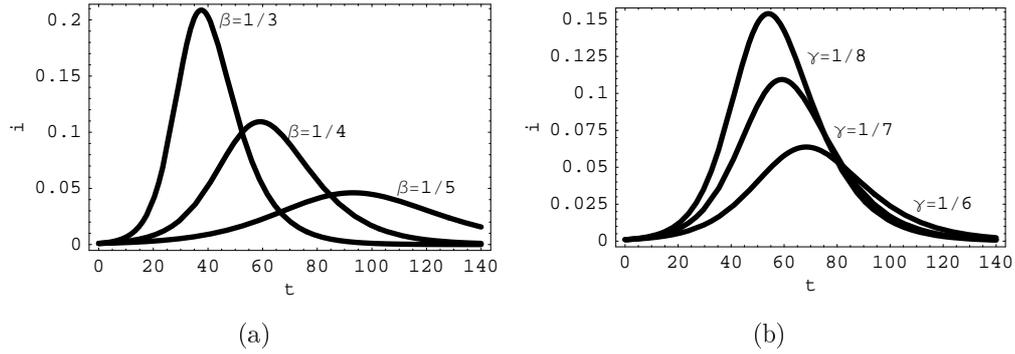


Figura 3.4: Dinâmica da fração de infectados para diferentes valores de β e γ : (a), $\gamma = 1/7$; (b), $\beta = 1/4$. As condições iniciais são $s(0) = 0.999$, $i(0) = 0.001$ e $r(0) = 0$.

Da equação 3.3, temos que a condição necessária é

$$(\beta S - \gamma)I > 0. \tag{3.5}$$

Contudo, temos sempre $I > 0$. Logo, essa relação é satisfeita quando

$$S > \frac{\gamma}{\beta}. \tag{3.6}$$

Então, para dados β e γ , existe um número crítico de suscetíveis para uma epidemia, e se $S(0)$ é o número inicial de suscetíveis, a condição para uma epidemia resulta ser $S(0) > \gamma/\beta$. A interpretação da inequação 3.6 fundamenta-se em dois aspectos: (i) um dado indivíduo infectado, em média, será contagioso por um tempo $1/\gamma$ e (ii) o número de indivíduos suscetíveis infectados por um infectado por unidade de tempo é βS . Em consequência o número total de infectados produzidos por um infectado é $\beta S/\gamma$. Assim, para que um evento seja epidêmico, este número

deve ser superior a 1. Introduzindo o número de reprodução básica, parâmetro adimensional

$$R_0 \equiv \frac{\beta S(0)}{\gamma}, \quad (3.7)$$

a condição necessária para a persistência da epidemia se expressa como $R_0 > 1$; o limiar de epidemia como $R_0 = 1$.

Um dos resultados interessantes do modelo SIR refere-se à descrição do tamanho final da epidemia. Tomando a razão entre as equações 3.2 e 3.4, e utilizando a definição 3.7, obtemos

$$\frac{dS}{dR} = -\frac{R_0}{S(0)} S \quad (3.8)$$

Assumindo que o tamanho inicial da infecção é pequeno, $S(0) \approx N$, e para tempos t' suficientemente grandes para que $I = 0$ (estado assintótico onde $R(t') + S(t') = N$), a solução de 3.8 é dada por

$$1 - r' = e^{-R_0 r'} \quad (3.9)$$

onde $r' = R(t')/N$ (fração final dos removidos pela epidemia). Trata-se de uma equação implícita, cuja curva, mostrada na figura 3.5, ilustra o efeito de se aumentar R_0 , além do limiar de epidemia $R_0 = 1$.

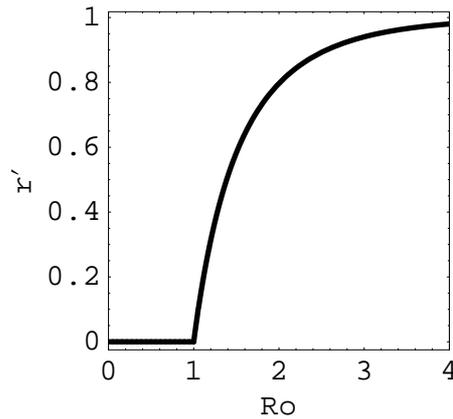


Figura 3.5: Fração da população infectada, r' , durante a epidemia em função do número de reprodução básica, R_0 .

No entanto, apesar das importantes contribuições desse modelo compartimental para o de-

envolvimento de teorias epidemiológicas, ele apresenta limitações, pois, por mais que nele sejam agregadas informações, sempre serão regidos por alguns pressupostos básicos que limitam sua abrangência. Com efeito, os modelos baseados em equações diferenciais assumem que a população é totalmente misturada, desprezando fatores espaciais ou de estrutura da população. Além disso, esse tipo de modelagem não leva em consideração o fato de que a população é composta de indivíduos interagentes; ela considera as populações como entidades contínuas, o que produz resultados não realísticos para pequenas populações. Em adição, nos modelos compartimentais todos os indivíduos dentro de um compartimento são tidos como semelhantes entre si. Sabe-se, entretanto, que há muitos fatores biológicos (idade, predisposição genética, dentre outros) e sociais (hábitos, costumes e círculo de amizades) que acabam influenciando para que um indivíduo adoça. Tais fatores podem ser, de certa forma, ainda inseridos nesse tipo de modelagem a partir da criação de sub-compartimentos que serviriam para classificá-los, contudo isso se demonstra inviável quando os fatores são muitos.

Portanto, de maneira geral, para uma melhor predição e para definir estratégias mais ou menos realistas no combate ao espalhamento de uma epidemia, é necessário desenvolver novas metodologias matemáticas e computacionais.

3.3 Epidemias em Redes Complexas

Modelos que incorporam a estrutura de redes evitam a suposição de classes de indivíduos homogeneamente misturados, designando a cada indivíduo um conjunto finito e permanente de contatos, para os quais eles podem transmitir infecções e dos quais eles podem ser infectados. Ainda que, na rede os indivíduos possam ter o mesmo número de contatos por unidade de tempo que teriam no sistema tradicional ¹, na primeira o conjunto de contatos é fixo, enquanto que no modelo SIR padrão está mudando continuamente.

Ainda que a literatura em epidemiologia matemática praticamente só se refere a versões do modelo SIR padrão, alguns progressos têm sido feitos na análise do comportamento do modelo

¹Modelo de classes SIR, homogeneamente misturado, descrito pelas equações diferenciais apresentadas na seção anterior, e que chamaremos de “SIR padrão”.

SIR na rede. Muitos destes trabalhos tem se focado em uma classe generalizada de modelos de redes aleatórias, cujas formas da distribuições de graus podem ser variadas arbitrariamente.

A progressão de uma epidemia é modelada inicialmente colocando um ou mais dos nós, escolhidos aleatoriamente, no estado I , e o resto no estado S . Em cada passo de tempo, cada vizinho de um nó infectado torna-se igualmente infectado com probabilidade $0 < \beta/N < 1$, onde β é a taxa de infecção da equação 3.2 do modelo SIR padrão e N o número de nós ou indivíduos ². Cada nó infectado, por sua vez, torna-se removido após um tempo τ . Modelos mais complexos podem ser feitos tomando β_{ij} e τ_i distribuídos, onde i e j indicam os nós, podendo assim representar uma população estratificada.

Pastor-Satorras e Vespignani [19] investigaram o crescimento de uma doença numa rede com distribuição de grau em lei de potência. O principal resultado obtido por eles, em redes grandes com distribuição de grau $P(k) \sim k^{-\alpha}$ com $\alpha < 3$, é que estas não exibem limiar epidêmico (figura 3.6). A explicação deste fato nas redes livre de escala deve-se a presença de *hubs* altamente

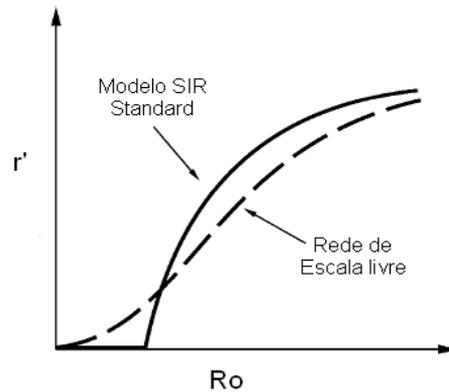


Figura 3.6: Comparação do crescimento de uma doença no modelo SIR, *standard* (padrão) e numa rede livre de escala. Em contraste com o caso padrão, com limiar de epidemia em $R_0 = 1$, na rede livre de escala não há limiar epidêmico. No gráfico, r' é a fração assintótica de removidos, enquanto que R_0 é o número de reprodução básica.

conectados, assegurando que a doença crescerá rapidamente se um desses vértice for infectado.

Outra área de investigação ativa relaciona-se com a assortatividade em redes, onde uma rede assortativa implica que nós com alto grau tendem a se conectar a nós que também possuem alto grau, em tanto que em redes dissassortativas, nós com alto grau tendem a se conectar a nós com baixo grau (Newman [2]). Analisando uma variedade de redes no mundo real, Newman e Park [26],

² β/N resulta ser a taxa de infecção por indivíduo, que não depende do número deles, N , só das características biológicas/sociais da doença. Por simplicidade de notação no que segue, no contexto de redes, chamaremos de β .

concluíram que redes sociais exibem grau de correlação positiva, são assortativas, enquanto que redes tecnológicas e biológicas (rede elétrica, Internet, cadeia alimentar, etc) são dissassortativas. Aplicando estes resultados para o crescimento de epidemias, Newman [27] mostrou que, em redes assortativas, a componente gigante conectada tende a ser menor quando comparada a outros tipos de redes. Segundo ele, redes sociais podem ser robustas para intervenções e ataques.

Alguns esforços têm sido feitos para estudar os efeitos da estrutura da rede local na propagação de doenças. Num modelo SIR adaptado a redes de mundo pequeno, Watts e Strogatz [5] mostraram, mediante simulações computacionais, que uma pequena probabilidade da reescrita das ligações entre vértices de uma rede regular reduz dramaticamente o valor da infecciosidade da doença necessário para infectar uma fração finita da população. Além disso, mostraram que, para uma doença altamente infecciosa, o tempo necessário para contaminar a população inteira é aproximadamente proporcional ao menor caminho médio da rede. Conseqüentemente, é altamente dependente da fração de atalhos aleatórios (figura 3.7). Referente ainda as redes de mundo

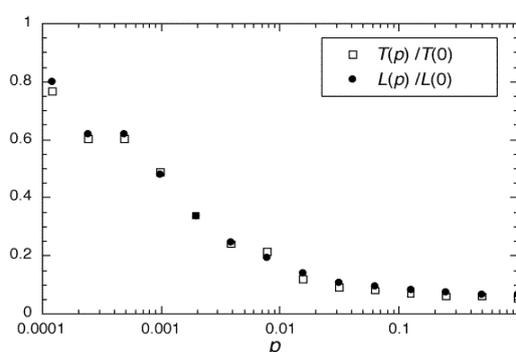


Figura 3.7: $T(p)$, tempo necessário para contaminação de toda a população, com uma taxa de infecção máxima ($\beta = 1$), tem a mesma forma funcional da distância média, $L(p)$ (Watts e Strogatz [5]).

pequeno, Kuperman e Abramson [20] encontraram que, quando os indivíduos recuperados no modelo SIR tornam-se instantaneamente em suscetíveis (modelo SIRS), uma baixa probabilidade de reescrita das ligações diminuem os níveis da doença (estado endêmico); por outro lado, alto grau de desordem, gera oscilações sustentadas da população infectada (figura 3.8).

Uma outra forma de investigar o crescimento de epidemias é a noção de comprimento de correlação da rede (Newman e Watts [28]), que regula quanto uma doença contagiosa deve crescer para sentir os efeitos dos atalhos (menores caminhos que ligam os vértices na rede). Com efeito, a

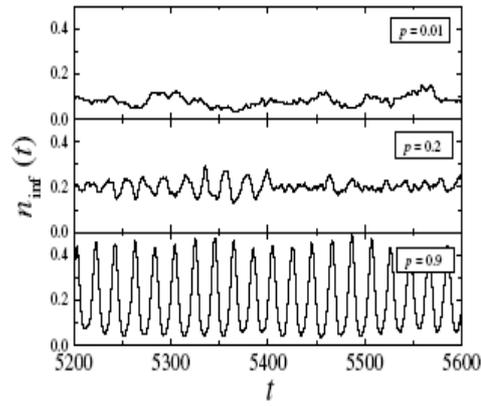


Figura 3.8: Fração de infectados em função do tempo no modelo SIRS em redes construídas de acordo com o modelo de Watts-Strogatz. Três curvas são mostradas para diferentes valores do parâmetro de desordem p , como mostram as legendas (Kuperman e Abramson [20]).

epidemia pode ser dividida em dois estágios: (a) Uma fase local de crescimento lento, pela presente redundância na topologia local da rede. (b) Uma fase de crescimento global, em que pode crescer exponencialmente por toda a rede via atalhos aleatórios. Estes modelos sugerem que uma das estratégias para prevenir epidemias deverá se focar na eliminação dos atalhos, prolongando assim o crescimento local o máximo possível.

Explorando o crescimento de uma epidemia em topologias distintas (figura 3.9), Shirley e Rushton [23] investigaram, por intermédio de simulações, a relação deste crescimento com as propriedades estatísticas das redes. O modelo que usaram, o SEIR, inclui uma nova classe, E, denominada de incubação. Os resultados apresentados por eles mostram que a taxa de infecção é inversamente proporcional ao menor caminho médio, por outro lado, o tamanho máximo da epidemia resulta ser diretamente proporcional. Verificaram também que as propriedades epidêmicas são sensíveis à variação da probabilidade de infecção, β (figura 3.10). Sintetizando, a performance em resposta às características epidêmicas é governada pelo tamanho do menor caminho médio.

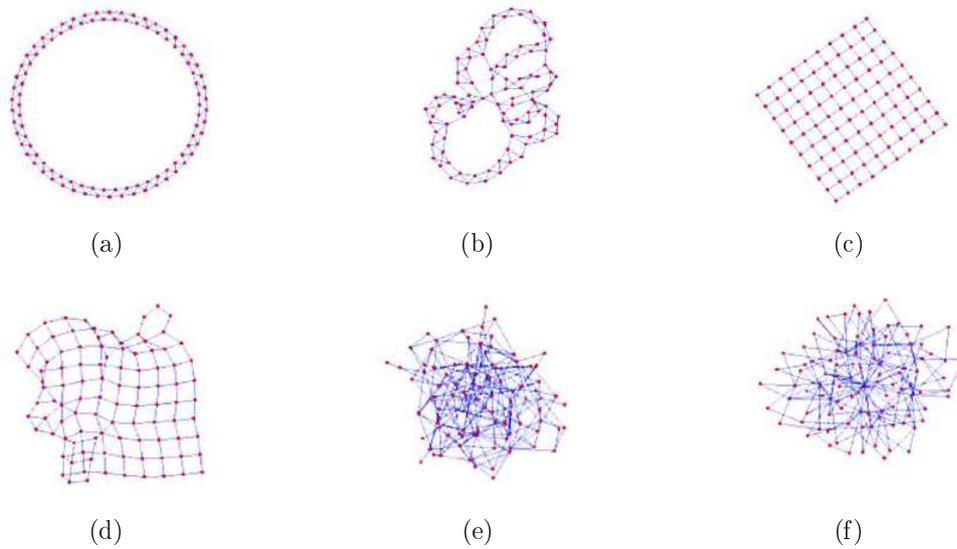


Figura 3.9: Ilustrações dos tipos de topologias investigadas no trabalho de Shirley e Rushton [23]. (a) Rede regular no anel (1D). (b) Rede de mundo pequeno no anel (1DR). (c) Rede regular no plano (2D). (d) Rede de mundo pequeno (2DR). (e) Rede aleatória (RG). (f) Rede de escala livre (SF).

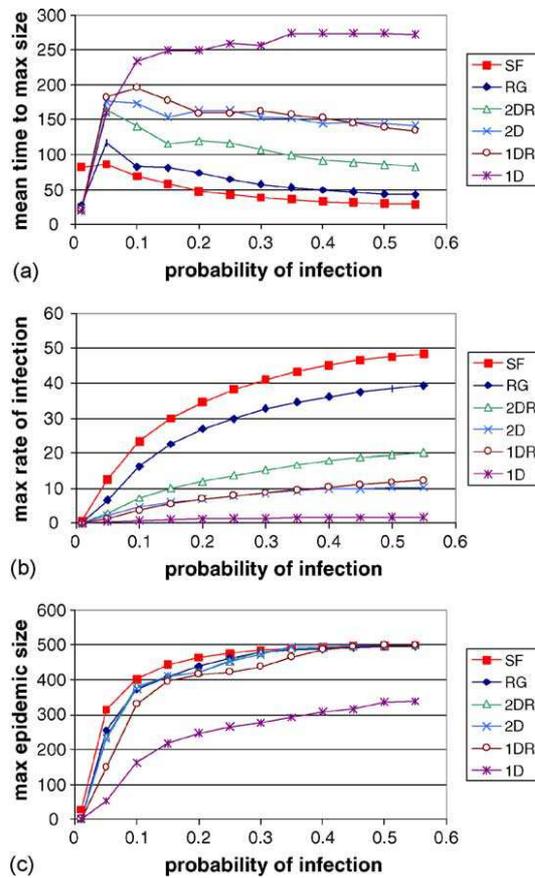


Figura 3.10: A sensibilidade do modelo SEIR à probabilidade de infecção β , nas diferentes estruturas de redes (Shirley e Rushton [23]). (a) Tempo médio para atingir o tamanho máximo da epidemia. (b) Máxima taxa de infecção. (c) Tamanho máximo da epidemia.

Capítulo 4

Dinâmica SIR nas Redes Paradigmáticas e Romântica

Como discutimos no capítulo anterior as redes paradigmáticas, como mundo pequeno e livre de escala, têm sido recentemente estudadas no contexto da transmissão de doenças [23]. Entretanto, devemos destacar que cada doença tem sua particular rede de transmissão. Por exemplo, as redes de contatos sexuais podem ter uma topologia bem diferente das redes de contatos que dão suporte a transmissão de doenças como a gripe. Esta distinção entre as estruturas das redes, estimula a investigação de redes realísticas e seu paralelo com as redes paradigmáticas, cujas propriedades topológicas são bem conhecidas. Assim, vemos que o conhecimento referente aos processos epidemiológicos amplia-se pelo conhecimento da topologia da rede associada à doença infecciosa em questão.

Devido ao recente levantamento detalhado de uma rede real de contatos amorosos, neste capítulo investigaremos a propagação de uma doença sexualmente transmissível hipotética nessa rede real, tomando um paralelo de tal investigação com as redes paradigmáticas. Os resultados que apresentaremos são todos dados em termos de simulações computacionais. Contudo, nas seções que antecedem a estes resultados, fazemos referência à medidas das propriedades topológicas das redes bem como uma descrição do modelo computacional que estamos utilizando.

4.1 Rede Romântica

Em 2004 foi publicado o primeiro levantamento completo de uma rede de contatos chamada pelos autores de “romântica” (Bearman et al. [29]). A rede em questão corresponde aos contatos

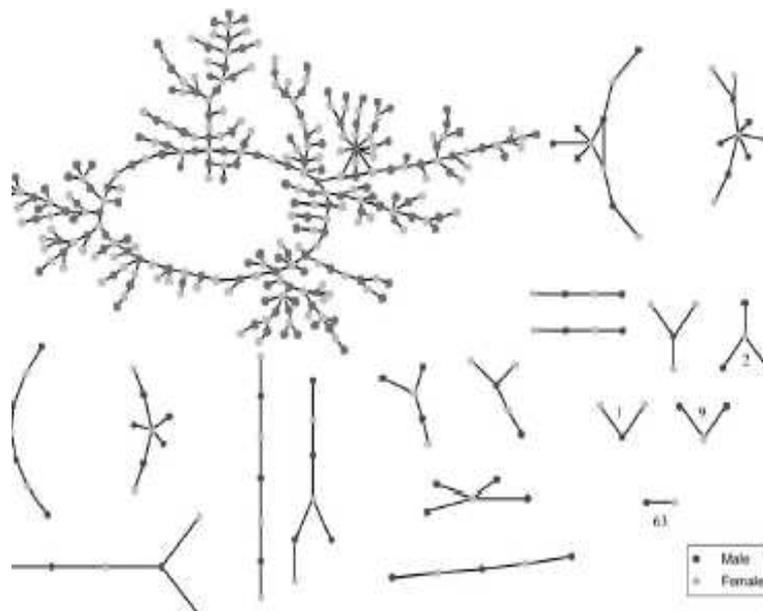


Figura 4.1: Rede romântica da Jefferson High School (Bearman et al [29]). Cada círculo representa um estudante, escuro do sexo masculino, claro do sexo feminino, e as linhas conectando os estudantes representam relações românticas. Os números abaixo da figura contabilizam o número de vezes que o padrão foi observado (por exemplo, 63 pares isolados).

entre adolescentes de uma escola mista de ensino médio de uma cidade de interior dos Estados Unidos, cujo nome fictício dado pelos autores é Jefferson High School. Todos os estudantes da escola (mais de 1000 alunos) foram entrevistados quanto aos seus parceiros na escola, sendo solicitado que identificassem o nome deles (foi garantido que os resultados seriam confidenciais e só usados para contrastar e validar as respostas), sendo que 90% responderam à pesquisa. Um *link* entre dois estudantes é criado quando ambos referem mutuamente o contato. Tanto os contatos românticos (namoros) como os sexuais foram relatados e identificados no tempo, abrangendo dezoito meses, sendo identificada uma rede de 573 indivíduos. O resultado mais importante foi o reconhecimento de uma componente gigante da rede (CG), envolvendo 288 estudantes, se bem que essa componente resulta quando são super impostos todos os contatos sem importar o tempo. É relevante acrescentar que cada integrante desta componente não percebe que pertence a mesma, visto que sua percepção de interação limita-se tanto ao local quanto ao tempo do contato.

4.2 Redes Estáticas

Várias topologias de redes, construídas por algoritmos computacionais, redes paradigmáticas, tem sido exploradas no contexto da transmissão de doenças na literatura, de acordo com o que vimos na seção 3.3. Motivados por estes trabalhos e pelo importante reconhecimento da componente gigante na rede romântica, compararemos os resultados, nas próximas seções, de uma propagação hipotética de uma doença sexualmente transmissível entre as estruturas modelo e real (figura 4.2). Para tanto, a construção das redes paradigmáticas é dada de acordo com os modelos de Watts-Strogatz e Albert-Barabási (seção 2.3), enquanto que a rede real é dada pelo mapeamento das interações de cada indivíduo na rede publicada no trabalho de Bearman et al. [29]. O número de vértices tomados em todas as redes é igual ao da componente gigante da rede romântica. Fixado o número de indivíduos, utilizamos o modelo de Watts-Strogatz para construir as redes regular com grau inicial $\kappa = 2$ ($p = 0$), mundo pequeno ($p = 0.05$) e aleatória ($p = 1$) em que p é o parâmetro de desordem. Com o modelo de Albert-Barabási construímos uma rede livre de escala ($N_0 = 2$ e $m = 1$) sendo que N_0 é o número de vértices ligados aleatoriamente (neste caso, sendo apenas dois, estão ligados inicialmente) e m é o número de *links* de cada vértice introduzido em cada passo de tempo.

4.3 Propriedades Estruturais das Redes Estáticas

As distribuições de grau das cinco redes são dadas na figura 4.3. Referente às redes geradas pelo modelo de Watts-Strogatz, embora o processo de reescrita das arestas não afete o grau médio $\langle k \rangle$ (tabela 4.1), a distribuição do grau muda, sendo no caso regular uma delta e no caso extremo de alta desordem uma distribuição de Poisson. A rede livre de escala, gerada pelo modelo de Albert-Barabási, é uma lei de potência, de acordo com as previsões do modelo. A distribuição do grau da rede real por sua vez é semelhante a da rede aleatória.

O desvio padrão de k , σ_k , é uma medida da variabilidade na conectividade. A rede regular possui $\sigma_k = 0$ uma vez que possui grau igual para todos os seus vértices, $k = 2$. Por outro lado, a rede livre de escala é que possui maior variação em sua conectividade; σ_k numa distribuição deste

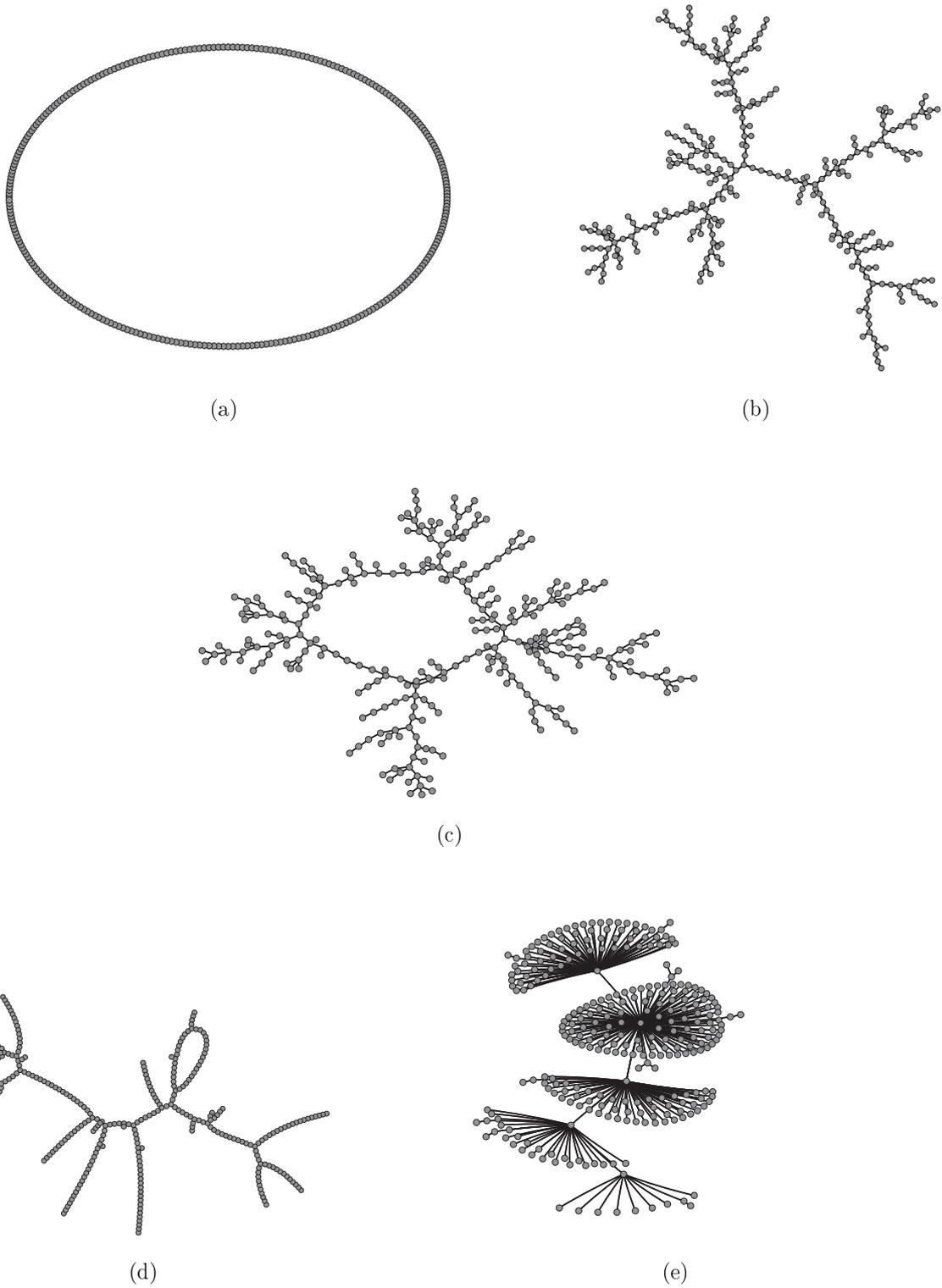


Figura 4.2: Exemplos de realizações das redes utilizadas nos cálculos: regular (a), aleatória (b), CG romântica (c), mundo pequeno (d) e livre de escala (e). Todas com o mesmo número de vértices ($N = 288$) e igual número médio de arestas por vértice ($\langle k \rangle = 2$). Figuras geradas com o PAJEK [30].

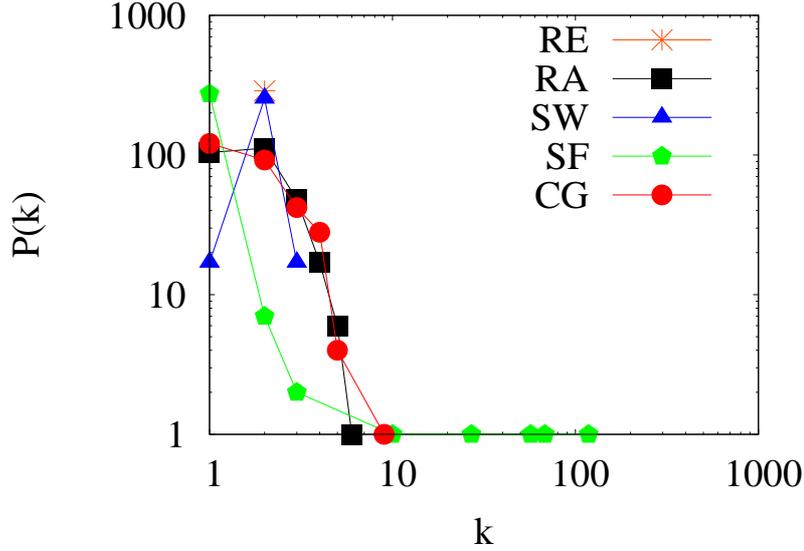


Figura 4.3: Distribuição de grau para cada uma das redes: (RE) regular, (RA) aleatória, (CG) CG romântica, (SW) mundo pequeno e (SF) livre de escala. Note que para a rede regular há somente um ponto ($k = 2$).

Tabela 4.1: Medidas básicas das propriedades topológicas: Grau médio $\langle k \rangle$, desvio padrão do grau σ_k , clusterização média C e menor caminho médio ℓ .

medidas	regular	aleatória	mundo pequeno	livre de escala	CG romântica
$\langle k \rangle$	2	2	2	~ 2	~ 2
σ_k	0	1.00	0.34	8.71	1.11
C	0	0	0	0	0
ℓ	72.25	14.70	19.32	3.16	16.14

tipo cresce como o número de vértices. Entre as demais redes, destacamos a proximidade de σ_k para a rede aleatória e CG romântica.

Em termos da clusterização média, C , todas as redes possuem valores nulos. Dentro do contexto de interações sexuais, $C = 0$ é resultado do comportamento majoritariamente heterossexual da população. Observando a rede real, vemos que a mesma é, na sua maioria, formada por indivíduos heterossexuais e, por conseguinte, tomamos as redes paradigmáticas dentro desse comportamento sexual.

Referindo-se ao menor caminho médio ℓ , vemos que a rede livre de escala possui o menor valor, fruto da ligação preferencial entre os vértices, por outro lado, a rede regular possui o maior, devido a estrutura altamente organizada das arestas. Entre as redes aleatória, mundo pequeno e CG romântica, observamos a proximidade de ℓ .

De acordo com os parágrafos anteriores, observamos que CG romântica não pode ser classificada como nenhuma das redes paradigmáticas apresentadas, embora pelas quatro medidas acima, ela está muito próxima de uma rede aleatória.

4.4 Simulação Computacional

As propriedades epidêmicas das redes são exploradas em simulações computacionais. Primeiramente, uma rede de uma determinada topologia é construída, onde cada vértice representa um indivíduo e cada aresta um contato entre os mesmos. Esta rede é representada (armazenada no programa computacional) por uma matriz de adjacência binária (seção 2.1). A seguir, um único

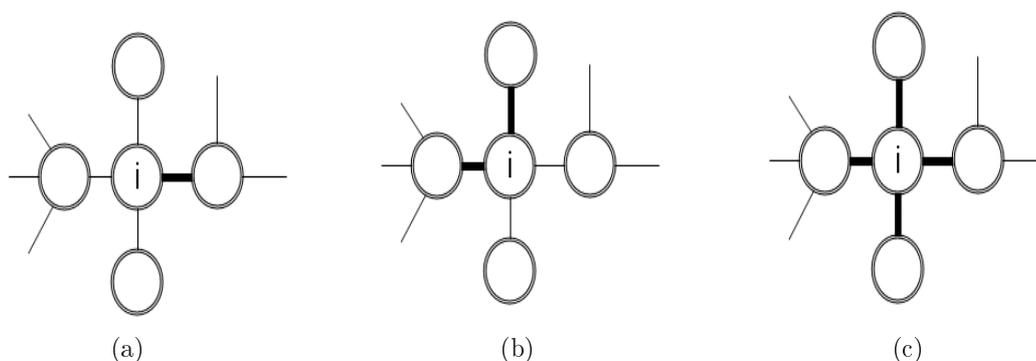


Figura 4.4: Representação esquemática das regras de transmissão para um indivíduo i e seus vizinhos; as arestas em destaque representam as interações entre os indivíduos num dado tempo. Em termos das regras de transmissão tomaremos: (a) aleatoriamente um vizinho (*one*); (b) uma fração aleatório de vizinhos (*fraction*); (c) todos os vizinhos (*all*).

indivíduo (vértice da rede) é escolhido aleatoriamente para ser infectado (estado I), enquanto os $N - 1$ restantes são colocados no estado suscetível (S). Esta configuração determina a condição inicial do sistema, a qual, pela escolha aleatória do indivíduo I, depende da semente do gerador de números aleatórios. Feito isto, a dinâmica de propagação da doença procede em passos discretos de tempo da seguinte forma: Em um passo de tempo, cada vértice da rede escolhe um vértice (entre a vizinhança) para interagir (os N vértices são atualizados em paralelo). Na interação, o indivíduo suscetível torna-se infectado com uma dada probabilidade, β , caso o vizinho selecionado esteja no estado infectado¹. O vértice infectado permanece assim até que transcorra um tempo

¹Na dinâmica o contrário é válido também, se o indivíduo está no estado I pode contaminar seu vizinho.

τ , no qual o indivíduo em questão torna-se imune ou removido; β e τ são constantes para todos os indivíduos na população. Estudamos a propagação de uma hipotética doença sexualmente transmissível na rede romântica (sífilis, AIDS, etc). Portanto, quando falamos de interações nos referimos explicitamente a contato sexual. O passo de tempo pode representar um dia ou uma semana; se um passo de tempo fosse um dia, estaríamos supondo que **todos** os indivíduos têm algum tipo de relação sexual com algum de seus parceiros (ou até com mais de um), todos os dias. Provavelmente isso é exagerado, o que levaria a sobre-estimação da gravidade da epidemia; mas esse mesmo passo de tempo pode ser assimilado a uma semana. Porém, não pretendemos aqui calibrar de maneira precisa os parâmetros da doença para predizer um surto, e sim comparar os resultados de rede para rede. Por isso mesmo toda uma gama de valores é explorada desde o mais exagerado até o altamente improvável, via variação de β e τ , incluindo valores que podem estar perto de parâmetros realistas. Adicionalmente, dentro de um contexto real esperamos, intuitivamente, que as interações românticas entre os indivíduos ocorram com uma variabilidade na frequência. Com efeito, investigamos situações distintas para o número de interações dos indivíduos dentro da sua rede de contatos, as quais denominamos de regras de transmissão. A primeira considera, por passo de tempo, uma interação para cada indivíduo, a qual é escolhida aleatoriamente dentro do seu conjunto de vizinhos (como foi descrito acima); denominamos esta regra de *one*. Na segunda, *fraction*, uma fração aleatória do número de vizinhos é considerada em cada passo. Na terceira, a que chamaremos de *all*, cada indivíduo interage simultaneamente com todos os vizinhos. Para cada uma destas regras, a dinâmica é repetida um número discreto de passos até o estado final onde não há mais indivíduos infectados. De interesse nesse procedimento são a dinâmica propriamente (evolução temporal do número de infectados) e o estado final e a sua relação com os parâmetros da infecção. Os resultados estão restritos a esses dois parâmetros da doença β e τ , ao indivíduo inicialmente tomado como infectado e as três regras de transmissão. O conjunto de resultados envolvidos corresponde a β entre 0 e 1 em passos de 0.01, enquanto que τ varia entre 1 e 30 em passos de 1. Ainda tomamos uma estatística sobre 100 sementes. Realizamos este conjunto de situações para cada regra de transmissão.

4.5 Resultados

Os resultados que apresentaremos foram obtidos considerando as interações entre os indivíduos da componente gigante da rede romântica, ora como estáticas ora como dinâmicas. Primeiramente investigamos as propriedades epidêmicas em todas as redes consideradas (redes paradigmáticas e CG romântica estática) tomando as arestas invariantes no tempo, conforme fixado na distribuição inicial (figura 4.3). Em seguida, analisamos os resultados considerando a formação dinâmica de pares na CG da rede romântica dinâmica.

4.5.1 Rede Estática

Partimos da investigação do comportamento da dinâmica das classes SIR nas diferentes topologias apresentadas na seção 4.2. Tomamos um caso particular por vez, fixando β , τ e a regra de transmissão para observarmos o efeito da estrutura das redes na propagação da doença. Na figura 4.5 vemos que a estrutura influencia significativamente na evolução temporal das diferentes classes. Verificamos ainda nesta figura que o número de infectados num determinado instante é maior na rede livre de escala. Referente ao tempo em que a rede permanece com indivíduos infectados, notamos que o tempo maior ocorre na estrutura regular. Em outras palavras, a dinâmica é mais lenta na rede regular, enquanto, no outro extremo temos a rede livre de escala com a dinâmica mais rápida. A tabela 4.2 reforça esta afirmação mesmo para uma estatística sobre 100 sementes. No entanto, as curvas das dinâmicas das classes apresentadas concebem

Tabela 4.2: Tempo máximo médio $\langle t_{max} \rangle$ e desvio padrão do tempo máximo $\sigma_{t_{max}}$ para cessar a dinâmica ($I = 0$) sobre 100 sementes.

medidas	regular	aleatória	mundo pequeno	livre de escala	CG romântica
$\langle t_{max} \rangle$	580.6	106.1	165.7	44.8	129.2
$\sigma_{t_{max}}$	201.2	39.9	45.3	5.8	46.5

uma especial situação para uma dada semente e, conseqüentemente, um particular indivíduo inicialmente infectado. Este particular indivíduo infectado é relevante pois o grau de cada um é diferente; só na rede regular é irrelevante, pois todos os vértices são equivalentes. Porém, mesmo

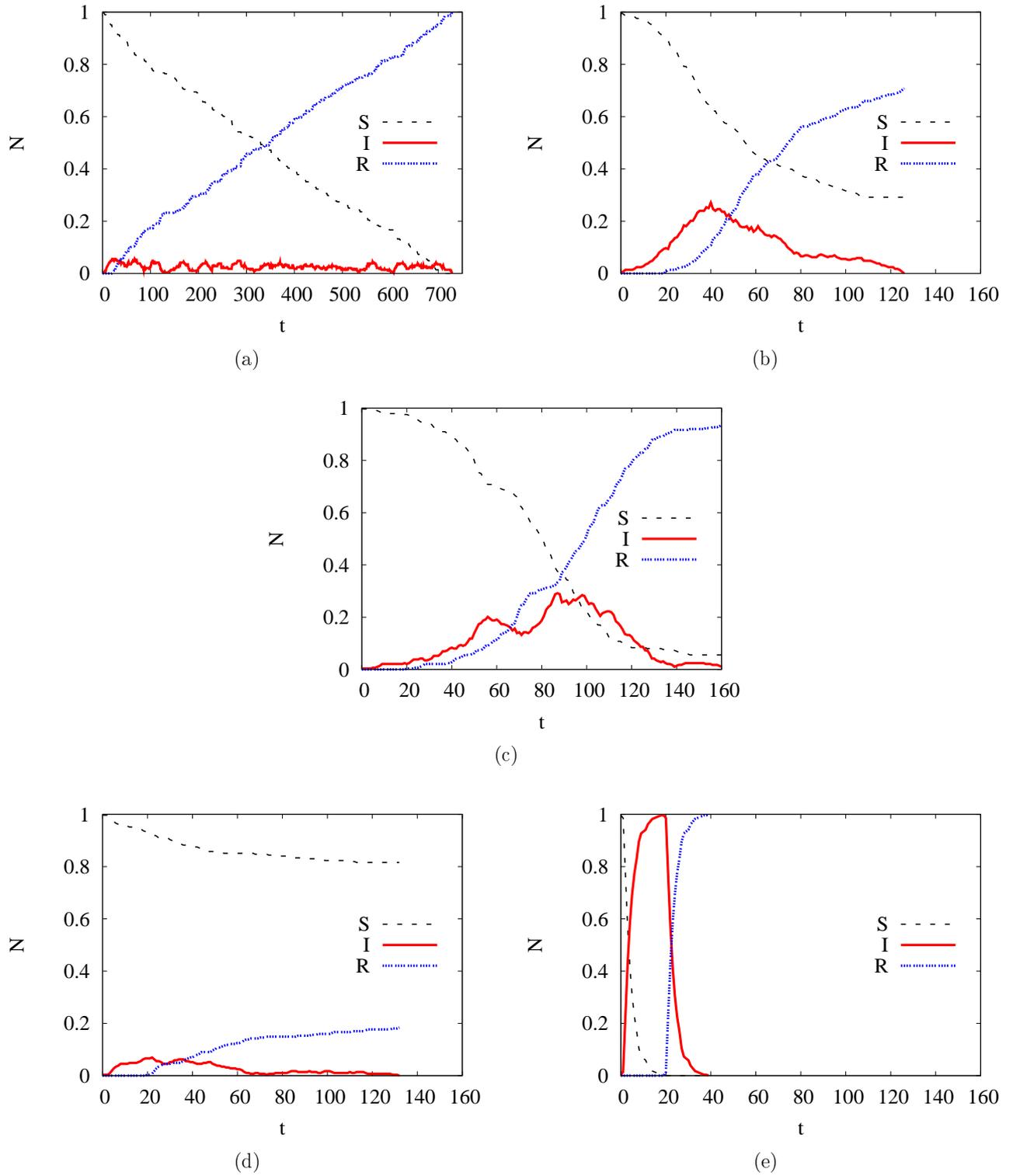


Figura 4.5: Dinâmica dos estados suscetível (S), infectado (I) e removido (R) nas diferentes topologias: (a) regular, (b) aleatória, (c) CG romântica, (d) mundo pequeno e (e) livre de escala. Simulações realizadas com $\beta = 0.2$ e $\tau = 20$ para a regra de transmissão *fraction*.

na rede regular, faz se necessário uma média sobre as sementes, pois elas determinam a seqüência de Markov. Averiguamos para os mesmos parâmetros o efeito da mudança na escolha da semente. Assim, mostramos na figura 4.6 a evolução temporal dos removidos para 100 sementes diferentes em cada uma das topologias estudadas. Podemos apreciar que considerar sementes distintas altera a dinâmica dos infectados e conseqüentemente o estado final da população.

Entre os resultados, percebemos que o impacto da troca de semente é maior, medido pela área cinza, nas redes aleatória e romântica, quando comparado as demais redes. Mas, analisando em detalhe esses resultados podemos extrair as seguintes conclusões:

- A influência na rede livre de escala se dá só no disparo da epidemia, sendo que, tanto a evolução posterior quanto o estado assintótico são invariantes. Isto é devido a que nessas redes a dinâmica está regida pelos *hubs*, uma vez infectados, a dinâmica é quase determinista fruto do grande número de conexões deles; o reflexo da distância entre o infectado inicial e o *hub* mais próximo.
- A rede regular, a despeito de ter todos seus vértices equivalentes, mostra a menor dispersão na dinâmica, dependendo da semente, e até do estado final da epidemia, provocando, com 100 sementes, entre 0-25% da população atingida pela epidemia. Isto ilustra o caráter local da disseminação da doença, visto que seus vértices interagem apenas a primeiros vizinhos.
- A rede aleatória apresenta, uma grande dispersão na dinâmica frente a escolha de sementes; mas, dada a falta de regularidade desta rede, apresentando estados assintóticos com 0-62% de removidos.
- A rede romântica, apresenta uma variabilidade da dinâmica muito parecida com a da rede aleatória, cai numa faixa de 0-77% de removidos.
- A rede de mundo pequeno, com uma dispersão entre a rede regular e aleatória, mostra um estado assintótico 0-35% de removidos.

Em resumo, no que diz respeito a variabilidade frente a sementes (neste problema isto tem o caráter de condições iniciais) a rede SF apresentam um estado assintótico bem definido, robusto

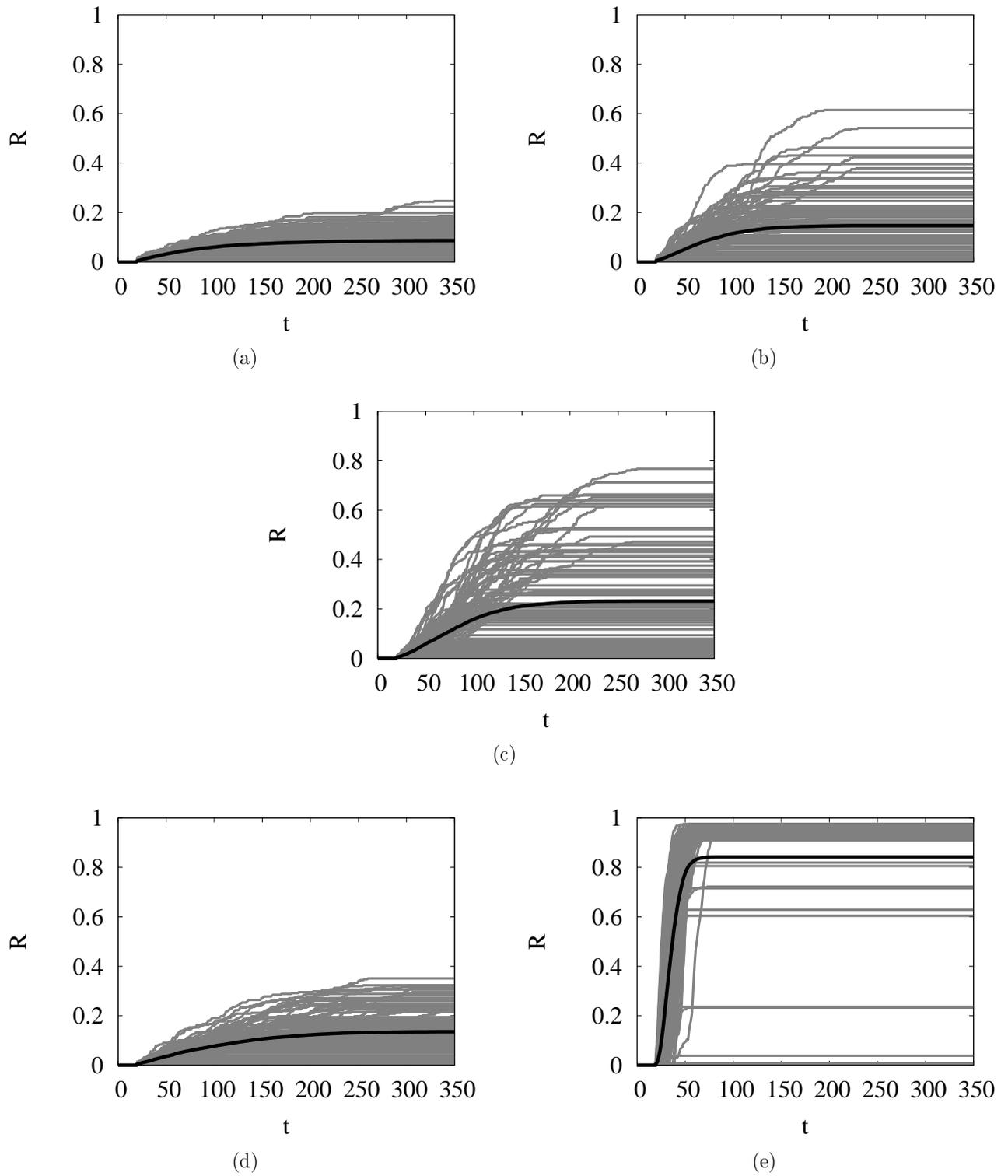


Figura 4.6: Evolução temporal da fração dos removidos para as diferentes redes: regular (a), aleatória (b), CG romântica (c), mundo pequeno (d) e livre de escala (e). As curvas em cinza são para semente distintas, total de 100 sementes, enquanto que a curva em preto é uma média sobre estas curvas. Simulações realizadas, particularmente, com $\beta = 0.2$ e $\tau = 20$ para a regra de transmissão *fraction*.

frente a variabilidade estocástica do modelo. Por outro lado, a rede aleatória, extremo da representação das redes reais, apresenta grande variabilidade na dinâmica que se projeta ao estado final (notar que esta situação é que torna quase inviável as predições sobre epidemias, pelo menos como gostaríamos de poder fazer). Um pouco acima, aparece a rede romântica, próxima das anteriores em relação a dinâmica com um percentual médio de removidos de $\sim 20\%$.

Feita esta primeira análise da variação da dinâmica dos infectados, dependente da estrutura das redes e da escolha da semente, passaremos a investigar a sensibilidade das topologias as regras de transmissão. Para tanto, apresentaremos os resultados da fração final média de removidos, $\langle R \rangle$, versus o número básico de reprodução $R_0 = \langle k \rangle \beta \tau$, onde $\langle k \rangle$ é o número médio de contatos dos indivíduos. O número de diferentes indivíduos tomados inicialmente e os valores de β e τ estão de acordo com o indicado na seção 4.4. Na figura 4.7 vemos que a mudança de regra de transmissão tem efeitos mais relevantes sobre as redes aleatória, livre de escala e CG romântica, com pouco efeito sobre as redes regular e mundo pequeno. Claramente, todas as regras de transmissão aplicadas a CG romântica e aleatória levam a um estado final bem distinto, enquanto que para a rede livre de escala as regras *fraction* e *all* são muito próximas. Na figura 4.8 podemos ver o comportamento das diferentes redes nas três regras de transmissão. Para valores de $R_0 \lesssim 1$, todas as redes, exceto a livre de escala, comportam-se similarmente, revelando que a topologia e a regra de transmissão adotada não influênciam na propagação da doença. Ao passo que aumentando R_0 ($R_0 \gtrsim 1$), tanto a estrutura como a regra de transmissão tornam-se importantes. Observamos que em todas as regras de transmissão $\langle R \rangle > 0$ inclusive para $R_0 < 1$, efeito típico de redes finitas. De forma geral, o estado final da epidemia da rede livre de escala é que possui mais acentuada diferença em todos cenários, quando comparada as demais redes. As redes geradas pelo modelo de Watts-Strogatz possuem estados finais da epidemia próximos para a regras de transmissão *one* e *fraction*. A componente gigante da rede romântica estática possui removidos assintóticos similares aos das redes regular, aleatória e mundo pequeno quando consideramos a regra de transmissão *one*. A CG romântica, para as regras *fraction* e *all*, comporta-se diferentemente no que refere-se ao tamanho final da epidemia. Referente aos cenários, podemos perceber que conforme vamos aumentando o número de interações dos indivíduos (*one* para *all*) por passo de tempo, mais

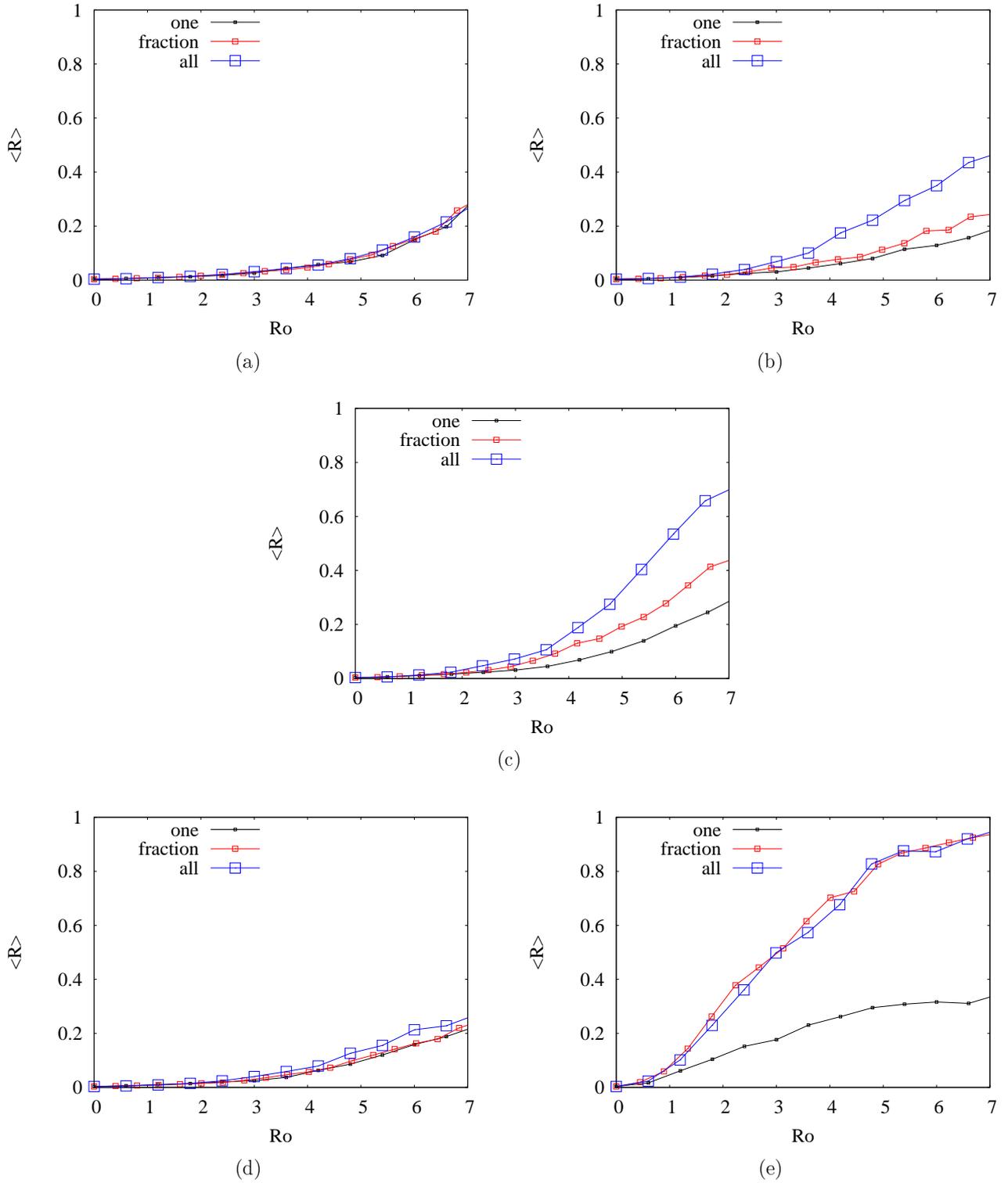


Figura 4.7: Regras de transmissão distintas aplicadas as redes: (a) regular, (b) aleatória, (c) CG romântica, (d) mundo pequeno e (e) livre de escala. Sendo *one* escolha aleatória de um vizinho, *fraction* um fração aleatória de vizinhos e *all* todos os vizinhos. Onde $\langle R \rangle$ são os removidos assintóticos e $R_0 = \langle k \rangle \beta \tau$ é o número básico de reprodução.

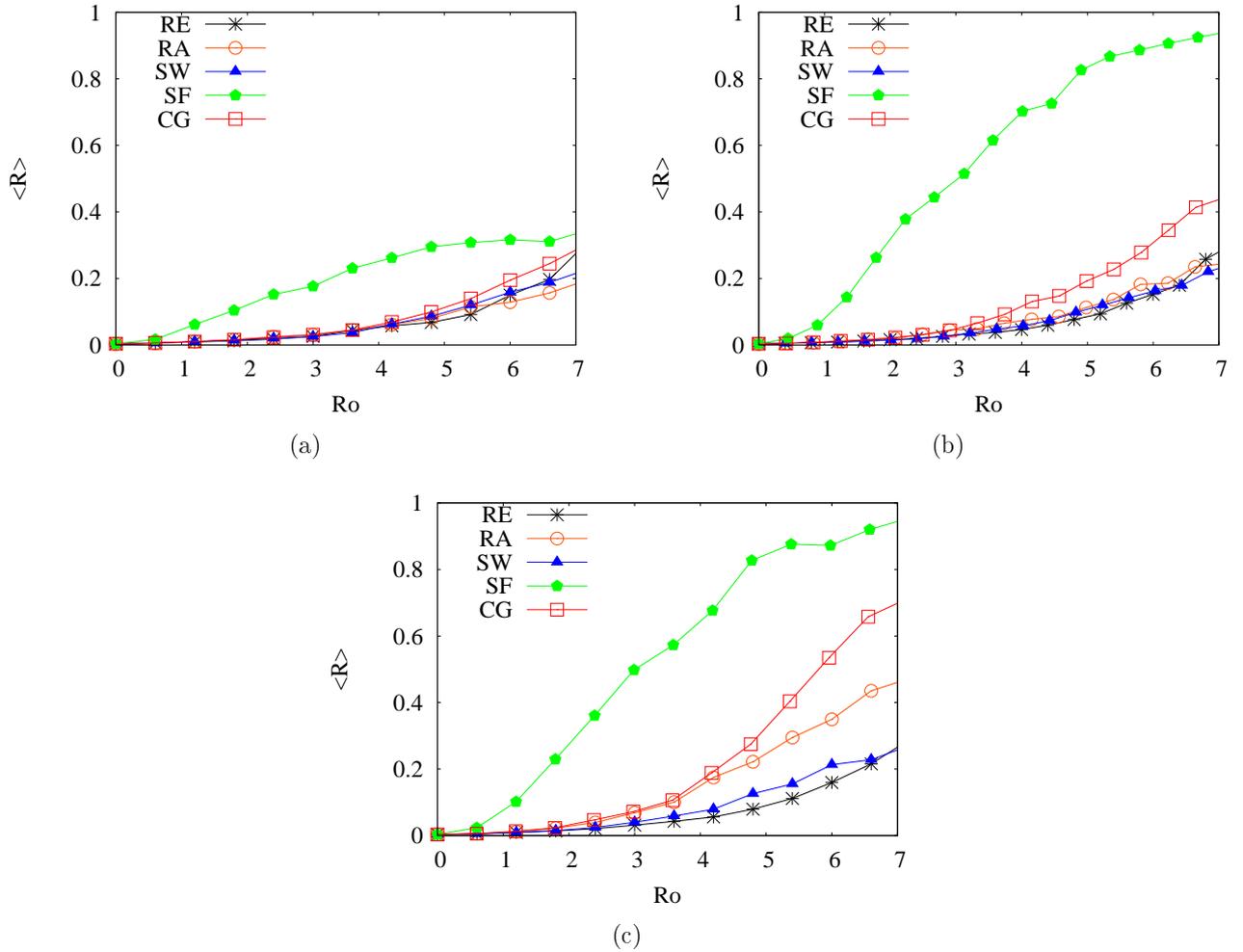


Figura 4.8: Curvas de infecção para as regras: (a) one, (b) fraction, (c) all. Redes regular (RE), mundo pequeno (SW), aleatória (RA), livre de escala (SF) e CG romântica (CG).

distinta a CG da rede real torna-se das redes paradigmáticas, dentro do contexto da propagação da epidemia.

4.5.2 Rede Dinâmica

Na rede romântica a característica mais relevante da rede estática é a componente gigante, conforme já mencionamos, conectando 288 estudantes. Entretanto, baseados na estrutura dinâmica publicada por James Moody (co-autor do trabalho [29]), vemos que a estrutura emerge completamente baseada na história local da rede. Portanto, não observamos alguns estruturas de interações ao longo do tempo. Em verdade, percebemos que as interações entre os alunos formam

grupos pequenos e isolados, quando tomamos um intervalo de tempo (figura 4.9).

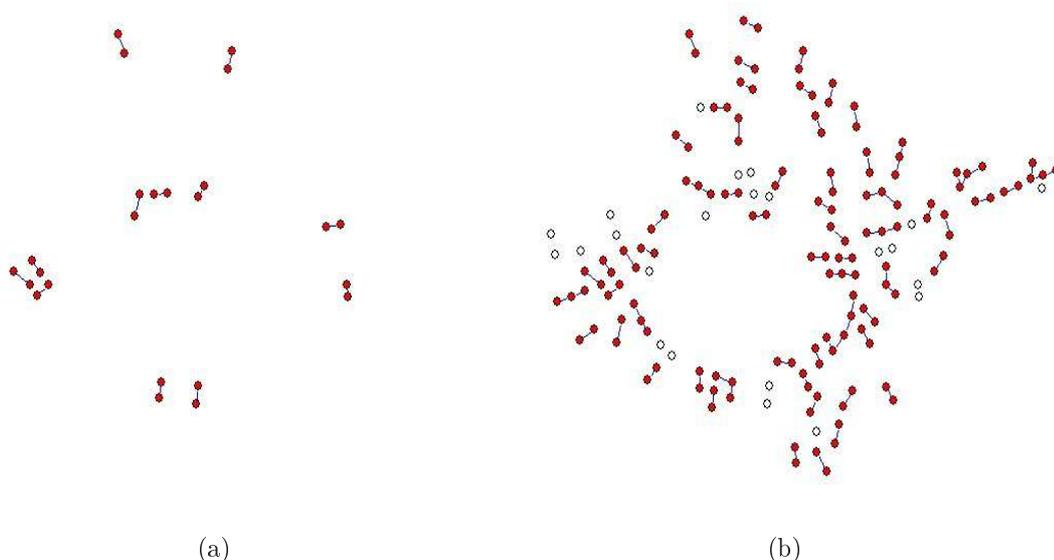


Figura 4.9: Dois tempos distintos na dinâmica da componente gigante da rede romântica na Jefferson High School. Os círculos hachurados representam alunos que estão interagindo, enquanto que os não hachurados representam alunos que não interagem para um dado tempo. Tempos: (a), *frame* 10 ou 120 dias e (b), *frame* 30 ou 360 dias. A dinâmica completa, *gif* animada, é composta por 45 *frames* equivalentes a 12 dias cada uma, e está publicada na página de Moody (<http://www.soc.duke.edu/~jmood77/NetMovies/index.htm>).

Estas verificações, referentes a dinâmica, são suficiente para alterar as propriedades estruturais da rede. A distribuição de grau dinâmica representada na figura 4.10(a), exibe um comportamento muito distinto na conectividade quando comparada a CG romântica estática (figura 4.3). Entre as diferenças sobressai-se a grande parcela de indivíduos sem contatos ($k = 0$), contemplada na distribuição para certo tempo. Ao longo do tempo destacamos: (a) o número de interações se dá, $t < 15$, entre pares de indivíduos isolados; (b) $30 < t < 40$ é o intervalo de tempo de maior atividade de interações, atingindo um máximo em $t = 35$ e (c) para $t > 40$ o grau diminui com o passar do tempo.

Outro aspecto que merece atenção é a fração do tempo da interação de cada aluno (figura 4.10(b)). Apuramos deste resultado que uma pequena fração dos alunos, $\sim 3\%$ da população, possui um grande tempo de interação, 90% do tempo total, enquanto que a maior parte da população, $\sim 70\%$ dos alunos, interage 40% do tempo total.

As descrições da distribuição de grau e da fração do tempo da interação dos indivíduos nos sugere que a propagação de uma doença, doença de dinâmica SIR, terá um impacto diferente

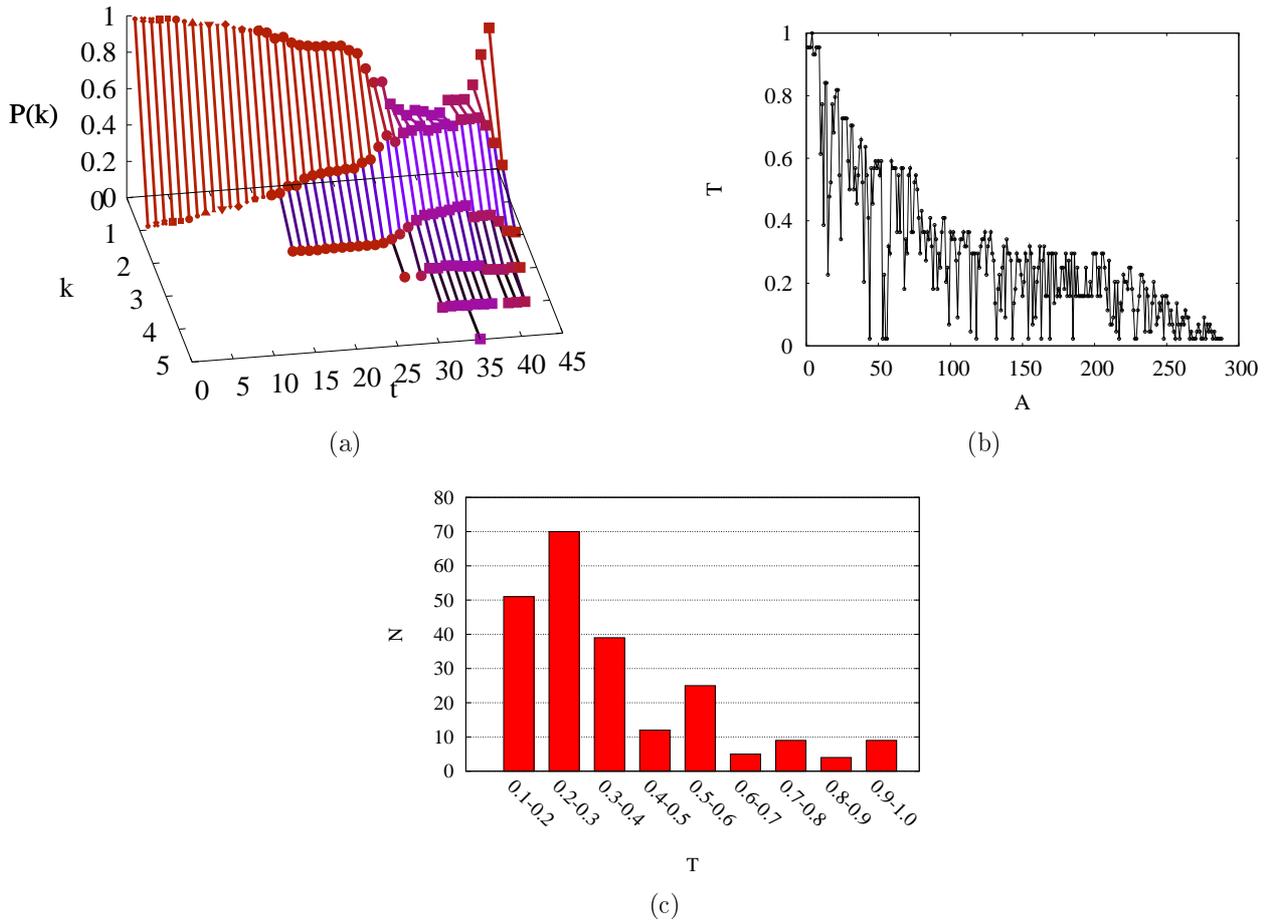


Figura 4.10: (a) Evolução temporal para distribuição de grau na CG romântica dinâmica. (b) Fração do tempo da interação, T , por aluno, A . (c) Histograma do número de alunos N por tempos de interação, como fração do tempo total do estudo.

quando confrontarmos a estrutura dinâmica e estática da CG romântica. Esta última frase justifica-se não só pela presença de pequenas componentes isoladas (grupo de indivíduos isolados) mas também pelo pequeno tempo, em média, de duração destas componentes. Desta maneira, o tamanho final da epidemia, $\langle R \rangle$, dependerá da relação entre o período de infecção τ e o tempo de interação dos indivíduos dentro da dinâmica da formação de contatos da rede. De fato, realizamos as simulações computacionais similarmente ao indicado na seção 4.4. Nesta simulação, cada estrutura de interações, *frame* da *gif*, permanece por 12 passos de tempo dentro da simulação. Este número de passos deve-se ao número de *frames* da *gif* animada, 45 no total, e o tempo de coleta dos dados, 540 dias.

No intervalo amplo de ρ explorado o tamanho final da epidemia, para todas as regras de transmissão, é regularmente pequeno ($\langle R \rangle \simeq 0.003$) e, notoriamente, independente dele acima do

valor crítico $R_0 = 1$. Este resultado para o número de removíveis assintóticos revela que nenhum indivíduo (diferente do inicialmente infectado), em média, é infectado. Portanto, não há epidemia neste cenário. Esta mudança brusca de comportamento, em contraste com a rede estática, deve-se exclusivamente à dinâmica das interações entre os indivíduos, tanto no que se refere ao tamanho das componentes quanto aos intervalos de tempo que elas perduram.

É possível diminuir o efeito da interação na transmissão da doença tomando um período de infecção (τ) muito maior que o tempo de permanência das componentes da rede. No limite de $\tau \rightarrow \infty$ (ou seja $R_0 \rightarrow \infty$), o modelo SIR se transforma em SI, no qual independente do valor de β toda a população termina no estado infectado (em algumas regiões este é o caso da infecção pelo HIV, não com τ infinito claro, mas igual à média de vida da população). Neste contexto, portanto, a quantidade relevante é a fração de infectados. Na figura 4.11(a) podemos observar que a dinâmica dos infectados é visivelmente sensível a semente. Desta dinâmica, temos que o

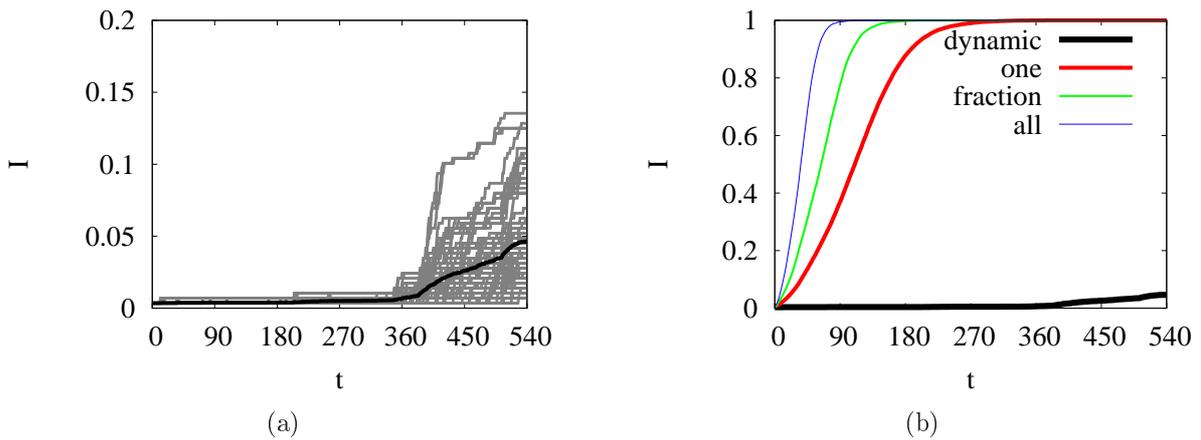


Figura 4.11: Evolução temporal da fração dos infectados para uma doença de dinâmica SI. As simulações foram realizadas para $\beta = 0.2$. (a) Rede CG romântica dinâmica. As curvas em cinza correspondem a distintas sementes, total de 100 sementes, enquanto que a curva em preto é uma média sobre estas sementes. (b) Redes CG romântica dinâmica (*dynamic*) e estática aplicada as regras de transmissão: *one*, *fraction* e *all*. Cada curva é uma média sobre 100 sementes.

número de infectados cresce substancialmente a partir de $t > 360$ este resultado está relacionado com a distribuição de grau (figura 4.10(a)) na qual, a partir deste tempo, o número de indivíduos com grau maior que zero torna-se expressivo. Na figura 4.11(b) comparamos a rede dinâmica e estática dentro do mesmo cenário, a diferença é evidente. A rede estática, aplicada as três regras de transmissão, espalha mais a doença que a rede dinâmica. Devido às nossas considerações

iniciais para a dinâmica SI da doença, a alteração do comportamento deve-se apenas a estrutura da evolução dos contatos (componentes isoladas e a fração dos indivíduos que não participam da dinâmica em cada tempo).

Capítulo 5

Conclusões

Neste trabalho, investigamos as propriedades epidêmicas de uma rede real de interações românticas (estática e dinâmica), recentemente publicada por Bearman et al [29]. Para tanto, incorporamos uma hipotética doença de dinâmica SIR. Confrontamos o tamanho final da epidemia e a dinâmica de infectados entre a rede romântica estática e as redes paradigmáticas construídas de acordo com o modelo de Watts-Strogatz [5] e Albert-Barabási [8]. No contexto das redes estáticas, investigamos a propagação da epidemia considerando três distintas formas de interação entre um indivíduo e seus vizinhos, a qual batizamos de regras de transmissão. Ainda, estudamos a dinâmica dos infectados considerando a formação dinâmica de pares na rede romântica.

Anterior as simulações da propagação da epidemia, extraímos informações sobre a topologia das redes real e paradigmáticas. Através de medidas como distribuição de grau, grau médio e menor caminho (veja figura 4.3 e tabela 4.1), mostramos que a topologia da componente gigante da rede romântica estática é distinta das demais redes. Comparando este conjunto de medidas entre a rede romântica e as demais, observamos que a diferença é maior para a rede livre de escala e, por outro lado, ela está mas próxima de uma rede aleatória.

Particularmente, Shirley e Rushton [23] afirmaram que o tempo para cessar a dinâmica da doença é determinado pelo menor caminho médio. De fato, comparando o menor caminho médio, na tabela 4.1, com o tempo máximo para cessar a dinâmica, na tabela 4.2, observa-se que os valores de ambas grandezas verificam a mesma relação de ordem: regular > mundo pequeno >

CG romântica > aleatória > livre de escala.

Mostramos o efeito das regras de transmissão no tamanho do estado final da epidemia, $\langle R \rangle$. Nossos resultados indicaram diferenças relevantes no número de removidos assintóticos para regras distintas (figura 4.7), nas redes livre de escala, CG romântica estática e aleatória. Observando a tabela 4.1 vemos que as redes anteriormente citadas são as que apresentam maior flutuação na conectividade; sugerindo que as diferenças no tamanho final da epidemia, devido a variabilidade na frequência de interações, estão ligadas ao desvio padrão do grau.

Comparamos $\langle R \rangle$ entre as diferentes topologias para uma mesma regra de transmissão (figura 4.8). Nesta comparação, observamos que a rede real em todas as regras de transmissão diferencia-se muito da rede livre de escala. Por outro lado, a CG romântica aproxima-se muito dos resultados obtidos para as redes geradas pelo modelo de Watts-Strogatz. Entre estas redes, vimos que a rede aleatória é a mais próxima em todas as regras de transmissão. Este ponto é muito relevante pois tem se falado muito sobre as redes livre de escala em referência a redes sociais. Porém na rede romântica, um exemplo de rede real, (pequena claro), ele se manifesta, quanto a propagação de doença na estrutura estática, como uma rede aleatória, como as introduzidas por Erdős e Rényi [4]. Existe um trabalho de levantamento de dados de contatos sexuais na Suécia (amostra da população, não apenas uma escola), onde emerge uma distribuição de contatos com lei de potência [31]. Porém, o levantamento dessa rede foi baseado nas respostas de questionários abrangendo 12 meses (ou seja apelando a subjetividade de cada indivíduo) e sem possibilidade de cruzamento de respostas como no trabalho da Jefferson School.

Sob o ponto de vista da dinâmica de formação de pares, medimos tanto a evolução da distribuição de grau bem como a fração de tempo de interação de cada indivíduo na rede romântica (figura 4.10). Da distribuição de grau verificamos que há intervalos de tempo em que o número de interações entre os indivíduos é maior, correspondente a 30% do tempo total. Baseados na fração de tempo da interação dos indivíduos, obtemos que mais do 50% dos alunos participam da dinâmica num tempo menor de 40% do tempo total.

O tamanho final da epidemia na CG romântica dinâmica revelou que nenhum indivíduo (diferente do inicialmente infectado) é infectado, em média. Justifica-se o resultado não só pela

presença de pequenas componentes isoladas (contempladas na distribuição de grau), mas também pelo curto tempo de duração destas componentes (devido ao tempo de interação dos indivíduos). Este resultado sugere que o tempo de interação dos indivíduos frente ao tempo de infecção deve ser considerado. Verificamos que o efeito desta relação de tempos diminui quando assumimos que o tempo de infecção é da ordem do tempo de observação da dinâmica da rede, dezoito meses. Em particular confere uma dinâmica de doença do tipo SI. Ao comparar com os resultados da estrutura estática (figura 4.11(b)), vemos que o cenário dinâmico traz predições bem menos graves, em termos de epidemia, que a versão estática (arestas permanentes).

A partir dos resultados que foram expostos, concluímos que a rede livre de escala é a rede paradigmática que mais diverge, em termos de propriedades epidêmicas, da rede real. Existem referências na bibliografia à possível relação entre as redes livres de escala e redes realísticas de interações sexuais, mostramos neste estudo que a rede aleatória é a que mais se aproxima da rede real de interações românticas. Por fim, incorporando a dinâmica real na formação de pares, obtivemos resultados distintos nas propriedades epidêmicas quando comparados a rede estática. A evolução temporal da epidemia na rede romântica dinâmica, apresentada neste trabalho, não pode ser considerada como uma predição precisa (o que por outro lado não existe em epidemias), mas é um forte indício de que a dinâmica nas interações entre agentes não pode ser omitida nos cálculos “realísticos” de epidemias. Por um lado isto relativiza a teoria padrão de epidemias, modelos tipo SIR, baseados em equações diferenciais. Por outro, é uma boa notícia, pois as predições de surtos e tamanhos de epidemias são menores do que as baseadas no modelo padrão.

Referências Bibliográficas

- [1] Albert, R. and Barabási, A.-L., *Rev. Mod. Phys.* **74** (2002) 47.
- [2] Newman, M., *SIAM Review* **45** (2003) 167.
- [3] Costa, L. d. F., Rodrigues, F. A., Travieso, G., and Boas, P. R. V., *Advances in Physics* **56** (2007) 167.
- [4] Erdos, P. and Renyi, A., *Publ. Math. Debrecen* **6** (1959) 290.
- [5] Watts, D. J. and Strogatz, S. H., *Nature* **393** (1998) 440.
- [6] Milgram, S., *Psychology Today* **2** (1967) 60.
- [7] Granovetter, M., *The American Journal of Sociology* **78** (1973) 1360.
- [8] Barabási, A.-L. and Albert, R., *Science* **286** (1999) 509.
- [9] Boccaletti, S., Latora, V., Moreno, Y., Chavez, M., and Hwang, D.-U., *Physics Reports* **424** (2006) 175.
- [10] Potterat, J. J. et al., *Sex Transm Infect* **78 Suppl 1** (2002).
- [11] Newman, M. E. J. and Girvan, M., *Physical Review E* **69** (2004) 026113.
- [12] Martinez, N. D., *Ecol.Monogr.* **61** (1991).
- [13] Rapoport, A., *Bulletin of Mathematical Biology* **19** (1957) 257.
- [14] Bernoulli, D. and Blower, S., *Reviews in Medical Virology* **14** (2004) 275.

- [15] Hamer, W., *Lancet* (1906) 733.
- [16] Ross, R., *Proceedings of the Royal Society of London* **92** (1916) 204.
- [17] Soper, M., *Journal Royal Statistical Society* **92** (1929) 34.
- [18] Kermack, W. and McKendrick, A., *Bulletin of Mathematical Biology* **53** (1991) 33.
- [19] Pastor-Satorras, R. and Vespignani, A., *Phys. Rev. Lett.* **86** (2001) 3200.
- [20] Kuperman, M. and Abramson, G., *Phys. Rev. Lett.* **86** (2001) 2909.
- [21] Newman, M. E. J., *Phys. Rev. E* **66** (2002) 016128.
- [22] Jones, J. H. and Handcock, M. S., *Nature* **423** (2003) 605.
- [23] Shirley, M. D. and Rushton, S. P., *Ecological Complexity* **2** (2005) 287.
- [24] Bansal, S., Grenfell, B. T., and Meyers, L. A., *Journal of The Royal Society Interface* **4** (2007) 879.
- [25] Ward, H., *Sex Transm Infect* (2007) sti.2006.023598.
- [26] Newman, M. E. J. and Park, J., *Phys. Rev. E* **68** (2003) 036122.
- [27] Newman, M. E. J., *Phys. Rev. Lett.* **89** (2002) 208701.
- [28] Newman, M. E. J. and Watts, D. J., *Physical Review E* **60** (1999) 7332.
- [29] Bearman, P. S., Moody, J., and Stovel, K., *American Journal of Sociology* **110** (2004) 44.
- [30] Batagelj, V. and Mrvar, A., *Pajek – program for large network analysis*, 1998.
- [31] Liljeros, F., Edling, C. R., Amaral, L. A. N., Stanley, H. E., and Aberg, Y., *Nature* **411** (2001) 907.