

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
PROGRAMA DE PÓS-GRADUAÇÃO EM FÍSICA
Tese de Doutorado

DINÂMICA DE DOENÇAS INFECCIOSAS
EM
REDES COMPLEXAS

Alexsandro Marian Carvalho

Tese realizada sob a orientação do Prof. Dr. Sebastián Gonçalves, apresentada ao Programa de Pós-Graduação em Física da UFRGS em preenchimento parcial dos requisitos para a obtenção do grau de Doutor em Ciências.

Porto Alegre
Novembro de 2012

Agradecimentos

Agradeço inicialmente ao meu orientador prof. Dr. Sebastián Gonçalves, cujos conhecimentos e habilidades não só enriqueceram esta tese 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 por todo o carinho.

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, ou seja, ausência de limiar de epidemia e limiar de epidemia dependente da topologia. Motivados por estes resultados, analisamos e implementamos modelos teóricos e computacionais de propagação de epidemias. Em termos de contribuição, primeiramente, apresentamos uma solução algébrica para o modelo SIR na aproximação de campo médio. Em seguida, desenvolvemos um modelo de rede de interação dinâmica a qual se desenvolve uma epidemia. Entre os resultados, achamos que o tempo característico de formação da rede frente ao tempo da doença influencia na dinâmica da epidemia. Por fim, utilizando da dinâmica da doença, comparamos uma rede real de contatos amorosos com as redes paradigmáticas. Concluimos, que estrutura estática real comporta-se como uma rede aleatória e que a inserção da formação dinâmica das interações atenua os efeitos da epidemia.

Abstract

Phenomena such as the rapid spread of epidemics on a global scale have drawn attention to the importance of the structure of the network, through which members of a community interact. The effect of the network topology of social contacts the dynamics of epidemics has been studied recently in general network models (small world, scale free, etc), showing effects that give another dimension to the results of typical mean field, or absence of epidemic threshold and epidemic threshold depends on the topology. Motivated by these results, we analyze and implement theoretical and computational models for the spread of epidemics. In terms of contribution, first, we present an algebraic solution for the SIR model in the mean field approximation. Further, we develop a network model for the dynamic interaction between individuals in which develops an epidemic. Among the results, we found that the characteristic time of formation of the network against the time of disease influences the dynamics of the epidemic. Finally, using the dynamics of the disease, we compared a real network of loving contacts with the network model. We conclude that real static structure behaves like a random network and that the insertion of the formation dynamics of interactions attenuates the effects of the epidemic.

Sumário

1	Introdução	1
2	Redes Complexas	3
2.1	Breve História	3
2.2	Caracterização da Topologia	6
2.2.1	Grau	6
2.2.2	Distribuição de Grau	6
2.2.3	Coeficiente de Agrupamento	7
2.2.4	Menor Caminho Médio	7
2.3	Modelos de Redes	8
2.3.1	Modelo de Erdős-Rényi	8
2.3.2	Modelo de Watts-Strogatz	11
2.3.3	Modelo de Barabási-Albert	15
3	Abordagens para Doenças Infecciosas	21
3.1	Modelo SIR	21
3.2	Aproximação de Campo Médio	22
3.2.1	Modelo Determinístico	22
3.2.2	Modelo Estocástico	25
3.3	Redes de Contatos	27
3.3.1	Rede Estática	28
3.3.2	Rede Dinâmica	34
4	Uma Solução Algébrica para o Modelo de Kermack-McKendrick	39
4.1	Nova Representação	40
4.2	Aproximação	41
4.3	Solução Algébrica	43

5	Modelo de Rede Dinâmica e Epidemias	46
5.1	Rede Dinâmica	46
5.2	Epidemia	48
5.2.1	Número Básico de Reprodução	48
5.2.2	Probabilidade da Epidemia	51
5.2.3	Dinâmica da Doença	52
6	Dinâmica de Epidemias em uma Rede Social Real	56
6.1	Redes	56
6.2	Simulações Computacionais	58
6.3	Resultados	59
6.3.1	Rede Estática	60
6.3.2	Rede Dinâmica	67
7	Conclusões	72
	Referências Bibliográficas	76

Capítulo 1

Introdução

A necessidade do homem primitivo de sobreviver a um ambiente hostil desencadeou uma apurada técnica de observação. Esta habilidade permitiu-lhe estender o conhecimento dos fenômenos naturais. Com o decorrer dos anos a compreensão humana da natureza tornou-se mais complexa. Dada a incapacidade humana de entender a ordem de complexidade das “leis” que governam a natureza, ele passou a simplificar sua realidade. Esta representação é o que denominamos de modelo. Uma visão não real, contudo às vezes bem estruturada da realidade.

Para que o modelo seja executável devemos subtrair alguns detalhes. O que, por sua vez, implica em um conhecimento dos mecanismos básicos do fenômeno em questão. Particularmente, em epidemiologia, este feito foi dado pelo avanço do conhecimento médico sobre os agentes causadores das infecções, no século XIX (Lechevalier [1]). Com efeito, foi possível sistematizar a dinâmica de transmissão de uma doença infecciosa em termos de modelos matemáticos. Provavelmente, o 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 (Kermack e McKendrick [2]). Praticamente todos os estudos realizados com esta dinâmica, desde o início do século passado, consideram as classes distribuídas homoganeamente 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 (Barabási [3]). A estrutura das interações pode ser representada por certas redes paradigmáticas (*small-world*, *scale-free*, etc) (Newman e Park [4]). A introdução da heterogeneidade espacial levou a uma conclusão distinta acerca do número crítico de infectados necessários para levar a população a uma epidemia (limiar da doença), sugerindo uma redução deste valor ou, em algumas redes *scale-free*, mesmo um valor nulo (Pastor-Satorras e Vespignani [5]). Em síntese, na última

década, percebemos que um modelo de propagação de doenças epidêmicas deve considerar a dinâmica da doença e a estrutura de contatos. No entanto, estes não são os únicos ingredientes relevantes. Recentemente, a inserção de dinâmica nas interações entre os agentes sugeriu uma relação entre os tempos característicos da formação dinâmica dos pares e da propagação da doença. Entre os principais resultados, observou-se que o limiar da doença não se difere do cenário estático desde que a dinâmica de formação da rede seja mais lenta do que o espalhamento da doença (Risau-Gusman [6]).

Neste trabalho implementamos e analisamos modelos teóricos e computacionais que descrevem a dinâmica de transmissão de uma doença dentro de redes complexas. A primeira contribuição está relacionada a uma análise do modelo SIR na aproximação de campo médio. Na literatura, muitos autores fazem menção a impossibilidade tanto de se obter uma solução exata como uma boa aproximação analítica para o modelo SIR padrão. Apresentamos, utilizando basicamente de uma nova representação, uma solução algébrica. Posteriormente, implementamos um modelo de interação dinâmica entre sujeitos e introduzimos a disseminação de uma doença, o que nos permite avaliar a relação entre os tempos característicos da rede e doença na dinâmica da epidemia. Por fim, desenvolvemos um conjunto de simulações computacionais para introduzir a propagação de uma epidemia em uma rede real. Estes estudos fornecem subsídios para um paralelo entre as redes paragnáticas e esta topologia real.

Em termos de estrutura, este trabalho está dividido em sete capítulos, dos quais esta introdução é o primeiro e as conclusões o último. Dos seis capítulos restantes, os dois seguintes apresentam uma revisão bibliográfica enquanto que os demais as nossas contribuições. Na sequência apresentamos uma pequena síntese de cada capítulo. 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 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 e a descrição da propagação de doenças em redes complexas. No capítulo quatro, apresentamos uma formulação algébrica para o modelo SIR na aproximação de campo médio. Avaliamos, no capítulo cinco, os efeitos de inserir dinâmica nos contatos entre os indivíduos tanto na evolução da rede bem como na propagação de uma epidemia. No capítulo seis, analisamos os resultados de simulações computacionais para a propagação de uma doença nas topologias real (estática e dinâmica) e paradigmáticas. Por fim, apresentamos as conclusões e perspectivas para trabalhos futuros.

Capítulo 2

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.

Este capítulo apresenta uma revisão dos conceitos básicos em redes complexas. A primeira seção destina-se à base histórica, na qual indicamos o desenrolar dos principais eventos relacionados a pesquisa em redes. Na segunda seção, veremos algumas propriedades estatísticas cujo exame se faz necessário para adequada caracterização da topologia. A terceira seção por sua vez apresenta as principais estruturas de redes e os modelos mais conhecidos por gera-las. Por fim, vale mencionar, que a nossa contribuição neste capítulo está associada a realização das simulações computacionais e o seu paralelo com as previsões analíticas.

2.1 Breve História

O termo redes complexas refere-se a um grafo que apresenta uma estrutura topográfica não trivial, composto por um conjunto de vértices (nós) que são interligados por meio de arestas (Albert e Barabási [7]). A teoria dos grafos, 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 há uma ilha entre os dois braços de um rio conectada a eles 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,

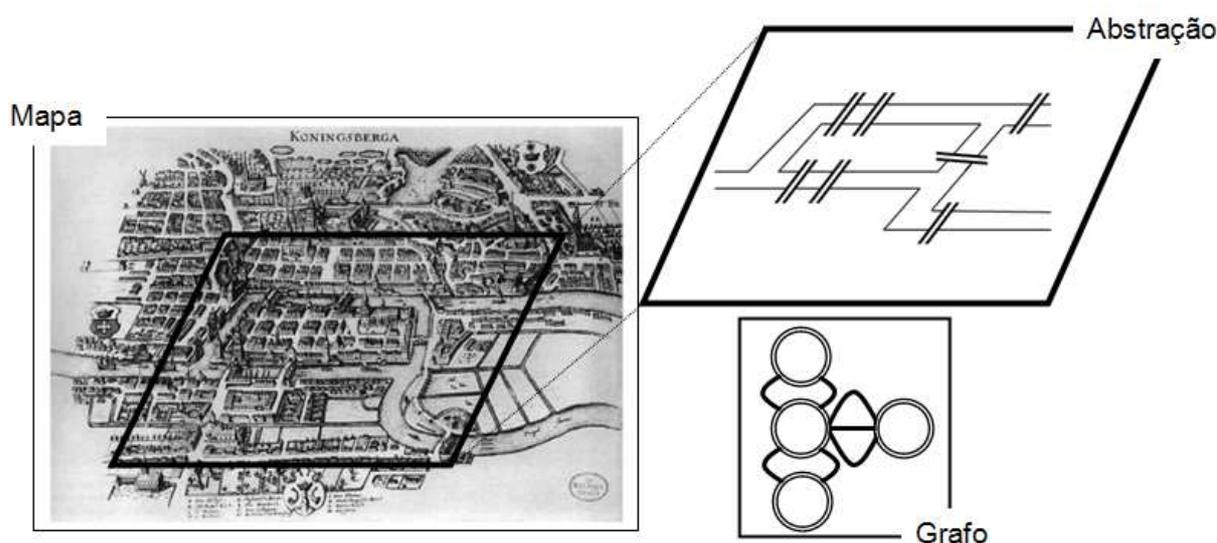


Figura 2.1: As sete pontes de Königsberg. Na linguagem dos grafos tal figura é constituída por quatro vértices e sete arestas.

criando assim o que pode ser o primeiro grafo da história.

Os estudos das redes complexas aplicados a sociologia foram iniciados em meados de 1930, quando sociólogos utilizavam essas redes com a finalidade de estudar o comportamento da sociedade e a relação entre os indivíduos. Contudo, foi só com o trabalho do psicólogo Stanley Milgram [8], nos anos 60, que as propriedades das redes começaram a ser reveladas. Milgram enviou 160 cartas para pessoas escolhidas ao acaso em dois estados norte-americanos, pedindo-lhes que as procurassem fazer chegar a um destinatário alvo, cujo nome, profissão e zona de residência eram dados. Caso não conhecesse, o destinatário deveria enviar a carta para algum amigo que pudesse eventualmente conhecê-lo. Das 160 cartas, 42 chegaram ao destino e nestas o número médio de intermediários foi seis pessoas. Esta experiência parece confirmar a idéia de que a distância, medida em número de ligações de conhecimento direto, ou graus de separação, entre dois elementos típicos de uma rede de ligações sociais é de fato bastante pequena, mesmo em redes com muitos elementos, como a sociedade americana dos anos sessenta. Existem outras experiências, em áreas tão diversas como a matemática ou a comunidade cinematográfica, que dão resultados semelhantes, por exemplo, o jogo de Bacon e o número de Erdös (Buzydowski [?]). Estes experimentos revelaram uma das características das redes sociais já a muito difundida pelo conhecimento popular - **o mundo é pequeno**. O modelo mais simples que permite compreender os efeitos de mundo pequeno são das redes aleatórias (veja subseção 2.3.1). As suas principais propriedades foram estudadas e descritas, em 1959, por Erdös e Rényi [9].

No decorrer dos anos, percebeu-se que as redes aleatórias não são totalmente reais,

em muito casos, devido a falta de estrutura local. Por exemplo, em 1973, Granovetter [10] estudou as relações entre pessoas na sociedade. Ele 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. É evidente, neste tipo de rede social, que a probabilidade de quaisquer dois amigos de um dado indivíduo se conhecerem é bastante superior à probabilidade de duas pessoas escolhidas ao acaso se conhecerem. Na terminologia das redes, esta propriedade mede-se a partir do coeficiente de agrupamento (*clustering*) C , que é a probabilidade de dois primeiros vizinhos do mesmo vértice estarem conectados entre si. Por exemplo, dada a rede de atores (Albert e Barabási [7]) constituída de 225226 indivíduos determina-se o coeficiente de agrupamento $C = 0.79$. Ao passo que uma rede aleatória de mesmo tamanho e grau de separação possui $C = 0.00027$. Como se pode ver, a rede de atores têm um agrupamento muito superior ao da rede aleatória. Esta característica, associado ao maior C , é denominado de **efeito da vizinhança**. Então, para modelar redes reais precisamos de um modelo mais sofisticado do que o das redes aleatórias, que reproduza o efeito de mundo pequeno sem sacrificar o efeito da vizinhança. Este modelo foi introduzido, em 1998, por Watts e Strogatz [11]. Levando em conta este fato, eles não só construíram um modelo, capaz de gerar uma rede de mundo pequeno (*small-world*) (em analogia ao fenômeno descoberto por Milgram), que comporta os dois efeitos, mundo pequeno e vizinhança, bem como formalizaram o problema de Granovetter. O trabalho de Watts e Strogatz traz a estrutura de mundo pequeno para redes sociais da sociologia para as comunidades formadas por físicos, inspirando novas investigações em redes complexas.

Em 1999, Barabási e Albert [12] decidiram verificar se a estrutura de mundo pequeno estava presente na *World Wide Web*. Para isso, mapearam a topologia das conexões entre as páginas da WWW. Os autores revelaram neste estudo que a distribuição de probabilidades para o número de ligações das páginas segue uma lei de potência. Isto significa que, ao contrário do que acontece nas redes aleatórias e nas redes de mundo pequeno, o número de ligações de um vértice não tem um valor típico. As redes com esta propriedade foram batizadas com o nome de redes livre de escala (*scale-free*), precisamente por não terem uma escala característica na conectividade. Inspirados por estas descobertas, eles propuseram um modelo de redes *scale-free* cujo ingrediente básico é a ligação preferencial: princípio de favorecer as ligações a vértices com maior número de ligações. Em alusão a uma parábola do Evangelho de Mateus, “aos que mais têm será dado em abundância e, aos que menos têm, até o que têm lhes será tirado” (Mateus, 25, 28-29), chama-se este **efeito de Mateus**.

Na última década, a teoria das redes complexas tem se desenvolvido expressivamente. Investigações de novas ferramentas (algébricas, computacionais, etc) auxiliam cada vez mais na adequada classificação das propriedades intrínseca de redes reais. Estes esforços, por um

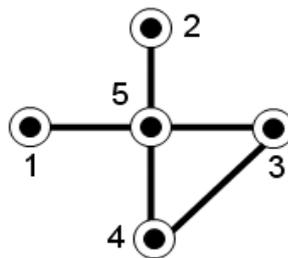


Figura 2.2: Exemplo de rede formada por 5 vértices e 5 arestas. Identificamos o grau de cada vértice como $k_1 = k_2 = 1$, $k_3 = k_4 = 2$ e $k_5 = 4$. A rede possui grau médio $\langle k \rangle = 2$.

lado, reiteram as características de mundo pequeno, vizinhança e Mateus numa diversidade maior de sistemas e, por outro lado, revelam novas propriedades, tais como: comunidade, ciclos, etc.

2.2 Caracterização da Topologia

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 o trabalho de Newman [13].

2.2.1 Grau

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, veja figura 2.2. A partir da medida do grau podemos determinar o grau médio, dado por

$$\langle k \rangle = \frac{1}{N} \sum_{i=1}^N k_i. \quad (2.1)$$

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.

2.2.2 Distribuição de Grau

Para obter uma caracterização global dos graus dos vértices da rede definimos uma quantidade chamada distribuição de grau, $P(k)$. Esta quantidade descreve a probabilidade de um vértice escolhido aleatoriamente ter grau k . Tal medida é importante, por exemplo,

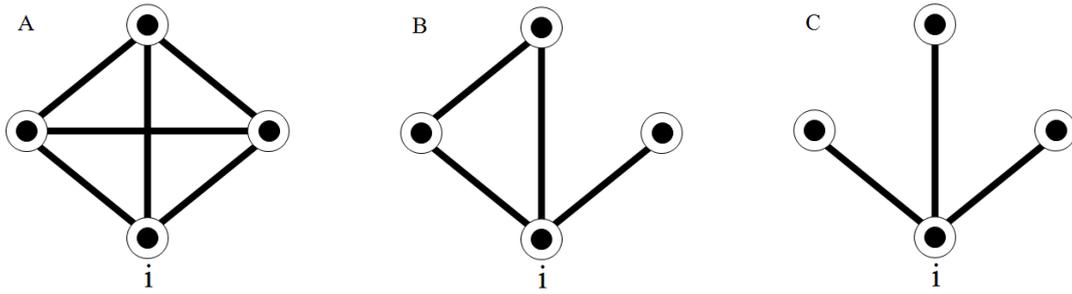


Figura 2.3: Exemplos de redes com diferentes coeficientes de agrupamento. Em (A) todos os vértices estão conectados entre si, $C_i = 1$, (B) $C_i = 1/3$ e em (C) os vizinhos do vértice i não possuem conexão entre eles, $C_i = 0$.

para determinar se as conexões entre os vértices são uniformes, servindo como uma medida básica de classificação.

2.2.3 Coeficiente de Agrupamento

Cada vértice possui uma quantidade associada que mede o nível de conexão entre seus vizinhos. Essa quantidade, chamada de coeficiente de agrupamento, indica a probabilidade de dois vértices vizinhos terem um terceiro vértice vizinho a ambos. Para tanto, medimos 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 a coeficiente de agrupamento, figura 2.3, resulta

$$C_i = \frac{2e_i}{k_i(k_i - 1)}. \quad (2.2)$$

Como medida de caracterização global da rede, é calculada a média da clusterização entre todos os vértices da rede,

$$C = \frac{1}{N} \sum_{i=1}^N C_i. \quad (2.3)$$

2.2.4 Menor Caminho Médio

O comprimento do caminho que conecta dois vértices i e j é dado pelo número de arestas ao longo deste caminho. Definimos d_{ij} como o menor caminho entre todos os caminhos que conectam os vértices i e j . Sua determinação é importante para caracterização da estrutura interna das redes e na investigação de efeitos dinâmicos relativos ao transporte e à comunicação. Os menores caminhos entre todos os pares de vértices em uma rede podem ser representados através de uma matriz de distâncias Ω . O valor $d_{max} = \max d_{ij}$ é chamado

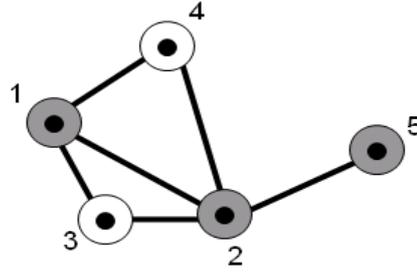


Figura 2.4: Exemplo de caminho na rede. Em cinza destacamos os vértices que pertencem ao menor caminho que separa os vértices 1 e 5.

diâmetro da rede. A média entre os valores na matriz Ω define o caminho característico da rede (menor caminho médio), sendo calculada por

$$L = \frac{1}{N(N-1)} \sum_{i \neq j} d_{ij} \quad (2.4)$$

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 contornar esta divergência é utilizar a chamada eficiência global, definida como

$$E = \frac{1}{N(N-1)} \sum_{i \neq j} \frac{1}{d_{ij}} \quad (2.5)$$

que corresponde a média dos inversos dos menores caminhos.

2.3 Modelos de Redes

Na literatura encontramos uma vasta variedade de modelos de redes, veja Costa et al [14]. Nesta seção apresentaremos os principais modelos e algumas de suas propriedades.

2.3.1 Modelo de Erdős-Rényi

O modelo Erdős-Rényi (ER) [9] gera redes aleatórias. Para tanto, ele assume que uma rede é formada a partir de uma coleção inicial de vértices, e que as conexões atribuídas são realizadas de forma aleatória. Desta maneira, não existe nenhum critério que privilegie umas ligações em relação a outras.

No que segue, apresentaremos a distribuição de grau, menor caminho médio e coeficiente de agrupamento para este modelo.

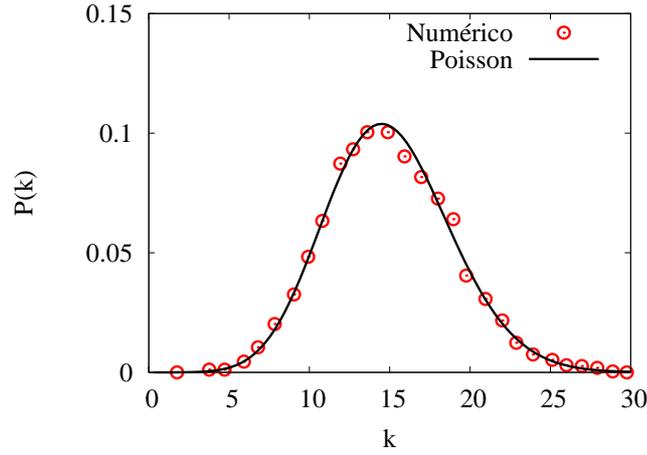


Figura 2.5: Distribuição de grau que resulta da simulação numérica (\odot) de uma rede aleatória. Rede gerada com $N = 10000$ e $p = 0.0015$. Comparamos a simulação com os resultados da distribuição de Poisson (linha contínua), podemos ver que o desvio é pequeno.

Distribuição de Grau

Uma rede aleatória com probabilidade de conexão p , o grau k_i para o vértice i segue uma distribuição binomial com parâmetros $N - 1$ e p ,

$$P(k) = C_{N-1}^k p^k (1-p)^{N-1-k} \quad (2.6)$$

Para $N \rightarrow \infty$, $p \rightarrow 0$ e $Np = cte$, observamos que

$$\begin{aligned} \lim_{N \rightarrow \infty} \left(1 - \frac{\langle k \rangle_t}{N}\right)^{N-k} &\rightarrow e^{-\langle k \rangle_t} \\ \lim_{N \rightarrow \infty} \frac{N!}{(N-k)! N^k} &\rightarrow 1. \end{aligned} \quad (2.7)$$

Como bases nestas aproximações, reescrevemos a equação (2.6) como

$$P(k) \simeq e^{-\langle k \rangle} \frac{\langle k \rangle^k}{k!}. \quad (2.8)$$

A última expressão é denominada de distribuição de Poisson, onde $\langle k \rangle = pN$ é o grau médio.

Na Fig. 2.5 mostramos a distribuição de grau para uma rede aleatória gerada com os parâmetros $N = 10000$ e $p = 0.0015$. Desta figura podemos observar que o pico da distribuição, $P(k) \sim 0.1$, está associado diretamente com o grau médio, $\langle k \rangle = 15$. Adicionalmente, verificamos que vértices com um grau muito superior ou muito inferior ao grau médio são pouco prováveis.

Menor Caminho Médio

Para caracterizar o tamanho de uma rede aleatória calculamos a distância média entre pares de vértices. Para tanto, escolhemos ao acaso um vértice da rede como ponto

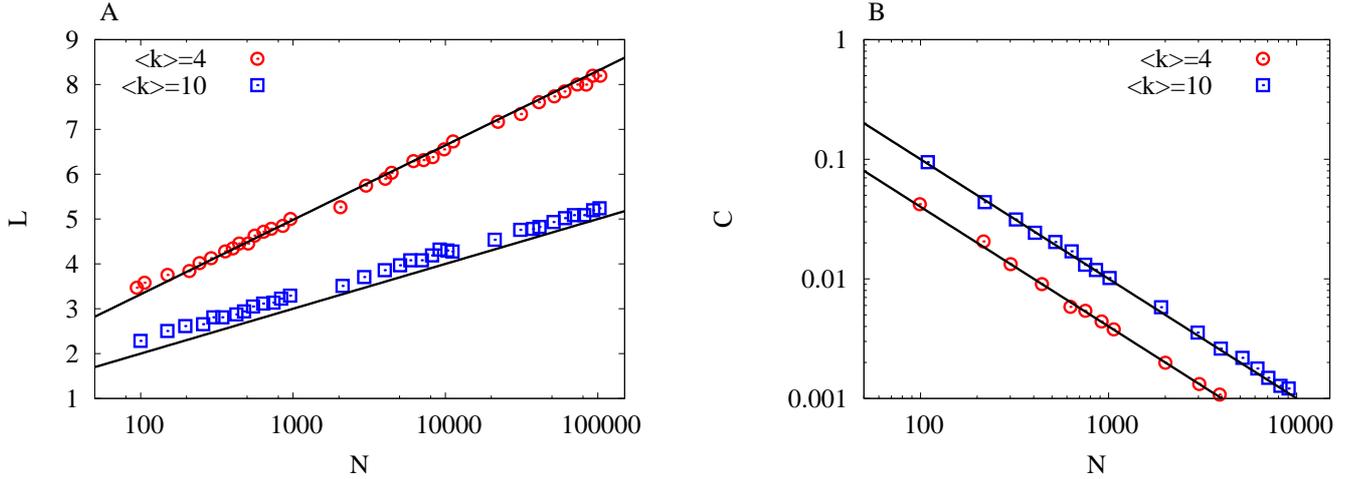


Figura 2.6: (A) O menor caminho médio L versus o tamanho da rede N . (B) O coeficiente agrupamento C versus o tamanho da rede N . Os símbolos referem-se a simulações numéricas e as linhas contínuas as previsões analíticas.

de partida. Este vértice terá em média $\langle k \rangle$ primeiros vizinhos, cada um dos quais também com $\langle k \rangle$ primeiros vizinhos, ou seja, o vértice inicial tem aproximadamente $\langle k \rangle^2$ segundos vizinhos. Seguindo o mesmo raciocínio, o vértice inicial terá cerca de $\langle k \rangle^n$ n-ésimos vizinhos, isto é, vértices a uma distância de n graus de separação. Quando tivermos $n = L$, que resulta em $\langle k \rangle^L \sim N$, a cadeia construída a partir do vértice inicial pode chegar a qualquer vértice da rede. Portanto, o menor caminho médio numa rede aleatória é

$$L \sim \frac{\ln N}{\ln \langle k \rangle}. \quad (2.9)$$

Na Fig.2.6A mostramos o comportamento do menor caminho médio frente ao tamanho da rede N e o grau médio $\langle k \rangle$. Em relação ao tamanho da rede, podemos observar que ao elevar o valor de N aumentamos o valor de L , em verdade, em uma relação logarítmica. Sendo assim, o menor caminho médio cresce lentamente com o tamanho da população. No entanto, esta mesma figura mostra que o acréscimo do $\langle k \rangle$ reduz L . A compreensão deste resultado é imediata, elevar o grau médio implica em criar um maior número de “caminhos” de conexão entre os vértices. Consequentemente, esta variabilidade nas conexões entres os vértices gera atalhos.

Coeficiente de Agrupamento

Consideramos um vértice na rede aleatória e seus primeiros vizinhos. A probabilidade de dois destes vizinhos estarem conectados é igual a probabilidade de dois vértices escolhidos ao acaso estarem conectados. Consequentemente, o coeficiente de agrupamento é

$$C = p = \frac{\langle k \rangle}{N}. \quad (2.10)$$

Na Fig. 2.6B apresentamos o coeficiente de agrupamento em função do tamanho da população para duas redes aleatórias de grau médio distintos. Da figura observamos que o coeficiente de agrupamento é diretamente proporcional ao grau médio. Esta proporcionalidade resulta, em parte, da redução da largura da distribuição de grau. Em outras palavras, em um sistema muito grande, podemos supor que cada indivíduo tem um grau próximo do grau médio. O que eleva consideravelmente a possibilidade de quaisquer vértices da vizinhança de um dado vértice compartilharem uma conexão. Contrário é o efeito quando fixamos o grau médio e variamos o tamanho da população, C é inversamente proporcional a N . A exemplo da dependência com o grau médio, o entendimento deste resultado passa também pela largura da distribuição de grau.

2.3.2 Modelo de Watts-Strogatz

O modelo de Watts-Strogatz (WS) [11] é útil para a construção de algumas redes de mundo pequeno. Genericamente, estas redes possuem um pequeno valor para o menor caminho médio e um alto valor para o coeficiente de agrupamento.

O modelo, em questão, consiste no seguinte algoritmo:

1. 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 $N \gg \kappa \gg \log N \gg 1$.
2. Randomizar: Aleatoriamente reescreva cada aresta da rede regular com probabilidade

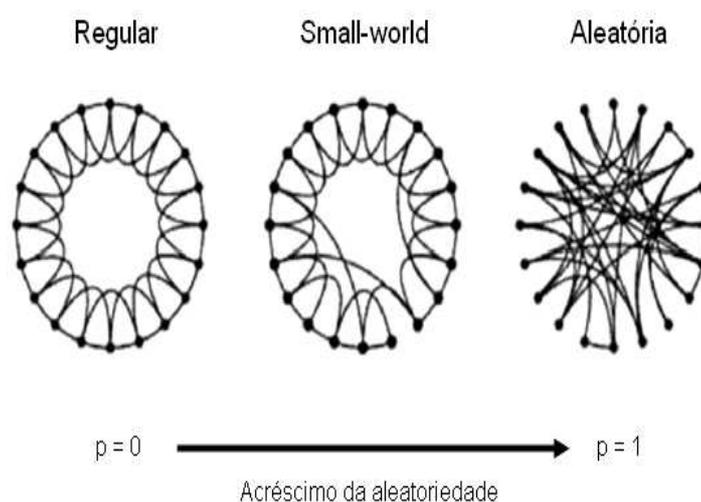


Figura 2.7: O processo aleatório de re-conexão das arestas no modelo de Watts e Strogatz, que interpola entre um anel regular e uma rede aleatória sem alterar o número de vértices e arestas. Iniciamos com $N = 20$ vértices cada um conectado aos seus primeiros quatro vizinhos. Para $p = 0$ o anel original 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 [11]).

p de modo que auto-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).

Variando p transitamos entre redes com topologias distintas, este processo é ilustrado na figura 2.7. Para baixos valores de p não alteramos a configuração inicial, ou seja, a rede é altamente estruturada localmente e possui um número elevado de graus de separação, caracterizando uma rede regular. No outro extremo, grandes valores de p , desordenamos a estrutura local (via as muitas ligações aleatórias) e, por conseguinte, reduzimos consideravelmente o grau de separação da rede. Associamos esta completa desorganização as redes aleatórias. Entre estes dois extremos, encontramos uma rede que, por um lado, apresenta uma estrutura local e, por outro, um baixo grau de separação, denominada de mundo pequeno. O limite entre as diferentes topologias não é bem estabelecido. Porém, na literatura encontramos sugestão que o regime de mundo pequeno está no intervalo $0.01 < p < 0.1$. A esquerda deste intervalo temos a região característica das redes regulares e a direita o regime das redes aleatórias.

Apresentamos na sequência os principais resultados para este modelo.

Distribuição de Grau

Para $p = 0$ cada vértice tem o mesmo grau κ . Então a distribuição de grau é uma função delta centrada em κ . Para $p > 0$, o grau k_i de um vértice i pode ser escrito (Barrat e Weigt [15]) $k_i = \kappa/2 + c_i$, onde c_i pode ser dividido em duas tipos: $c_i^1 \leq \kappa/2$ arestas não removidas (com probabilidade $1 - p$), enquanto $c_i^2 = c_i - c_i^1$ arestas têm sido religadas em i , cada uma com probabilidade $1/N$. As distribuições de probabilidade de c_i^1 e c_i^2 são

$$P_1(c_i^1) = C_{\kappa/2}^{c_i^1} (1-p)^{c_i^1} p^{\kappa-c_i^1} \quad (2.11)$$

$$\begin{aligned} P_2(c_i^2) &= C_{pN\kappa/2}^{c_i^2} \left(\frac{1}{N}\right)^{c_i^2} \left(1 - \frac{1}{N}\right)^{pN\kappa/2 - c_i^2} \\ &\simeq \frac{(p\kappa/2)^{c_i^2}}{c_i^2!} e^{-p\kappa/2} \end{aligned} \quad (2.12)$$

para grande valores de N . Combinando estes dois comportamentos, temos a distribuição de grau

$$P(k) = \sum_{n=0}^{\min(k-\kappa/2, \kappa/2)} C_{\kappa/2}^n (1-p)^n p^{\kappa/2-n} \frac{(p\kappa/2)^{k-\kappa/2-n}}{(k-\kappa/2-n)!} e^{-p\kappa/2} \quad (2.13)$$

para $k \geq \kappa/2$.

Na Fig. 2.8, podemos notar que a forma da distribuição é similar a de uma rede aleatória. Neste sentido, a topologia da rede é relativamente homogênea, todos os vértices

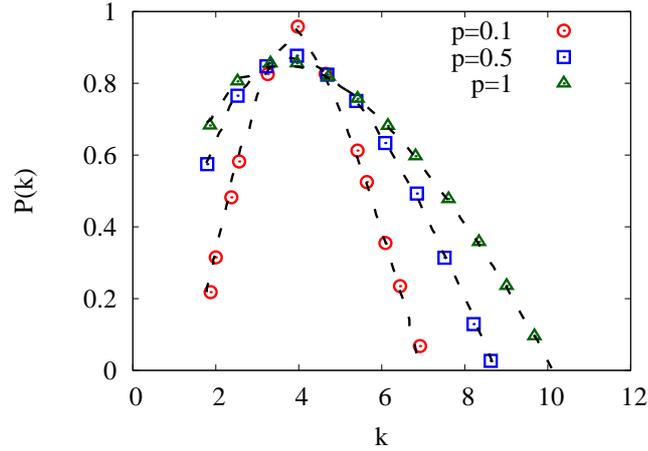


Figura 2.8: Distribuição de grau do modelo WS para $\kappa = 4$ e diversos p . Os símbolos são obtidos de simulações numéricas do modelo WS para $N = 1000$ e as linhas pontilhadas corresponde a equação (2.13).

tem aproximadamente o mesmo número de arestas. Além disso, o efeito do parâmetro de desordem p sobre a distribuição está claro na figura; aumentando a desordem elevamos a largura da distribuição. Em outras palavras, variando p , transitamos de uma distribuição delta (rede regular) a uma distribuição de Poisson (rede aleatória).

Menor Caminho Médio

Consideramos um vizinhança de raio r centrada em torno de um ponto aleatoriamente escolhendo numa rede *small-world* de N vértices. Seja $m(r)$ o número de vértices na rede que não pertence a vizinhança e $n(r)$ sendo o número de *clusters* na rede. Segundo Newman et al [16], a taxa em que m decresce com acréscimo de r é igual ao número $2n$ de arestas dos *clusters* na rede vezes o número de ligações k no anel

$$\frac{d\mu}{dr} = -2\kappa v, \tag{2.14}$$

onde $\mu = m/N$ e $v = n/N$. A taxa em que o número de *clusters* n muda tem duas contribuições. Primeiro, o número de *clusters* aumenta como um resultado dos *shortcuts* (atalhos). Isto significa que como r aumenta, novos *shortcuts* são encontrados com uma taxa $4\kappa n/\xi$, onde $\xi = 1/(\kappa p)$ é o comprimento característico. Para cada *shortcut* encontrado, um novo *cluster* se formará em uma posição aleatória da rede. A probabilidade deste acontecimento é m/N . Então a taxa em que *clusters* são criado é $4\kappa mn/\xi N$. O número de *clusters* decresce quando as arestas de um *clusters* encontram outro. A probabilidade deste evento acontecer é $2\kappa(n-1)n/m$, no limite de um pequeno δr . Desta maneira,

$$\frac{dv}{dr} = \frac{4\kappa\mu v}{\xi} - \frac{2\kappa v(v-1/N)}{\mu}. \tag{2.15}$$

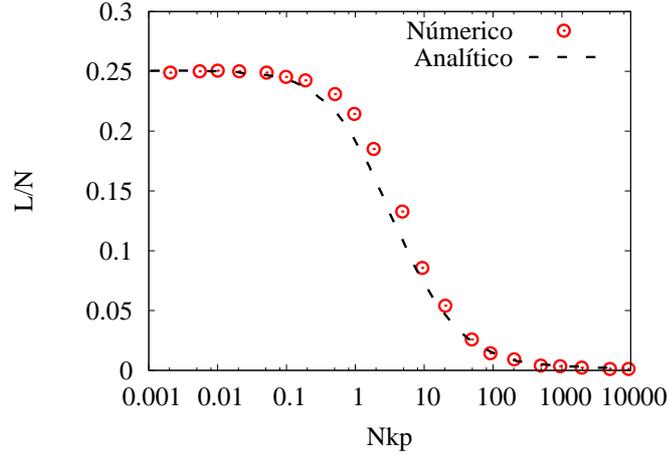


Figura 2.9: O menor caminho médio como uma fração do tamanho do sistema. O símbolo é uma medida numérica para o modelo discreto e a linha pontilhada a solução analítica para o modelo contínuo, equação (2.19) e $\kappa = 2$. Tamanho do sistema $N = 10^6$.

Resulta das equações (2.14) e (2.15),

$$\frac{dv}{d\mu} = -\frac{2\mu}{\xi} + \frac{v - 1/N}{\mu}. \quad (2.16)$$

A solução desta equação, para as condições iniciais $\mu(0) = 1$ e $v(0) = 1/N$, é

$$v = \frac{2}{\xi}(\mu - \mu^2) + \frac{1}{N}. \quad (2.17)$$

Após a substituição deste resultado na equação (2.14), obtemos a seguinte solução

$$r = \frac{\xi}{2\kappa\sqrt{1 + 2\xi/N}} \left(\tanh^{-1} \frac{1}{\sqrt{1 + 2\xi/N}} - \tanh^{-1} \frac{2\mu - 1}{\sqrt{1 + 2\xi/N}} \right). \quad (2.18)$$

Portanto, o menor caminho médio é dado por

$$L = \frac{1}{N} \int_0^\infty r A(r) dr = \frac{\xi}{2k\sqrt{1 + 2\xi/N}} \tanh^{-1} \frac{1}{\sqrt{1 + 2\xi/N}}, \quad (2.19)$$

onde $A(r) = 2N\kappa v(r)$ é a “área da superfície” da vizinhança.

Na Fig. 2.9 comparamos as soluções discreta (simulação) e contínua (equação (2.19)). Ainda em relação a esta figura, destacamos os fatores de escala do menor caminho médio L (N) e da desordem p ($1/N\kappa$). Assim, em mais detalhes, podemos verificar que para um anel regular o menor caminho médio escala linearmente com o tamanho do sistema, $L(0) = N/2 \langle k \rangle$. No caso limite de $p \rightarrow 1$, a rede converge para uma rede aleatória com $L(1) = \ln N / \ln \langle k \rangle$. Porém, na região intermediária $0 < p < 1$ o menor caminho médio cai rapidamente com o incremento de p .

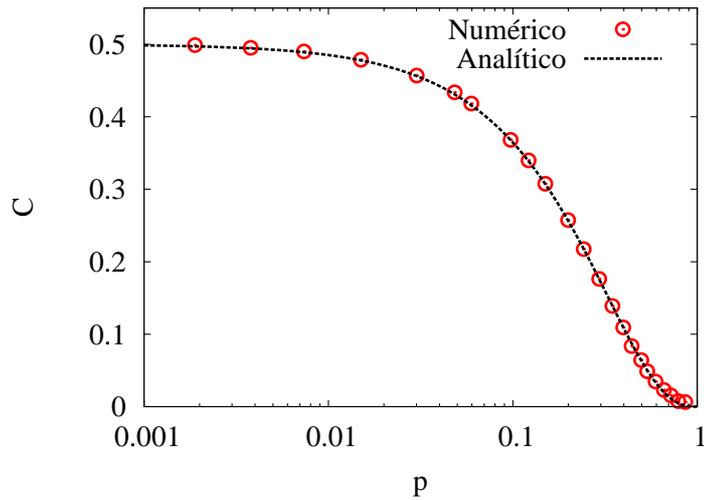


Figura 2.10: Clusterização versus p , para $\kappa = 4$ e $N = 5000$. O símbolo refere-se a simulação numérica e a linha contínua ao resultado analítico, $C(0)(1 - p)^3$.

Coeficiente de Agrupamento

De acordo com Barrat e Weigt [15], para calcular C para o modelo SW, começamos com um anel com coeficiente de agrupamento $C(0)$. Para $p > 0$, dois vizinhos de um vértice i que estava conectado em $p = 0$ são ainda vizinhos e conectados por uma aresta com probabilidade $(1 - p)^3$, desde que existam três arestas que permanecem intactas. Consequentemente, $C = C(0)(1 - p)^3$, para $N \gg \kappa$. Podemos afirmar, então que

$$C = \frac{3(\kappa - 2)}{4(\kappa - 1)}(1 - p)^3. \quad (2.20)$$

Na Fig. 2.10 podemos observar os principais resultados em relação ao coeficiente de agrupamento. Primeiro, a figura apresenta que para pequenos valores de desordem, $p \sim 0$, C é constante, $C(0) = 3(\kappa - 2)/4(\kappa - 1)$. Nesta condição, em verdade, o coeficiente de agrupamento independe tanto de p como do tamanho do sistema. Já no caso limite $p \rightarrow 1$, C tende ao comportamento das redes aleatórias, isto é, inversamente proporcional ao tamanho do sistema, $C(1) = \langle k \rangle / N$. Na região intermediária o coeficiente de agrupamento permanece bastante próximo do seu valor para o anel regular, e só cai para um valor relativamente alto de p . Este resultado na região em que o menor caminho médio cai rapidamente mas o coeficiente de agrupamento não, refere-se ao regime de mundo pequeno.

2.3.3 Modelo de Barabási-Albert

O modelo de Barabási-Albert (BA) [12] gera redes *scale-free* com base a dois processos:

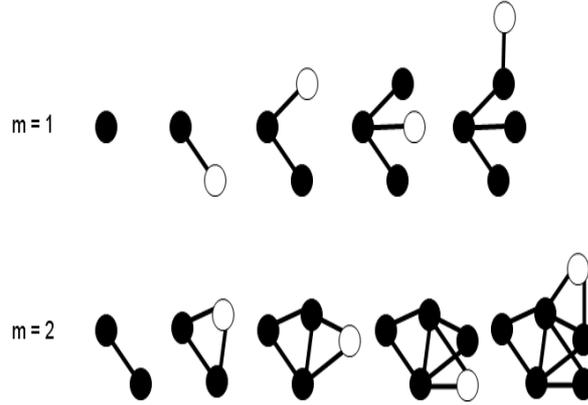


Figura 2.11: Dinâmica de anexação preferencial do modelo BA com parâmetros $m = N_0 = 1$ e $m = N_0 = 2$. O vértice branco representa o novo vértice adicionado em cada passo de tempo.

1. 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.
2. Ligação preferencial: O novo vértice, que vai ser adicionado à rede, tende a se conectar com os vértices mais conectados, ou seja, a probabilidade Π de um novo nó se conectar ao vértice i depende do grau k_i do vértice i , de modo que

$$\Pi(k_i) = \frac{k_i}{\sum_j k_j}. \quad (2.21)$$

Depois de t passos de tempo este processo resulta numa rede com $N = t + N_0$ vértices e mt arestas, figura 2.11. Interessante notar que estes dois mecanismos de construção não estão presentes no modelo ER e no modelo WS. Nestes não há adição de novos vértices e as conexões são estabelecidas de forma homogênea. Além disso, as redes *scale-free* serão 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.

No que segue, analisaremos as propriedades características do modelo BA.

Distribuição de Grau

Assumindo que k_i é uma variável contínua real, a taxa com que k_i muda é proporcional a $\Pi(k_i)$. Consequentemente, k_i satisfaz a equação dinâmica (Barabási e Albert [12])

$$\frac{\partial k_i}{\partial t} = m\Pi(k_i) = m \frac{k_i}{\sum_{j=1}^{N-1} k_j} = \frac{k_i}{2t}. \quad (2.22)$$

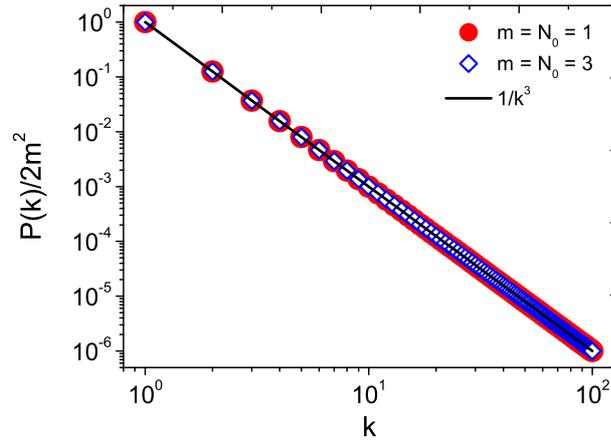


Figura 2.12: Distribuição de grau para o modelo BA. Os pontos referem-se a simulações numérica realizadas $N = 10^5$. A linha contínua representa o resultado de acordo com as previsões analíticas, veja equação (2.27).

A solução desta equação, com a condição inicial $k_i(t_i) = m$ é

$$k_i(t) = m \left(\frac{t}{t_i} \right)^{1/2}. \quad (2.23)$$

A equação acima indica que o grau de todos os nós é descrito por uma lei de potência. Usando esta equação, podemos escrevermos a probabilidade $P[k_i(t) < k]$ como

$$P[k_i(t) < k] = P\left(t_i > \frac{m^2 t}{k^2}\right). \quad (2.24)$$

Considerando que adicionamos novos vértices em intervalos de tempos iguais, o valor t_i tem uma probabilidade constante

$$P(t_i) = \frac{1}{N_0 + t}. \quad (2.25)$$

Substituindo na equação (2.24) e após algumas operações, obtemos

$$P(k) = \frac{\partial P[k_i(t) < k]}{\partial k} = \frac{2m^2 t}{N_0 + t} \frac{1}{k^3} \quad (2.26)$$

tomando o limite $t \rightarrow \infty$

$$P(k) \sim \frac{2m^2}{k^3}. \quad (2.27)$$

Na Fig. 2.12 percebemos que esta abordagem analítica está em acordo com os resultados numéricos. Desta maneira, podemos observar que tanto $P(k) \propto m^2$ como $\ln P(k) \propto -3 \ln k$. A última observação justifica a reta decrescente na figura.

Menor Caminho Médio

Dado um vértice escolhido aleatoriamente, então o número de primeiros vizinhos (número de vértices que estão diretamente ligado a este vértice) segue uma lei de potência

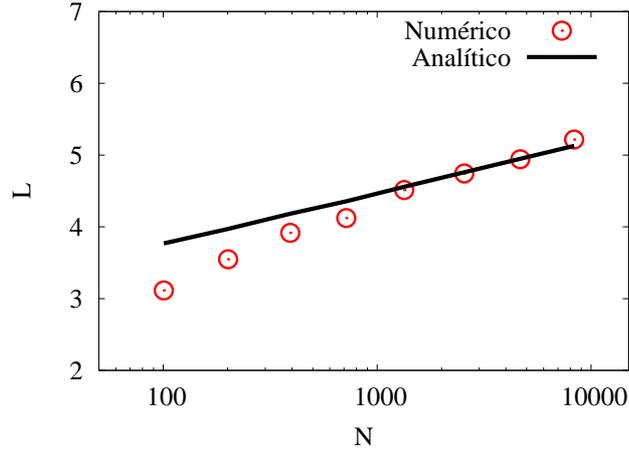


Figura 2.13: Menor caminho médio L versus o tamanho da rede N . Comparamos os resultados analíticos, equação (2.34), com o numérico, simulação. Redes geradas com $m = N_0 = 2$.

conforme descrevemos na subseção anterior. Porém, qual é a distribuição dos segundos vizinhos? Chen et al [17] afirmam que o número de segundos vizinhos é igual a soma do grau de todos os primeiros vizinhos. Sendo assim, a distribuição de grau de cada um destes primeiros vizinhos é dada por

$$q(k) = \frac{kP(k)}{\int kP(k)} = \frac{mk^{-2}}{1 - \frac{1}{\sqrt{N}}}. \quad (2.28)$$

Então, o valor esperado do grau de cada um dos primeiros vizinhos é

$$\langle k_q \rangle = \frac{m \ln N}{2(1 - \frac{1}{\sqrt{N}})}. \quad (2.29)$$

Usando o teorema de Wald [18], o valor esperado de segundos vizinhos é obtido

$$\langle N_2 \rangle = \langle N_1 \rangle \langle k_q - 1 \rangle \quad (2.30)$$

onde N_1 e N_2 denotam, respectivamente, o número de primeiros e segundos vizinhos. Sendo $\langle N_1 \rangle = \langle k \rangle$, então

$$\langle N_2 \rangle = \langle k \rangle \langle k_q - 1 \rangle. \quad (2.31)$$

Podemos generalizar a equação (2.31) para o n -ésimo vizinho como

$$\langle N_n \rangle = \langle k \rangle \langle k_q - 1 \rangle^{n-1}. \quad (2.32)$$

Sendo o número de todos os vizinhos igual a menor caminho médio L_{sf} , temos

$$1 + \sum_{i=1}^{L_{sf}} \langle N_i \rangle = 1 + \sum_{i=1}^L \langle k \rangle \langle k_q - 1 \rangle^{i-1} = N \quad (2.33)$$

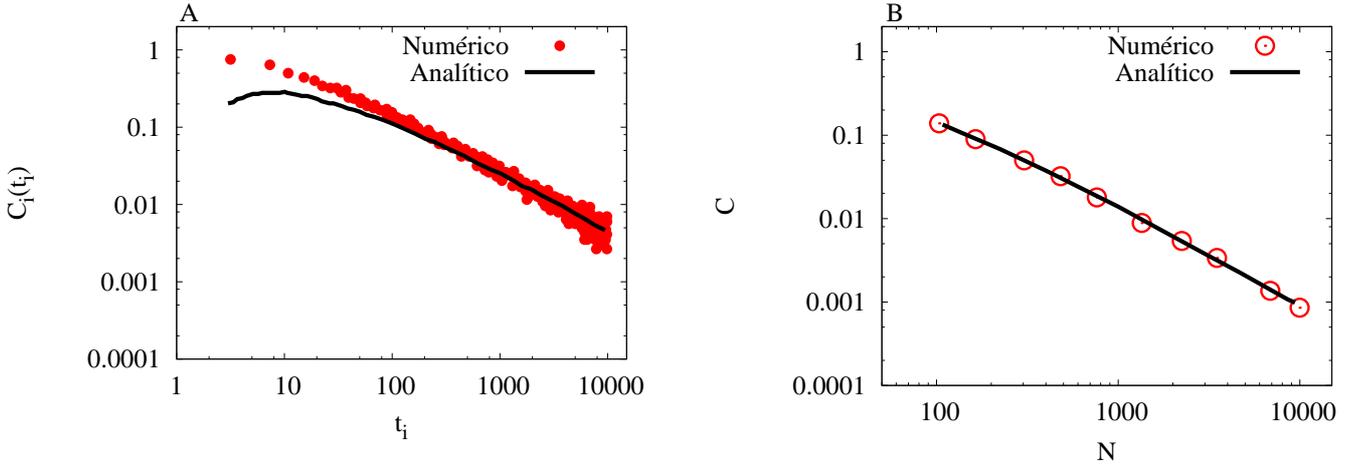


Figura 2.14: (A) Valor inicial do coeficiente de agrupamento local $C_i(t_i)$. (B) O coeficiente de agrupamento C como função do tamanho da rede. O termos numérico e analítico referem-se, respectivamente, a simulação e a equação (2.40). Redes realizadas para $m = N_0 = 2$ (resultados correspondes a uma média sobre 100 redes).

Resolvendo a equação (2.33) e tomando $N \gg 1$, obtemos

$$L \sim \frac{\ln N}{\ln \ln N}. \quad (2.34)$$

A Fig. 2.13 mostra o menor caminho como função do tamanho da rede. Comparando a predição analítica (linha contínua) com os resultados das simulações (pontos) observamos que a curva analítica superestima a solução numérica no limite de pequenas redes, $N \sim 100$. No entanto, no outro limite, $N \sim 10000$, notamos uma boa correspondência entre as duas soluções.

Coeficiente de Agrupamento

Assumimos que $m \geq 2$. O caso $m = 1$ é trivial, as redes são árvores e, portanto, $C = 0$. Por definição, o coeficiente C_i depende de duas variáveis e_i e k_i , veja equação (2.2). A equação apropriada para a mudança de C_i , segundo Fronczak et al [19], é

$$\frac{dC_i}{dt} = \sum_{x=0}^{m-1} p_{ix} \Delta C_{ix}. \quad (2.35)$$

O primeiro termo do lado direito corresponde a probabilidade de um novo vértice de m arestas estar conectado a i e a x primeiros vizinhos de i , dada por

$$p_{ix} = \frac{k_i}{2N} C_x^{m-1} \left(\frac{k_i \ln N}{4N} \right)^x \left(1 - \frac{k_i \ln N}{4N} \right)^{m-1-x}. \quad (2.36)$$

Já o segundo termo denota a mudança no coeficiente de agrupamento quando um novo vértice é conectado ao vértice i e no x primeiros vizinhos de i , isto é,

$$\Delta C_{ix} = -\frac{2C_i}{k_i + 1} + \frac{2x}{k_i(k_i + 1)}. \quad (2.37)$$

Com efeito, a equação (2.35) pode ser reescrita (após um pouco de álgebra)

$$\frac{dC_i}{dt} = -\frac{m}{mt + \sqrt{tt_i}}C_i + \frac{m(m-1)\ln t}{4(mt^2 + t\sqrt{tt_i})}. \quad (2.38)$$

Resolvendo a equação sob a condição inicial $C_i(t_i) = m^2(\ln t_i)^2/8(m-1)t_i$ que descreve o coeficiente de agrupamento do vértice i exatamente no momento de adição t_i , obtemos

$$C_i(t) = \frac{m-1}{8(\sqrt{t} + \sqrt{t_i/m})^2} \left[(\ln t)^2 + \frac{4m}{(m-1)^2} (\ln t_i)^2 \right] \quad (2.39)$$

Para obter o coeficiente de agrupamento médio C da rede, calculamos $C = \int_1^t C_i(t) dt_i/t$. Resulta desta operação, em $t = N$,

$$C = \frac{m^2(m+1)^2}{4(m-1)} \left[\ln \left(\frac{m+1}{m} \right) - \frac{1}{m+1} \right] \frac{(\ln N)^2}{N}. \quad (2.40)$$

Na Fig. 2.14 confrontamos a predição dos resultados das duas últimas equações frente as simulações. A partir da Fig. 2.14A notamos que nosso tratamento analítico para o coeficiente de agrupamento local C_i não condiz com a solução numérica para redes de tamanho pequeno, $N < 100$. Consequentemente, podemos afirmar que a solução analítica para C também não será adequada para este tamanho de rede. Porém, verificamos uma correspondência entre as duas soluções para ambas as quantidades, C_i (Fig 2.14A) e C (Fig 2.14B), para redes de tamanho no intervalo $N > 100$.

Capítulo 3

Abordagens para Doenças Infecciosas

As epidemias não são oriundas do contexto contemporâneo. Dentro da história é possível observar 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 da propagação de uma epidemia 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 (revisão) destes modelos. Entre as possíveis dinâmicas de uma doença, enfatizamos um modelo que divide a população nas distintas classes suscetível, infectado e removível, bastante difundido na literatura, a qual abordamos na primeira seção. Seguimos para a segunda seção, na qual apresentaremos o modelo SIR na aproximação de campo médio. Na terceira seção, indicaremos os principais resultados dos estudos teóricos de epidemias em redes complexas. Por fim, vale mencionar, que a nossa contribuição neste capítulo está associada a realização das simulações computacionais e o seu paralelo com as previsões analíticas.

3.1 Modelo SIR

O modelo SIR é um modelo utilizado para descreve a evolução de uma epidemia numa população dividida em três classes: 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 imunes), ou seja, daqueles que já foram infectados pela doença e tornaram-se imunes. Assim, em um instante de tempo t , uma população é caracterizada

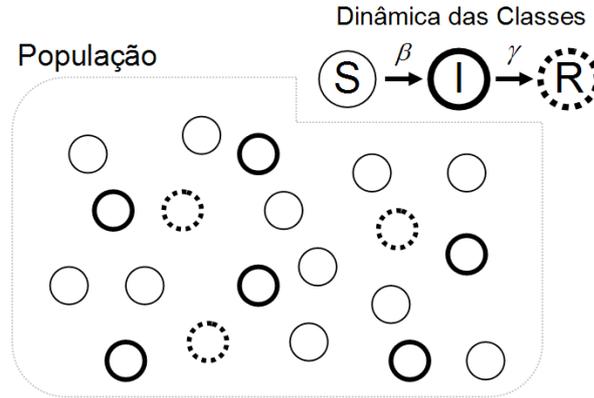


Figura 3.1: Esquema de uma população isolada, classificada de acordo com o modelo SIR; note que as classes estão misturadas aleatoriamente. No topo direito, temos a representação da dinâmica das classes do modelo SIR.

pelos quantidades $S(t)$, $I(t)$ e $R(t)$ sendo que cada componente representa o número de indivíduo na particular categoria, onde $N(t) = S(t) + I(t) + R(t)$ é o número total de indivíduos da população a cada instante (tomado como constante no que segue). A dinâmica entre as classes procede de tal forma que um indivíduo saudável, S , torna-se infectado com uma dada probabilidade β após entrar em contato com um indivíduo que pertença a classe dos infectados, I , até que transcorra um tempo τ_d , no qual o indivíduo em questão torna-se imune, R . Podemos sintetizar esta dinâmica da doença, pelas equações

$$\begin{aligned} S + I &\xrightarrow{\beta} II \\ I &\xrightarrow{\gamma} R \end{aligned} \tag{3.1}$$

onde $\gamma = 1/\tau_d$.

3.2 Aproximação de Campo Médio

Como uma primeira abordagem, supomos que 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 assumimos que os indivíduos das diferentes classes estão homogeneamente misturados, de maneira que qualquer indivíduo de uma classe pode estar em contato com qualquer outro indivíduo de outra classe em qualquer instante, veja figura 3.1. Denominamos esta abordagem de aproximação de campo médio.

3.2.1 Modelo Determinístico

Nosso objetivo aqui é apresentar o modelo determinístico que descreve a dinâmica da epidemia para o modelo SIR na aproximação de campo médio. De acordo como Kermack e

McKendrick [2], para tanto devemos realizar duas hipóteses básicas para definir as transições entre os estados:

1. O número de agentes na classe de infectados aumenta a uma taxa proporcional ao número de agentes na citada classe e ao número de agentes na classe suscetível. Matematicamente, isso é representado pelo termo $\beta S(t)I(t)$. Esta hipótese é herdada do estudo de cinética química, ao 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 $\gamma I(t)$.

Esses mecanismos são descritos pelas equações

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

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

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

dadas as condições iniciais $S(0) = N - i_0$, $I(0) = i_0$ e $R(0) = 0$. Em algumas versões do modelo, o parâmetro β/N é substituído por β . A formulação dada aqui é mais apropriada pois um indivíduo faz contatos potencialmente infecciosos em uma taxa β independente do tamanho da população. A Fig. 3.2A representa a solução numérica (método de Euler) para a dinâmica da epidemia. Desta figura, podemos ver, entre outras coisas, a presença de um estado assintótico

$$(S(\infty), I(\infty), R(\infty)) = (N - R(\infty), 0, R(\infty)),$$

onde a primeira coordenada deve-se ao fato da população ser constante. Além disso, podemos notar que a curva dos infectados apresenta um pico, o qual denominaremos de I_{max} .

Em relação a $R(\infty)$, perceba que um valor não nulo determina a condição para que uma epidemia se estabeleça. Em termos da equação (3.3), representamos tal condição por $dI/dt > 0$. Sendo assim, desta equação, temos que a condição necessária é $(\beta S/N - \gamma)I > 0$, com $I > 0$. Sempre essa relação é satisfeita quando $S > \gamma N/\beta$. Então, para dados β e γ , existe um número crítico de suscetíveis para uma epidemia, e se $S(0) = S_0$ é o número inicial de suscetíveis, a condição para uma epidemia resulta ser $S_0 > \gamma N/\beta$. A interpretação deste resultado 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

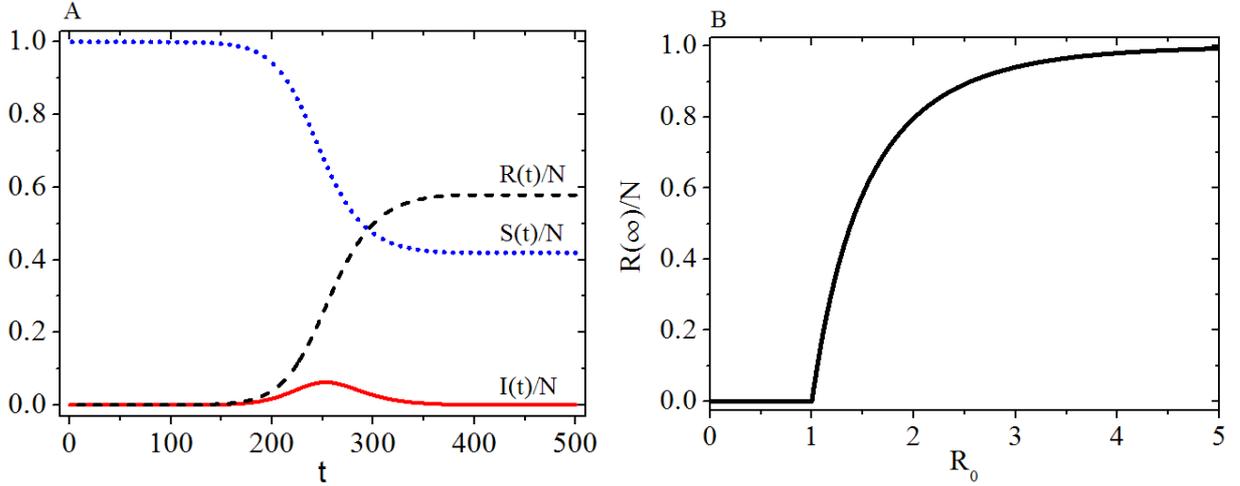


Figura 3.2: (A) Dinâmica das classes do modelo SIR para os parâmetros $\beta = 0.15$ e $\gamma = 0.1$. (B) Fração da população infectada, $R(\infty)/N$, durante a epidemia em função do número básico de reprodução, R_0 . Os resultados correspondem a uma população de $N = 10^6$ sobre a condição inicial de um único indivíduo infectado.

produzidos por um infectado é $\beta S/(N\gamma)$. Assim, para que um infectado introduzido em uma população de suscetíveis resulte em uma epidemia, este número deve ser superior a 1. Introduzindo o número básico de reprodução, parâmetro adimensional

$$R_0 = \frac{\beta}{\gamma}, \quad (3.5)$$

para $S_0 \simeq N$. Observe que a condição necessária para a persistência da epidemia se expressa como $R_0 > 1$ e, portanto, o limiar da epidemia com $R_0 = 1$.

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 (3.5), obtemos

$$\frac{dS}{dR} = -\frac{R_0}{S_0} S \quad (3.6)$$

Assumindo que o tamanho inicial da infecção é pequeno, $i_0 = 1$, e um tempo suficientemente grande para que $I = 0$ (estado assintótico onde $R(\infty) + S(\infty) = N$), a solução de (3.6) é dada por

$$1 - r = e^{-R_0 r}, \quad (3.7)$$

onde $r = R(\infty)/N$. Trata-se de uma equação implícita, cuja curva, mostrada na figura 3.2B, ilustra o efeito de se aumentar R_0 , além do limiar de epidemia $R_0 = 1$. Observamos, claramente, nesta figura duas regiões; uma sem epidemia ($R_0 \leq 1 \rightarrow R(\infty) = 0$) e outra com epidemia ($R_0 > 1 \rightarrow R(\infty) > 0$).

Outro resultado analítico útil que podemos derivar é o pico de infecção, I_{max} . Das equações (3.2) e (3.3), obtemos

$$\frac{dI}{dS} = -1 + \frac{1}{R_0} \frac{1}{S}. \quad (3.8)$$

Integrando a última equação, resulta

$$\Gamma = I + S - \frac{1}{R_0} \ln S = \text{constante} = N - \frac{1}{R_0} \ln(N - i_0) \quad (3.9)$$

onde a última igualdade deve-se as condições iniciais (note que $d\Gamma/dt = 0$). Ainda da equação (3.3), vemos que o máximo ocorre em $S = 1/R_0$ onde $dI/dt = 0$. Assim, da equação (3.9), com $S = 1/R_0$ obtemos

$$I_{max} = N - \frac{1}{R_0} (\ln(R_0(N - i_0)) + 1). \quad (3.10)$$

Com efeito, percebemos que para qualquer valor inicial $i_0 > 0$ (com $R_0 > 1$), I aumenta partindo de i_0 e cresce até uma valor máximo I_{max} . Após atingir o valor de pico, o número de infectados decai até zero, instante em que a epidemia cessa.

3.2.2 Modelo Estocástico

Muitos processos biológicos têm um elemento de incerteza. Por exemplo, o período de contágio da gripe ocorre de um até cinco dias após o início dos sintomas, mas o tempo exato em que um indivíduo infectado pode contagiar os demais indivíduos saudáveis da população não pode ser previsto. É possível incorporar esta aleatoriedade ou estocasticidade no formalismo matemático.

Para tanto, na versão estocástica do modelo SIR, as variáveis contínuas presentes no modelo determinístico são substituídas por número discretos, e as taxas β e γ são substituídas por probabilidades. Embora que o modelo SIR estocástico é conceitualmente diferente da sua versão determinística, uma comparação entre ambos é possível. Em verdade, o modelo de epidemias estocástico possui uma contraparte determinística, obtido por definir os incrementos populacionais determinísticos iguais aos valores esperados dos incrementos no modelo estocástico, para mais detalhes veja o trabalho de Abbey [20]. Assim, informações úteis sobre o modelo estocástico muitas vezes pode ser obtida a partir do seu modelo determinístico correspondente, desde que o tamanho da população seja grande. No modelo SIR estocástico, dado o estado $(S(t), I(t), R(t))$ e a correspondente equação determinística (3.3), temos

$$\langle I(t + dt) \rangle - \langle I(t) \rangle = \left(\frac{\beta}{N} S(t) - \gamma \right) I(t) dt + o(dt) \quad (3.11)$$

Porém, segue da equação anterior que

$$\frac{d\langle I(t) \rangle}{dt} = \left(\frac{\beta}{N} \langle S(t) \rangle - \gamma \right) \langle I(t) \rangle + \frac{\beta}{N} \sigma_{S(t), I(t)}^2, \quad (3.12)$$

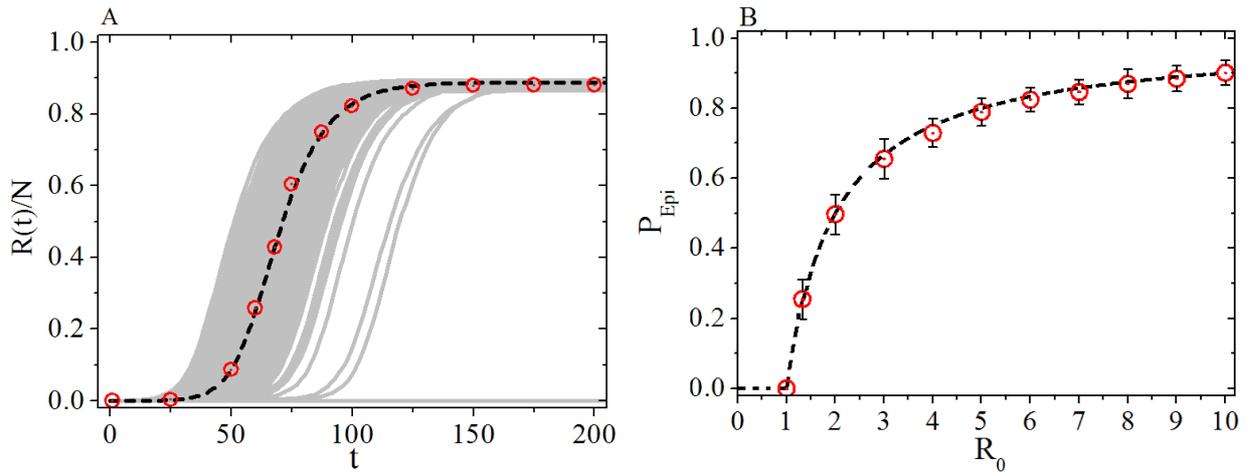


Figura 3.3: (A) Dinâmica dos removidos dadas as probabilidades $\beta = 0.26$ e $\gamma = 0.1$. (B) Probabilidade de ocorrência de uma epidemia. Os resultados correspondem a uma população de $N = 10^4$ sobre a condição inicial de um único indivíduo infectado. As linhas contínuas referem-se a distintas realizações, 500 no total. Os símbolos representam a média sobre as realizações. As linhas pontilhadas representam em (A) a solução numérica do sistema de equações (3.2-3.4) e em (B) a equação (3.17).

a partir da qual fica claro que a solução das equações deterministas não é simplesmente a média do processo estocástico. Esta é uma consequência inevitável da não-linearidade das taxas de transição do modelo estocástico. No entanto, para uma ampla classe de processos, a solução determinística é uma boa aproximação para a média estocástica das realizações que resultam em uma epidemia quando N é grande. Na Fig. 3.3A confrontamos os modelos estocástico e determinístico. Esta figura evidencia a variabilidade das muitas realizações. No entanto, a média sobre as realizações que resultam em uma epidemia, praticamente, não se diferencia da solução determinística.

Um dos principais aspectos em que modelos determinísticos e estocásticos se diferem é a extinção. Em um modelo SIR determinístico, como apresentado na subseção anterior, uma epidemia verdadeiramente nunca se extingue em um período de tempo limitado. De fato, o número de infectados diminui exponencialmente e chega a zero somente no infinito. Em contraste, no modelo estocástico uma epidemia pode extinguir-se, ou seja, o número de infectados podem tornar-se zero em um tempo finito.

Uma das alternativas para determinar a probabilidade de extinção P_{Ext} é usar o processo de ramificação de Galton-Watson [21]. Este processo trata-se de uma versão estocástica de um crescimento geométrico. Dentro deste contexto, supomos um cenário em que $R_0 > 1$ no qual um indivíduo infectado causa i novas infecções com probabilidade dada por uma distribuição geométrica, ou seja,

$$P_i = \frac{R_0^i}{(R_0 + 1)^{i+1}} \quad (3.13)$$

onde R_0 é o número médio de infecções secundárias. A função geratriz desta distribuição é

$$G(s) = \sum_{j=1}^{\infty} P_j s^j = \frac{1}{1 + R_0(1 - s)}. \quad (3.14)$$

Agora, podemos facilmente determinar a distribuição de infectados secundários na m -ésima geração $G^{(m)}(s)$ (a função G aplicada m vezes a s). Para determinar a probabilidade de extinção, por fim, devemos resolver

$$\lim_{m \rightarrow \infty} G^{(m)}(0) = q. \quad (3.15)$$

Segue imediatamente que q é solução da equação $G(q) = q$. Portanto, resulta

$$P_{Ext} = q = \frac{1}{R_0}. \quad (3.16)$$

Com base neste resultado, podemos definir a probabilidade de ocorrência de uma epidemia P_{Epi} . Para tanto, tomamos o complementar da probabilidade de extinção, isto é,

$$P_{Epi} = 1 - P_{Ext} = 1 - \frac{1}{R_0}. \quad (3.17)$$

Na Fig. 3.3B mostramos o comportamento da probabilidade da epidemia. Nesta figura observamos que uma epidemia é provável apenas para $R_0 > 1$.

3.3 Redes de Contatos

Modelos que incorporam a estrutura de redes evitam a suposição de classes de indivíduos homogeneamente misturados, designando a cada indivíduo um conjunto finito e

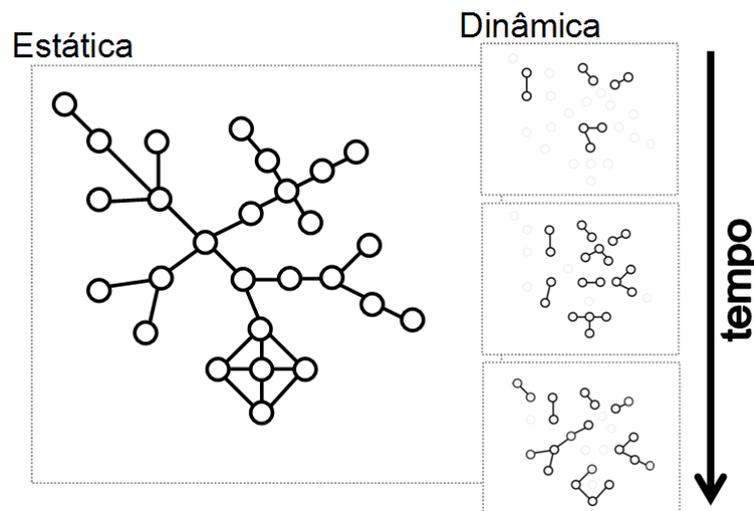


Figura 3.4: Representação de uma dinâmica social hipotética e as redes fundamentadas tanto pela formação dinâmica de pares em tempos distintos (Dinâmica) como pela história completa da rede (Estática), ou seja, superposição de todas as interações ao longo do tempo.

permanente de contatos, para os quais eles podem transmitir infecções e a partir 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 campo médio, na primeira o conjunto de contatos é fixo, enquanto que no campo médio está mudando continuamente. Ao considerar uma rede de contatos devemos ter em mente que as interações sociais ocorrem com uma certa variabilidade no tempo e, conseqüentemente, a estrutura da vizinhança de um dado indivíduo emerge completamente baseado na história local da rede, veja figura 5.1. Nas subseções seguintes discutiremos o resultado de aplicar uma doença infecciosa nestes diferentes cenários, estático e dinâmico.

3.3.1 Rede Estática

Na última década, tornou-se evidente que a transmissão de uma doença é dependente da rede contatos entre indivíduos, a qual depende da estrutura de interação deles (Draief [22]). Discutiremos na sequência a influência da topologia, primeiramente, no estado final e, posteriormente, na dinâmica da doença.

Estado Final da Doença

Considere um par de indivíduos que estão conectados, um i infectado e outro j suscetível. Suponha que a taxa média de transmissão da doença, causada pelo contato entre eles, seja β^{ij} e o indivíduo infectado permaneça neste estado por um tempo τ_d^i . Então, a probabilidade que a doença não será transmitida de i para j é $1 - T_{ij} = (1 - \beta^{ij})^{\tau_d^i}$, particularmente, para tempo discreto (Newman [23]). Com efeito, podemos caracterizar a probabilidade média que um indivíduo infectado transmite a doença para um suscetível de sua vizinhança

$$T = \langle T_{ij} \rangle. \quad (3.18)$$

Chamamos T de “transmissibilidade” da doença, $0 \leq T \leq 1$.

Agora imagine-se assistindo a propagação da doença, que iniciou com um único indivíduo infectado, espalhando-se através de uma rede. Se marcamos ou “ocuparmos” cada aresta na rede através do qual a doença é transmitida, com probabilidade T , no final do processo observaremos os vértices (tamanho do *cluster*) que podem ser alcançados, a partir do vértice inicial, ao percorrer as arestas marcadas. Desta maneira, determinar o tamanho final da doença em redes é equivalente ao problema de percolação na rede.

Para determinar o comportamento do tamanho do cluster usamos ferramentas baseadas na teoria de percolação (percolação sobre arestas), usando a técnica de função geratriz.

Seguindo Newman et al [24], a função geratriz para a distribuição de grau

$$G_0(x) = \sum_{k=1}^{\infty} P(k)x^k, \quad (3.19)$$

e os momentos da distribuição de grau, dados por

$$\langle k^n \rangle = \left[\left(x \frac{d}{dx} \right)^n G_0(x) \right]_{x=1}. \quad (3.20)$$

Assim, dado um vértice escolhemos uma das suas aresta ao acaso e, posteriormente, seguimos esta para o vértice vizinho. Resulta que o número de arestas que emanam deste vértices para outros é dado por

$$G_1(x) = \frac{G'_0(x)}{G'_0(1)} = \frac{1}{\langle k \rangle} G'_0(x) \quad (3.21)$$

Ainda, para resolver o problema de percolação, precisamos das funções geratrizes $G_0(x; T)$ e $G_1(x; T)$ para a distribuição do número de arestas ocupadas ligadas aos vértices, como uma função da transmissibilidade T . A probabilidade de um vértice ter m das k arestas ocupadas é dado pela distribuição binomial $C_m^k T^m (1 - T)^{k-m}$ e então a distribuição de probabilidade de m é gerada por

$$\begin{aligned} G_0(x; T) &= \sum_{m=0}^{\infty} \sum_{k=m}^{\infty} P(k) C_m^k T^m (1 - T)^{k-m} x^m \\ &= \sum_{m=0}^{\infty} P(k) (1 - T + xT)^k \\ &= G_0(1 + (x - 1)T). \end{aligned} \quad (3.22)$$

Similarmente, a distribuição de probabilidade de arestas ocupadas que levam a um vértice seguindo uma aresta aleatoriamente escolhida é gerada por

$$G_1(x; T) = G_1(1 + (x - 1)T). \quad (3.23)$$

Feita esta introdução das ferramentas, passamos a determinar a distribuição do tamanho s dos *clusters* de vértices conectados por arestas ocupadas (*outbreak*). Seja $H_0(x; T)$ a função geratriz desta distribuição,

$$H_0(x; T) = \sum_{s=0}^{\infty} P_s(T) x^s. \quad (3.24)$$

Por analogia com o parágrafo anterior, definimos $H_1(x; T)$ a função geratriz para os *clusters* de vértices conectados que alcançamos seguindo uma aresta escolhida aleatoriamente. É possível expressar $H_1(x; T)$ por uma forma auto-consistente (veja a representação geométrica na figura 3.5), dado por

$$H_1(x; T) = xG_1(H_1(x; T); T). \quad (3.25)$$

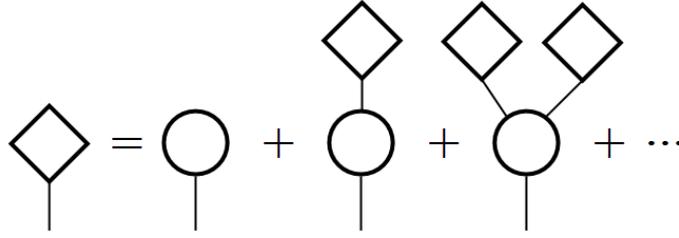


Figura 3.5: Representação dos possíveis *clusters* alcançados dada a escolha de um vértice aleatório.

Então, o tamanho do cluster alcançado para um vértice inicialmente escolhido é distribuído de acordo

$$H_0(x; T) = xG_0(H_1(x; T); T). \quad (3.26)$$

Podemos obter o *outbreak* médio, $\langle s \rangle$, usando a equação (3.24), temos

$$\langle s \rangle = H'_0(1; T) = 1 + G'_0(1; T)H'_1(1; T) \quad (3.27)$$

Diferenciando a equação (3.25) e usando os resultados (3.22) e (3.23), obtemos

$$\langle s \rangle = 1 + \frac{T \langle k \rangle}{1 - T(\langle k^2 \rangle / \langle k \rangle - 1)}. \quad (3.28)$$

Como exemplo, aplicamos este resultado a uma rede *scale-free* de distribuição de grau dada por $P(k) = k^{-\mu} e^{-k/\alpha} / Li_\mu(e^{-1/\alpha})$ com $\mu = 2$ (onde $Li_n(x)$ é o n -ésimo polilogaritmo de x), veja figura 3.6A para dois valores diferentes de α .

Notamos que a equação (3.28) diverge quando o denominador é nulo, desta maneira, temos

$$T_c = \frac{\langle k \rangle}{\langle k^2 \rangle - \langle k \rangle}. \quad (3.29)$$

Para $T > T_c$, temos uma epidemia, ou “componente gigante” na linguagem de percolação. Podemos calcular o tamanho da epidemia. Sobre o limiar de epidemia a equação (3.25) não é válida visto que a componente gigante é extensiva e, portanto, contém *loops*, que não estão de acordo nas suposições em que a equação (3.25) esta baseada. A equação é válida se redefinimos H_0 sendo a função geratriz para *clusters* de vértices isolados que não estão conectados a componente gigante. Então, sobre a transição da epidemia, temos

$$H_0(1; T) = \sum_s P_s = 1 - R(T) \quad (3.30)$$

onde $R(T)$ é a fração da população afetada pela epidemia. Ainda, podemos reescrever R dadas as equações (3.24), (3.22) e (3.19), como

$$R = 1 - G_0(u; T) = 1 - \sum_{k=1}^{\infty} P(k)(1 + (u - 1)T)^k \quad (3.31)$$

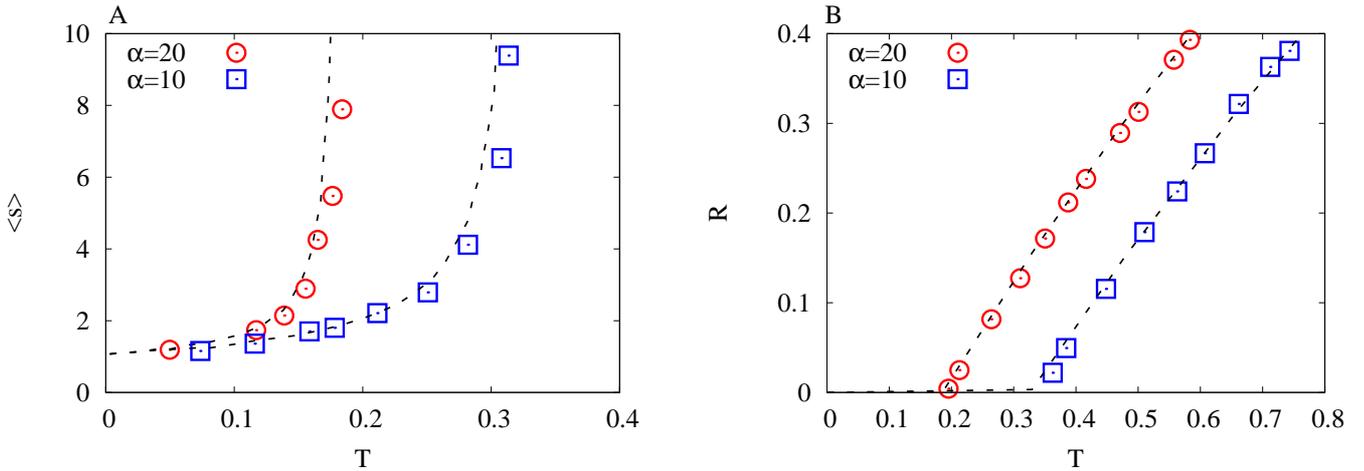


Figura 3.6: (A) Tamanho médio do *Outbreak* e (B) tamanho final da epidemia para o modelo SIR. As linhas correspondem aos resultados analíticos, equações, e os pontos as simulações de Monte Carlo para redes de tamanho $N = 10^4$.

onde u , em termos de epidemia, é a probabilidade que um dado vizinho de uma pessoa não esteja infectado, sendo solução da equação

$$u = G_1(u; T) = \frac{\sum_{k=1}^{\infty} k P(k) (1 + (u - 1)T)^{k-1}}{\langle k \rangle}. \quad (3.32)$$

Note que, mesmo acima de T_c , todos os *outbreaks* dados levam a epidemia da doença. Existem ainda *outbreaks* finitos mesmo no regime de epidemia. Enquanto que isso parece muito natural, no modelo de equações diferenciais (como discutido na seção anterior), para que todos os focos dão origem a uma epidemia, desde que estejamos acima do limiar epidemiológico. No presente caso, a probabilidade de um *outbreak* tornar-se uma epidemia num dado T é simplesmente igual a $R(T)$. A figura 3.6B trata de um exemplo quando submetemos as duas últimas expressões a uma distribuição de grau $P(k) = k^{-\mu} e^{-k/\alpha} / Li_{\mu}(e^{-1/\alpha})$ com $\mu = 2$ (onde $Li_n(x)$ é o n -ésimo polilogaritmo de x).

Ainda, relacionamos a transmissibilidade T com número de reprodução básica R_0 . Vimos, na seção anterior, que R_0 é o número de infecções secundárias causado por um indivíduo infectado numa população de suscetíveis. No contexto da teoria de percolação, corresponde ao número médio de aresta ocupadas que emanam de um vértice, que é de acordo com a equação (3.23)

$$R_0 = G'_1(1; T) = T G'_1(1) = T \left(\frac{\langle k^2 \rangle - \langle k \rangle}{\langle k \rangle} \right). \quad (3.33)$$

Observamos, da última expressão, que se a transmissibilidade de uma doença é igual ao limiar da epidemia ($T = T_c$), então $R_0 = 1$. Note que o número de reprodução básica depende explicitamente da topologia da rede (em $\langle k \rangle$ e $\langle k^2 \rangle$). Se duas redes tem o mesmo

grau médio mas diferem na variância do grau, será mais vulnerável para a propagação da doença a que apresentar maior variância.

Dinâmica da Doença

Todas as ferramentas que utilizamos até o momento, via teoria de percolação, mostraram-se adequadas na caracterização do estado final da epidemia, $R(t \rightarrow \infty)$. Porém, muito do que se desenvolve em política de prevenção pública depende da previsão da dinâmica da progressão da doença, $(S(t), I(t), R(t))$. No que segue, estenderemos estas ferramentas para descrever a dinâmica de uma doença em uma rede de contatos.

De acordo com Volz [25], para analisar os efeitos da estrutura de contatos na propagação da doença, consideramos um vértice suscetível i no tempo t com grau k em uma rede de distribuição de grau $P(k)$. Destes k vizinhos do indivíduo suscetível, supomos que M_{SS} corresponde a fração de suscetíveis enquanto que M_{SI} a fração de infectados. Assim, definimos que a probabilidade de encontrar um indivíduo j vizinho do indivíduo i no estado infectado é dado por $p_I = M_{SI}/M_S$, onde M_S representa a fração de vértices suscetíveis na rede. De mesma forma, temos $p_S = M_{SS}/M_S$ a probabilidade de um indivíduo i suscetível ter um vizinho j suscetível. Consequentemente, existe um número médio de kp_I vizinhos infectados. Em um tempo dt , o número esperado $\beta kp_I dt$ infectará o indivíduo i . Decorre disto, que a probabilidade do indivíduo i ser infectado no tempo t é dado por

$$\lambda_k(t) = \beta kp_I(t). \quad (3.34)$$

Agora, definimos $u_k(t)$ a fração de vértices de grau k que permanece suscetível no tempo t . Usando a equação anterior,

$$u_k(t) = \exp\left(-\int_0^t \lambda_k(t') dt'\right) = \theta^k \quad (3.35)$$

onde $\theta = \exp\left(-\int_0^t \beta kp_I(t') dt'\right)$. Dado θ , a fração de vértices que permanece suscetível no tempo t é dado por

$$S = \sum_{k=1}^{\infty} P(k)\theta^k = G(\theta). \quad (3.36)$$

Além disso, para a dinâmica de θ temos que

$$\frac{1}{\theta} \frac{d\theta}{dt} = -\lambda_1(t) \longrightarrow \dot{\theta} = -\beta p_I \theta \quad (3.37)$$

Infelizmente, isto não especifica completamente a dinâmica para θ e por extensão S , que também depende da variável p_I . Segue que,

$$\dot{p}_I = \frac{d}{dt} \left(\frac{M_{SI}}{M_S} \right) = \frac{\dot{M}_{SI}}{M_S} - \frac{\dot{M}_S M_{SI}}{M_S^2}. \quad (3.38)$$

O nosso objetivo, agora, é colocar a equação anterior em termos de p_I , p_S e θ . Para tanto, dado que P_k^S representa a distribuição de grau dos indivíduos suscetíveis, exibimos M_S como

$$M_S = \frac{\sum_k k p_k P_k^S}{\sum_k k P(k)} = \theta \frac{G'(\theta)}{G'(1)}. \quad (3.39)$$

Deste resultados, obtemos para

$$M_{SI} = M_S \frac{M_{SI}}{M_S} = p_I \theta \frac{G'(\theta)}{G'(1)} \quad (3.40)$$

Com base nos dois últimos resultados, após de um pouco de manipulação algébrica, obtemos

$$\dot{M}_S = \frac{d}{dt} \left(\theta \frac{G'(\theta)}{G'(1)} \right) = -\beta p_I \theta G'(\theta) - \beta p_I \theta^2 \frac{G''(\theta)}{G'(1)} \quad (3.41)$$

$$\dot{M}_{SI} = \frac{d}{dt} \left(p_I \theta \frac{G'(\theta)}{G'(1)} \right) = \beta p_I (p_S - p_I) \theta^2 \frac{G''(\theta)}{G'(1)} - (\beta + \gamma) M_{SI} \quad (3.42)$$

Com efeito, obtemos a dinâmica de p_I , ou seja,

$$\dot{p}_I = \beta p_S p_I \theta \frac{G''(\theta)}{G'(\theta)} - \beta p_I (1 - p_I) - \gamma p_I. \quad (3.43)$$

De maneira similar, a dinâmica de p_S é dada por

$$\dot{p}_S = \beta p_S p_I \left(1 - \theta \frac{G''(\theta)}{G'(\theta)} \right). \quad (3.44)$$

Em resumo, a dinâmica de uma infecção neste cenário, é dado pelo sistema de equações

$$\begin{aligned} \dot{\theta} &= -\beta p_I \theta \\ \dot{p}_I &= \beta p_S p_I \theta \frac{G''(\theta)}{G'(\theta)} - \beta p_I (1 - p_I) - \gamma p_I \\ \dot{p}_S &= \beta p_S p_I \left(1 - \theta \frac{G''(\theta)}{G'(\theta)} \right) \\ S &= G(\theta) \end{aligned} \quad (3.45)$$

$$(3.46)$$

sujeitas as condições iniciais

$$\begin{aligned} \theta(0) &= 1 - i_0 \\ p_I(0) &= \frac{i_0}{1 - i_0} \\ p_S(0) &= \frac{1 - 2i_0}{1 - i_0}, \end{aligned}$$

onde i_0 representa uma pequena fração de vértices escolhidos aleatoriamente como infectados iniciais. Adicionalmente, podemos contabilizar a fração de vértices infectados. Perceba

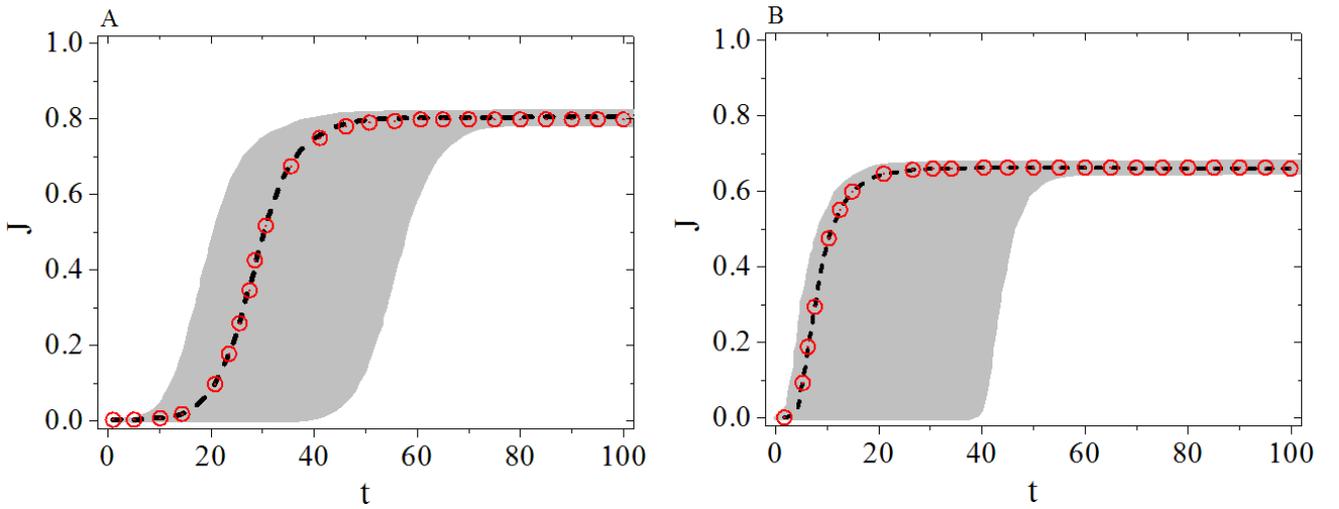


Figura 3.7: Dinâmica da fração de incidentes ($J = 1 - S$) para o modelo SIR em duas redes. (A) Poisson ($z = 3$) e (B) livre de escala ($\alpha = 20$ e $\mu = 1.615$). Resultados obtidos para $N = 10^4$, $\epsilon = 10^{-4}$, $\beta = 0.2$ e $\gamma = 0.1$. A linha tracejada corresponde a solução numérica do sistema de equações 3.45. Os símbolos correspondem a uma média sobre 500 realizações. A região sombreada indica a variabilidade na dinâmica dos incidentes.

que a fração de infectados aumenta com uma taxa $-\dot{S}$ e decresce com uma taxa de γI . Consequentemente,

$$\dot{I} = \beta p_I G'(\theta) - \gamma I. \quad (3.47)$$

Como exemplo, tomaremos redes cujos vértices possuem distribuição de grau dados por: (i) Poisson ($P(k) = z^k e^{-z}/k!$) e (ii) livre de escala ($P(k) = k^{-\mu} e^{-k/\alpha}/Li_\mu(e^{-1/\alpha})$). As Fig. 3.7A e 3.7B apresentam os resultados de 500 realizações para redes cuja a distribuição são dadas, respectivamente, por uma Poisson e livre de escala. Em ambos os casos, a trajetória analítica (linha tracejada) coincide com a média sobre as muitas realizações provenientes das simulações (símbolos). A trajetórias das simulações também apresentam uma variabilidade significativa no tempo necessário para atingir a fase de expansão e o tamanho final. Isto principalmente deve-se ao impacto significativo da aleatoriedade no início da epidemia. Por exemplo, um primeiro infectado com um grau médio baixo, ou um que tenha uma enorme quantidade de tempo para infectar o próximo infectado pode retardar o início da fase de expansão.

3.3.2 Rede Dinâmica

Um aspecto que usualmente é ignorado na abordagem de redes, como desenvolvemos na seção anterior, é a dinâmica das interações sociais (Read [26]). Em verdade, supor

que a rede estática só se justifica quando a dinâmica das interações sociais varia em escalas temporais que são maiores que a dinâmica da própria doença. Desta maneira, para averiguar a influência da variabilidade das interações frente a doença devemos considerar a dinâmica dos parceiros, que na linguagem de redes, corresponde a formação e dissolução das arestas entre os vértices.

De acordo com Volz e Meyers [27], nós podemos introduzir um modelo de contatos dinâmicos como uma simples extensão do modelo apresentado na subseção anterior. Para tanto, consideramos que o número de contatos dos indivíduo permanece fixo enquanto que a composição destes contatos muda com uma taxa específica. O modelo assume que em um dado tempo, um indivíduo possui um número de contatos (vizinhos) com os quais pode transmitir ou receber uma infecção. Cada contato é temporário com duração de uma certa quantidade de tempo. Após cessado este tempo, o vizinho em questão é substituído por um indivíduo diferente. Neste modelo, um dado contato (i, j) será substituído por outro (i, j') em uma taxa ρ .

Agora, nós estenderemos o modelo dado pelo sistema de equações (3.45) para permitir uma mudança de vizinhos com uma taxa ρ . Inicialmente perceba que uma troca de aresta afetará o arranjo das arestas sobre os vértices suscetíveis, infectados ou removidos. Porém, esta mudança nunca modificará o estado da doença. Com efeito, a dinâmica de θ é indiretamente influenciado pela mudança de um contato, através da mudança de p_I . Decorre disto que a equação para θ não se altera, ou seja, a dinâmica permanece dada pela equação 3.37.

Ao contrário de θ , a troca de contatos afeta a composição de contatos entre vértices suscetíveis e infectados. Para quantificar este efeito, Volz e Meyers postularam que as equações para \dot{p}_I (3.43) e \dot{p}_S (3.44) podem ser expressos pela seguinte modificação

$$\dot{p}_I = \beta p_S p_I \theta \frac{G''(\theta)}{G'(\theta)} - \beta p_I (1 - p_I) - \gamma p_I + f_I(p_I, M_I) \quad (3.48)$$

$$\dot{p}_S = \beta p_S p_I \left(1 - \theta \frac{G''(\theta)}{G'(\theta)} \right) + f_S(p_S, M_S) \quad (3.49)$$

onde a função $f_x(\cdot, \cdot)$ representa a contribuição da dinâmica de contatos para o sistema.

A dinâmica de contatos aumenta ou reduz os valores de p_I e p_S . Primeiro, consideramos o decréscimo de p_I :

- Em uma taxa ρ , um dado contato (i, j) será transformado em um novo contato (i, j') .
- Dado o contato (i, j) , $j \in I$ com probabilidade p_I
- Seja o novo contato (i, j') , $j' \notin I$ com probabilidade $1 - M_I$.

Então, p_I decresce com uma taxa $\rho p_I (1 - M_I)$. Similarmente, p_I cresce como segue:

- Em uma taxa ρ , uma dado contato (i, j) será transformado em um novo contato (i, j') .
- Dado o contato (i, j) , $j \notin I$ com probabilidade $1 - p_I$
- Seja o novo contato (i, j') , $j' \in I$ com probabilidade M_I .

Então, p_I cresce com uma taxa $\rho(1-p_I)M_I$. Superpondo os efeitos de acréscimo e decréscimo, obtemos que

$$f_I(p_I, M_I) = \rho(M_I - p_I). \quad (3.50)$$

No entanto, perceba que a variável dinâmica M_I não está em termos de p_I , p_S e θ . Portanto, uma quarta equação deverá ser incluída na dinâmica. A dinâmica de M_I pode ser determinada similarmente a dinâmica de M_S , tal como realizamos na subseção anterior, ou seja,

$$\dot{M}_I = \rho p_I \left(\frac{\theta^2 G''(\theta) + \theta G'(\theta)}{G'(1)} \right) - \gamma M_I. \quad (3.51)$$

Pelas mesmas razões, concluímos que

$$f_S(p_S, M_S) = \rho(M_S - p_S) = \rho \left(\frac{G'(\theta)}{G'(1)} - p_S \right). \quad (3.52)$$

Finalmente, a propagação de uma doença tipo SIR em uma rede dinâmica (número de contatos fixo) em que a vizinhança muda a uma taxa constante ρ , é dada pelo conjunto de equações

$$\begin{aligned} \dot{\theta} &= -\beta p_I \theta \\ \dot{p}_I &= \beta p_S p_I \theta \frac{G''(\theta)}{G'(\theta)} - \beta p_I (1 - p_I) - \gamma p_I + \rho(M_I - p_I) \\ \dot{p}_S &= \beta p_S p_I \left(1 - \theta \frac{G''(\theta)}{G'(\theta)} \right) + \rho \left(\frac{G'(\theta)}{G'(1)} - p_S \right) \\ \dot{M}_I &= \rho p_I \left(\frac{\theta^2 G''(\theta) + \theta G'(\theta)}{G'(1)} \right) - \gamma M_I \\ S &= G(\theta), \end{aligned} \quad (3.53)$$

sujeitas as condições iniciais

$$\begin{aligned} M_I(0) &= i_0 \\ \theta(0) &= 1 - i_0 \\ p_I(0) &= \frac{i_0}{1 - i_0} \\ p_S(0) &= \frac{1 - 2i_0}{1 - i_0} \end{aligned}$$

onde i_0 representa uma pequena fração de vértices escolhidos aleatoriamente como infectados iniciais.

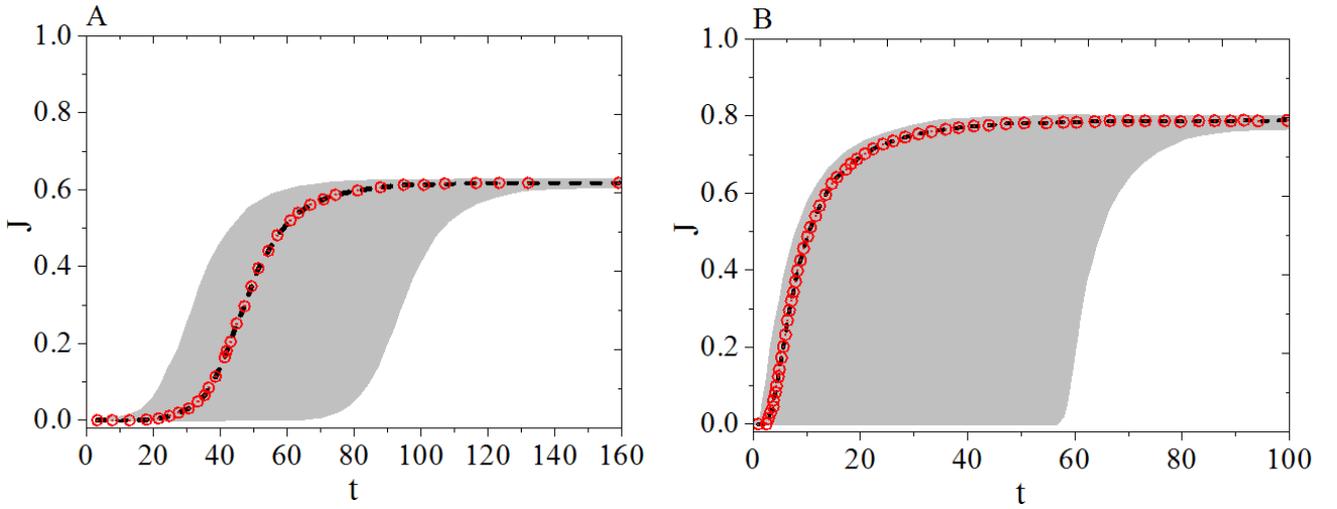


Figura 3.8: Dinâmica da fração de incidentes ($J = 1 - S$) para o modelo SIR em duas redes. (A) Poisson ($z = 3$) e (B) livre de escala ($\alpha = 20$ e $\mu = 1.615$). Resultados obtidos para $N = 10^4$, $\epsilon = 10^{-4}$, $\beta = 0.2$, $\gamma = 0.1$ e $\rho = 0.2$. A linha tracejada corresponde a solução numérica do sistema de equações (3.53). Os símbolos correspondem a uma média sobre 500 realizações. A região sombreada indica a variabilidade na dinâmica dos incidentes.

As Fig. 3.8A e 3.8B apresentam uma comparação da simulação (500 realizações) como o modelo determinístico dado pelo sistema de equações (3.53) para duas distribuições de grau: (i) Poisson ($P(k) = z^k e^{-z}/k!$) e (ii) livre de escala ($P(k) = k^{-\mu} e^{-k/\alpha}/Li_\mu(e^{-1/\alpha})$). Desta comparação podemos observar que o modelo determinístico tem uma boa concordância com a média destas realizações e com o tamanho final da epidemia. Em contraste a rede aleatória, a rede livre de escala fornece uma fase de expansão quase que imediata.

Um importante resultado para este modelo trata-se da convergência das soluções para a o modelo de campo médio (veja seção 3.2). No que segue, mostraremos que se a taxa de mudança entre os contatos cresce, a estrutura da rede torna-se irrelevante e o modelo converge para a aproximação de campo médio. No limite $\rho \rightarrow \infty$, a probabilidade de estar ligado a um vértice suscetível, infectado ou removido é diretamente proporcional ao grau dos vértices em cada estado. Consequentemente, temos que p_I converge para M_I e p_S converge para M_S . Sendo assim, temos que

$$\dot{\theta} = -\beta M_I \theta \quad (3.54)$$

$$\dot{M}_I = \rho M_I \left(\frac{\theta^2 G''(\theta) + \theta G'(\theta)}{G'(1)} \right) - \gamma M_I \quad (3.55)$$

Note que nenhuma destas equações depende de p_S e, portanto, juntas elas formam um sistema de equações que descreve a dinâmica da epidemia. Para ilustrar isto, definimos $S = G(\theta) = \theta$, ou seja, cada vértice tem exatamente um contato. Em uma população, o número de contatos de indivíduos infectados é exatamente igual ao número de vértices

infectados, isto é, $p_I = I$. Consequentemente, obtemos a dinâmica de campo médio, ou seja,

$$\dot{\theta} = -\beta M_I \theta = -\beta SI \quad (3.56)$$

$$\dot{M}_I = \beta M_I \theta - \gamma M_I = \beta SI - \gamma I. \quad (3.57)$$

Portanto, verificamos que o modelo de campo médio é uma boa aproximação para populações com contatos de curta duração.

Capítulo 4

Uma Solução Algébrica para o Modelo de Kermack-McKendrick

O modelo de Kermack-McKendrick ou modelo SIR, veja seção 3.2, é um modelo matemático aplicado a dinâmica de propagação de doenças infecciosas. O número básico de reprodução $R_0 = \beta\tau$ e a curva de infecção (que é o número assintótico de removidos como função de R_0) são conceitos fundamentais em epidemiologia. É verdade, porém, que epidemias reais são dificilmente descritas em termos do modelo SIR (ou uma das variantes SIS, SIRS, etc). Apesar disso, é quase impossível “falar” a linguagem da descrição matemática de epidemias, em qualquer nível de complexidade, sem ter o modelo SIR, suas equações, seu limiar e estado assintótico em mente. Se uma metáfora com a física pudesse ser aplicada, diríamos que o modelo SIR é para a dinâmica de epidemias como o oscilador harmônico é para física. Não podemos sequer pensar em um problema de mecânica quântica, por exemplo, sem ter o oscilador harmônico em mente. No entanto, a metáfora não pode ser sustentado em todos os níveis, porque diferente do exemplo físico, as equações do modelo SIR carecem de uma solução analítica. Perseguido por mais de 80 anos depois de terem sido apresentada, a fim de se obter o comportamento dinâmico das quantidades S , I , e R a solução numérica por computador — usando uma integração numérica tipo Euler — era a única maneira de tratá-la. Felizmente a partir de 1980 com o poder crescente de computadores fez esta uma solução simples e rápida, evitando toda o trabalho de cálculos numéricos com calculadoras. Mas a presente facilidade não diminui a beleza e a satisfação intelectual de uma solução analítica.

Neste capítulo nós apresentamos uma solução analítica para o clássico modelo SIR como formulado por Kermack-McKendrick. Para realizar esta tarefa, primeiro apresentamos uma solução exata e fechada, exceto por uma integral. Em seguida, realizamos uma expansão em série que nos permite apresentar uma solução aproximada para a referida integração. Por

fim, fundamentados nesta formulação indicamos uma conveniente aproximação analítica para evolução temporal dos estados da doença.

4.1 Nova Representação

O modelo de Kermack-McKendrick é dado por

$$\begin{aligned}\frac{dS}{d\tau} &= -R_0SI \\ \frac{dI}{d\tau} &= R_0SI - I \\ \frac{dR}{d\tau} &= I,\end{aligned}\tag{4.1}$$

onde S , I , e R representa, respectivamente, a fração de suscetíveis, infectados e removidos dentro de uma população de tamanho fixo ($S + I + R = 1$), $R_0 = \beta/\gamma$ é o número básico de reprodução e $\tau = \gamma t$.

No artigo original de Kermack-McKendrick, depois de apresentar as equações (4.1), os autores obtêm uma expressão para $S(R)$ partindo da primeira e terceira equação diferencial. Assumindo que não há sujeitos removidos no início da infecção ($R(0) = 0$) eles concluem que

$$S(\tau) = (1 - i_0)e^{-R_0R(\tau)}\tag{4.2}$$

onde $I(0) = i_0$ é a fração inicial de sujeitos infectados. Usando deste último resultado, eles apresentam uma equação diferencial independente para os sujeitos removidos

$$\frac{dR}{d\tau} = I = 1 - R - S = 1 - R - (1 - i_0)e^{-R_0R}\tag{4.3}$$

Aqui nós fazemos uma mudança de variável $w = R_0(1 - R)$, o que resulta

$$\frac{dw}{d\tau} = R_0(1 - i_0)e^{-R_0}e^w - w\tag{4.4}$$

dada a condição inicial $w(0) = R_0$. Assim, dada a definição de w e as equações (4.1) e (4.1), podemos expressar as classes da dinâmica da doença em termos de w , ou seja,

$$S(\tau) = (1 - i_0)e^{-R_0}e^{w(\tau)}\tag{4.5}$$

$$I(\tau) = \frac{w(\tau)}{R_0} - (1 - i_0)e^{-R_0}e^{w(\tau)}\tag{4.6}$$

$$R(\tau) = 1 - \frac{w(\tau)}{R_0}\tag{4.7}$$

Portanto, o problema de resolver as equações (4.1) foi transformado em resolver uma única equação diferencial para w (equação 4.1). Consequentemente, podemos expressar está solução

em termos de uma integral

$$\int \frac{dw(\tau)}{R_0(1-i_0)e^{-R_0}e^{w(\tau)} - w(\tau)} = \tau + k. \quad (4.8)$$

onde k é uma constante de integração. Vamos ver na próxima seção como lidar com esta integral.

4.2 Aproximação

Nós vimos na seção anterior que para resolver o modelo SIR nós precisamos integrar a equação (4.8), que não pode ser realizada analiticamente. Para superar este problema, nós introduziremos uma forma funcional equivalente a função de integração (equação 4.8). A ideia é usar uma forma funcional que pode ser integrada. Adicionalmente, afim de simplificar a notação, utilizaremos a usual condição inicial em que todos os sujeitos são suscetíveis exceto por um indivíduo infectado, que significa que $i_0 \simeq 0$.

Para este objetivo, nós representamos o termo exponencial no denominador da integral por uma série de potência, $e^w = \sum_{j=0}^{\infty} w^j/j!$. Nesta forma o denominador da integral (4.8) é um polinômio

$$P(w) = R_0 e^{-R_0} \sum_{j=0}^{\infty} \frac{w^j}{j!} - w. \quad (4.9)$$

Com efeito, nós podemos usar a técnica de fração parciais para obter a integral, ou seja,

$$\frac{1}{P(w)} = \sum_{j=1}^{\infty} \frac{A(z_j)}{w - z_j} \quad (4.10)$$

onde z_i são as raízes de $P(w)$ e os coeficientes $A(z_j)$ são dados por

$$A(z_j) = \left(\frac{dP(w)}{dw} \right)_{w=z_j}^{-1} = \frac{1}{R_0 e^{-R_0} e^{z_j} - 1}. \quad (4.11)$$

Agora o problema é obter todas as infinitas raízes do polinômio $P(w)$. Porém, uma simples inspeção numérica deste polinômio (Eq.4.10) mostra que existe somente uma raiz real se $R_0 = 1$ (condição do limiar). Para outros valores do número básico de reprodução acima do limiar ($R_0 > 1$) existem exatamente duas raízes reais para $P(w)$. Uma das raízes reais é $w_1 \approx R_0(1 + \frac{i_0}{R_0-1})$. A outra raiz real pode ser encontrada exatamente pela definição de w e a relação dada pela equação (4.1), ambas no limite assintótico (tempo tendendo ao infinito, assim $I(\infty) = 0$ e $S(\infty) + R(\infty) = 1$). Então a segunda raiz é $w_2 = w(\infty) = R_0 S(\infty)$, sendo que $S(\infty)$ é solução (de acordo com a equação (4.5)) da expressão

$$S(\infty) = (1 - i_0)e^{R_0(1-S(\infty))}. \quad (4.12)$$

Depois destas considerações, nós podemos escrever a soma na equação (4.10) como duas somas, uma sobre as raízes reais e outra sobre as complexas, ou seja,

$$\frac{1}{P(w)} = \sum_{z_j \in \mathfrak{R}} \frac{A(z_j)}{w - z_j} + \sum_{z_j \in \mathfrak{C}} \frac{A(z_j)}{w - z_j}. \quad (4.13)$$

Dentro do contexto do presente problema, nós sabemos que a soma que envolve as raízes complexas deve convergir para um valor real. Então, nós podemos escrever

$$\sum_{z_j \in \mathfrak{C}} \frac{A(z_j)}{w - z_j} = \alpha(w), \quad (4.14)$$

onde $\alpha(w) \in \mathfrak{R}$, e que dada a equação (4.13) pode ser expressado como

$$\alpha(w) = \frac{1}{R_0 e^{-R_0} e^w - w} - \sum_{z_j \in \mathfrak{R}} \frac{A(z_j)}{w - z_j}. \quad (4.15)$$

Agora nós realizamos uma expansão de Taylor de primeira ordem em torno do pico de infecção $w^* = w(\tau_p) = R_0 - \ln R_0$, obtemos

$$\begin{aligned} \alpha(w) &\simeq \alpha(w^*) + (w - w^*) \left(\frac{d\alpha(w)}{dw} \right)_{w=w^*} \\ &\simeq \frac{1}{1 - R_0 + \ln R_0} + \frac{w - R_0 + 2 \ln R_0}{(-1 + R_0) (\ln R_0)^2} + \frac{e^{R_0} (w + R_0(-2 + S(\infty)) + 2 \ln R_0)}{(e^{R_0} - R_0 e^{R_0 S(\infty)}) (R_0(-1 + S(\infty)) + \ln R_0)^2} \end{aligned} \quad (4.16)$$

Portanto,

$$\tau + k = \int \frac{dw}{P(w)} \simeq \int \overbrace{\sum_{z_j \in \mathfrak{R}} \frac{A(z_j)}{w - z_j}}^{F(w)} dw + \int \overbrace{\alpha(w)}^{G(w)} dw \quad (4.17)$$

onde

$$F(w) = \frac{\ln \left(w - \left(1 + \frac{i_0}{R_0 - 1} \right) R_0 \right)}{(1 - i_0) R_0 e^{\frac{R_0 i_0}{R_0 - 1}} - 1} + \frac{\ln (w - R_0 S(\infty))}{(1 - i_0) R_0 e^{(-1 + S(\infty)) R_0} - 1}, \quad (4.18)$$

$$G(w) \simeq \frac{w}{2} \left(\frac{2}{1 - R_0 + \ln R_0} + \frac{w - 2R_0 + 4 \ln R_0}{(-1 + R_0) (\ln R_0)^2} + \frac{e^{R_0} (w + 2R_0(-2 + S(\infty)) + 4 \ln R_0)}{(e^{R_0} - R_0 e^{R_0 S(\infty)}) (R_0(-1 + S(\infty)) + \ln R_0)^2} \right).$$

e

$$k \simeq F(w(0)) + G(w(0)) = F(R_0) + G(R_0). \quad (4.19)$$

Por fim, podemos determinar, aproximadamente, a integral (4.8) como

$$\begin{aligned} \tau &\simeq \frac{R_0 - w}{2} \left(\frac{2}{1 - R_0 + \ln R_0} + \frac{4}{(1 - R_0) \ln R_0} - \frac{R_0 - w}{(1 - R_0) (\ln R_0)^2} \right) + \frac{\ln \left[\frac{(1 - R_0)(R_0 - w)}{R_0^{i_0}} - 1 \right]}{(1 - i_0) R_0 e^{\frac{R_0 i_0}{R_0 - 1}} - 1} + \\ &+ \frac{(R_0 - w) e^{R_0} (w + R_0(-3 + 2S(\infty)) + 4 \ln R_0)}{2 (e^{R_0} - R_0 e^{R_0 S(\infty)}) (R_0(-1 + S(\infty)) + \ln R_0)^2} + \frac{\ln \left[\frac{w - R_0 S(\infty)}{(1 - S(\infty)) R_0} \right]}{(1 - i_0) R_0 e^{-(1 - S(\infty)) R_0} - 1} \end{aligned} \quad (4.20)$$

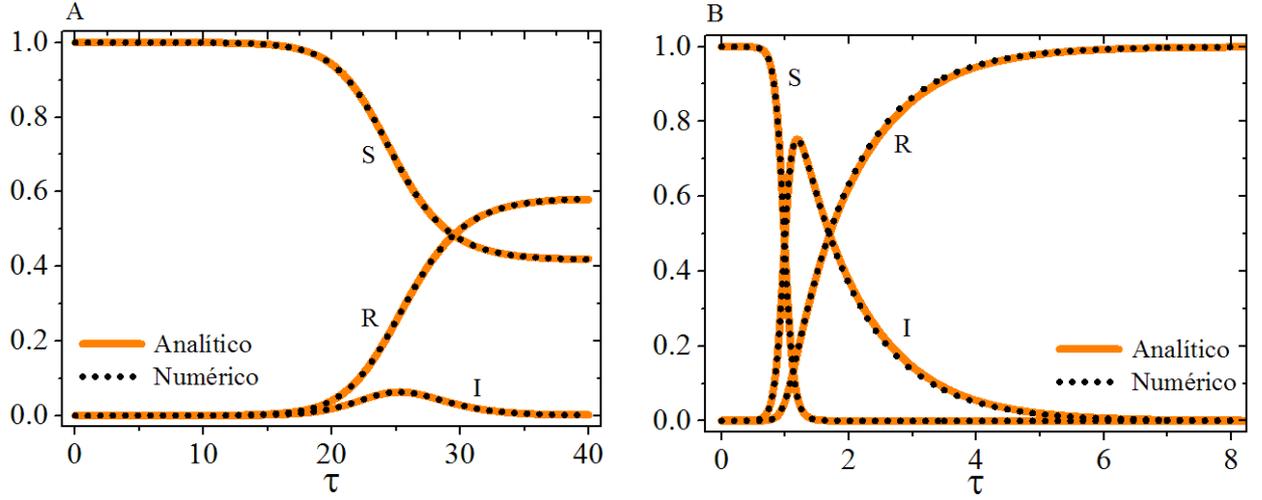


Figura 4.1: Resultado numérico versus analítico para a dinâmica das classes S , I e R . (A) $R_0 = 1.5$ e (B) $R_0 = 15$. Resultados obtidos para $i_0 = 10^{-6}$.

4.3 Solução Algébrica

Com base nas equações (4.5)-(4.7) e (4.20), a dinâmica SIR é determinada algebricamente pelo seguinte conjunto de equações

$$S(\tau) = (1 - i_0)e^{-R_0}e^{w(\tau)} \quad (4.21)$$

$$I(\tau) = \frac{w(\tau)}{R_0} - (1 - i_0)e^{-R_0}e^{w(\tau)} \quad (4.22)$$

$$R(\tau) = 1 - \frac{w(\tau)}{R_0} \quad (4.23)$$

$$\tau \simeq \frac{R_0 - w}{2} \left(\frac{2}{1 - R_0 + \ln R_0} + \frac{4}{(1 - R_0) \ln R_0} - \frac{R_0 - w}{(1 - R_0) (\ln R_0)^2} \right) + \frac{\ln \left[\frac{(1 - R_0)(R_0 - w)}{R_0 i_0} - 1 \right]}{(1 - i_0) R_0 e^{\frac{R_0 i_0}{R_0 - 1}} - 1} + \frac{(R_0 - w)e^{R_0} (w + R_0(-3 + 2S(\infty)) + 4 \ln R_0)}{2(e^{R_0} - R_0 e^{R_0 S(\infty)}) (R_0(-1 + S(\infty)) + \ln R_0)^2} + \frac{\ln \left[\frac{w - R_0 S(\infty)}{(1 - S(\infty)) R_0} \right]}{(1 - i_0) R_0 e^{-(1 - S(\infty)) R_0} - 1} \quad (4.24)$$

sujeita as condições iniciais $S(0) = 1 - i_0$, $I(0) = i_0$ and $R(0) = 0$. Note que a função $w(\tau)$ não é inversível. O que nos impossibilita de representar a dinâmica do estado da doença (S, I e R) explicitamente dependente do tempo.

Na figura 4.1 confrontamos a solução numérica e a analítica. Percebemos que há uma boa correspondência entre os resultados. No entanto, uma simples inspeção na figura 4.1B indica que a coincidência entre os resultados não é exata. Isto sugere que existe uma erro intrínseco a aproximação, o que é razoável. Verificamos os erros relativos caracterizando na curva dos removidos a diferença entre áreas obtidas via aproximação, equações (4.21)-(4.24),

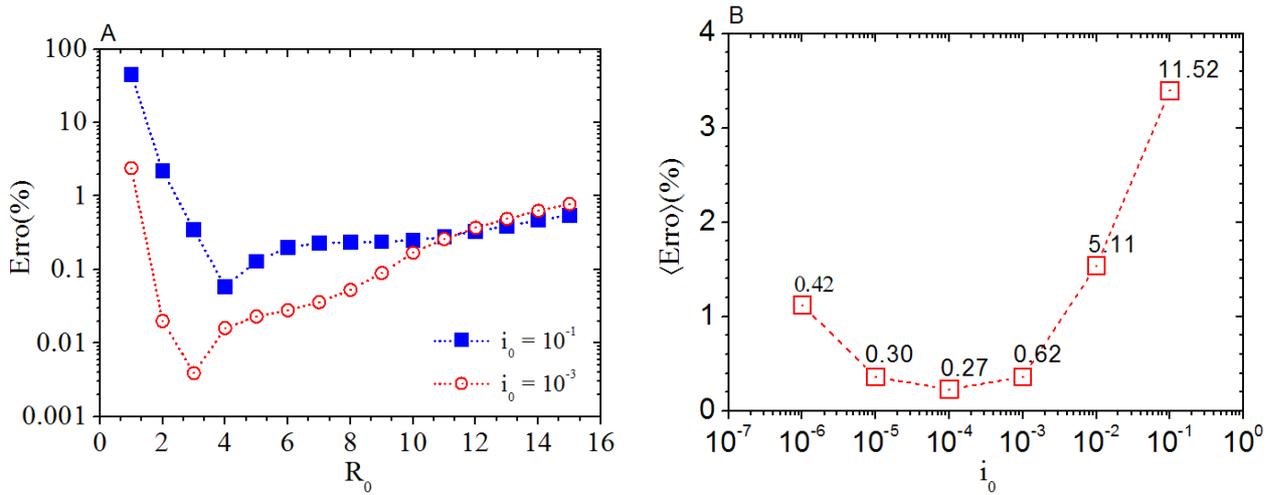


Figura 4.2: (A) Erro relativo em função de R_0 . (B) Erro médio relativo versus i_0 . Os números em destaque sobre os símbolos representam o desvio padrão. Média realizada para $R_0 \in [1, 15]$.

e a solução numérica do sistema (4.1). Para tanto, realizamos

$$\begin{aligned}
 \text{Erro} &= \frac{|\Delta_N - \Delta_A|}{\Delta_N} \\
 \Delta_x &= \int_0^\infty R_x(\tau) d\tau
 \end{aligned} \tag{4.25}$$

onde os índices N e A referem-se, respectivamente, as soluções numérica e analítica. Apresentamos na figura 4.2A o erro relativo versus o parâmetro R_0 . As curvas evidenciam que o erro depende tanto de R_0 como de i_0 . Em relação a dependência do erro com R_0 , por um lado, verificamos que para ambas as curvas existe um valor mínimo. Por outro, a esquerda e a direita deste R_0 crítico o valor do erro eleva-se a declividade distintas. Por exemplo, para uma fração de infectados inicial de $i_0 = 10^{-1}$ temos um mínimo em $R_0 = 3$ ($\text{Erro} = 0.0039\%$) e declividades de 1.198% ($1 < R_0 < 3$) e 0.064% ($3 < R_0 < 15$). Adicionalmente, comparando as curvas presentes nesta figura, vemos que o erro eleva-se com o acréscimo de i_0 . Na figura 4.2B indicamos o erro médio relativo para diferentes valores de i_0 .

Por fim, dados os resultados anteriores, podemos obter uma solução analítica para o tempo do pico de infecção, τ_p . Para tanto, substituímos as expressões $w(\tau_p) = R_0 - \ln R_0$ e $R(\infty) = 1 - S(\infty)$ na equação (4.20). Este resultado aliado as expressões do número básico de reprodução R_0 , número assintótico de removidos $R(\infty)$ e número máximo de infectados

I_{max} (apresentadas no artigo original de Kermack-McKendrick), ou seja,

$$R_0 = \frac{\beta}{\gamma} \quad (4.26)$$

$$I_{max} = 1 - \frac{1}{R_0} (1 + \ln[R_0(1 - i_0)]) \quad (4.27)$$

$$R(\infty) = 1 - (1 - i_0)e^{-R_0 R(\infty)} \quad (4.28)$$

$$\begin{aligned} \tau_p \simeq & \frac{1}{2} \left(\frac{\ln R_0 (3 \ln R_0 - 2R_0 R(\infty))}{(1 - R_0 e^{-R_0 R(\infty)}) (\ln R_0 - R_0 R(\infty))^2} + \frac{3 - 3R_0 + \ln R_0 + 2R_0 \ln R_0}{(-1 + R_0)(-1 + R_0 - \ln R_0)} \right) + \\ & + \frac{\ln \left[1 + \frac{(1-R_0) \ln R_0}{R_0 i_0} \right]}{(1 - i_0) R_0 e^{\frac{R_0 i_0}{R_0 - 1}} - 1} + \frac{\ln \left[1 - \frac{\ln R_0}{R_0 R(\infty)} \right]}{(1 - i_0) R_0 e^{-R_0 R(\infty)} - 1} \end{aligned} \quad (4.29)$$

representam de forma algébrica as principais quantidades do modelo SIR.

Capítulo 5

Modelo de Rede Dinâmica e Epidemias

A estrutura de contatos das populações influenciam a propagação de doenças infecciosas. Métodos matemáticos para previsão de epidemias em redes, em grande parte, assumem que os contatos entre os indivíduos são fixos. Na realidade, indivíduos, frequentemente, criam e aniquilam interações sociais ao longo do tempo. Sendo assim, os contatos sociais emergem da formação dinâmica das relações sociais.

Neste capítulo, desenvolvemos uma abordagem matemática para prever a transmissão de uma doença em redes dinâmicas na qual os contatos de cada indivíduo mudam no tempo. Inicialmente, primeira seção, apresentamos um modelo de dinâmica de contatos. Na segunda seção, utilizamos o modelo dinâmico de interação para implementar um modelo de infecção.

5.1 Rede Dinâmica

Consideramos uma população de N indivíduos isolados no tempo zero. Num dado instante de tempo, cada indivíduo tem uma probabilidade ρ/N de criar uma conexão com cada um dos demais indivíduos (tomado aleatoriamente entre os $N - 1$) e uma probabilidade σ de aniquilar cada uma das ligações formadas em tempo anteriores, veja a figura 5.1. Os parâmetros ρ e σ são constantes para todos os indivíduos sobre a evolução do tempo. Com efeito, dado que os indivíduos são idênticos, podemos determinar a dinâmica do grau médio pela equação

$$\frac{d\langle k \rangle_t}{dt} = \frac{\rho}{N} (N - \langle k \rangle_t) - \sigma \langle k \rangle_t. \quad (5.1)$$

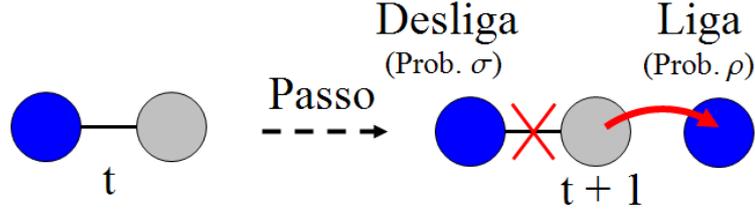


Figura 5.1: Dinâmica de interação entre os indivíduos. O modelo assume que cada indivíduo, em um dado passo de tempo, possui uma probabilidade ρ/N de fazer uma conexão com outro sujeito e uma probabilidade σ de remover uma conexão feita anteriormente.

dada a condição inicial $\langle k \rangle_0 = 0$. A solução desta equação diferencial é

$$\langle k \rangle_t = \frac{\rho}{\sigma} (1 - e^{-\sigma t}), \quad (5.2)$$

para $N \gg 1$. Note que $\tau_N = 1/\sigma$ representa o tempo característico de formação das interações. Adicionalmente, em relação ao grau médio $\langle k \rangle_t$, verificamos um comportamento assintótico

$$\lim_{t \rightarrow \infty} \langle k \rangle_t = \langle k \rangle_\infty = \frac{\rho}{\sigma}. \quad (5.3)$$

Aqui, vale salientar, se tomarmos $\langle k \rangle_0 = \langle k \rangle_\infty$ o grau médio permanece constante. Perceba que esta condição implica em um cenário ainda com dinâmica nas interações, no entanto, com o grau médio fixo ($\langle k \rangle_t = \langle k \rangle_\infty$). Desta maneira, analisaremos a evolução temporal das conexões em dois cenários: equilíbrio dinâmico ($\langle k \rangle_0 = \langle k \rangle_\infty$) e fora do equilíbrio ($\langle k \rangle_0 = 0$). Na figura 5.4B indicamos os dois cenários.

Retomando a dinâmica de formação, observamos que a mesma é dada em termos das probabilidades de criação, ρ , e aniquilação, σ . Contudo, a construção da rede pode ser pensada em uma forma alternativa. Supomos que a cada passo de tempo todos os indivíduos estão isolados e cada um destes tem uma probabilidade dependente do tempo, $\mathbf{p}(t)$, de criar uma interação com cada um dos demais. Em outras palavras, esta regra de interação não se difere da proposta por Erdos-Renyi a uma probabilidade $\mathbf{p}(t)$. Com efeito, a distribuição de grau é dada por uma distribuição binomial

$$P_{k'}(t) = \binom{N}{k'} \mathbf{p}(t)^{k'} (1 - \mathbf{p}(t))^{N-k'}. \quad (5.4)$$

Podemos estimar $\mathbf{p}(t)$ tomando a razão entre o número total de arestas e todas as possibilidades de conexão para um determinado tempo t , ou seja,

$$\mathbf{p}(t) = \frac{\langle k \rangle_t \frac{N}{2}}{\frac{N^2}{2}} = \frac{\langle k \rangle_t}{N} \quad (5.5)$$

e, conseqüentemente,

$$P_{k'}(t) = \binom{N}{k'} \left(\frac{\langle k \rangle_t}{N} \right)^{k'} \left(1 - \frac{\langle k \rangle_t}{N} \right)^{N-k'} \quad (5.6)$$

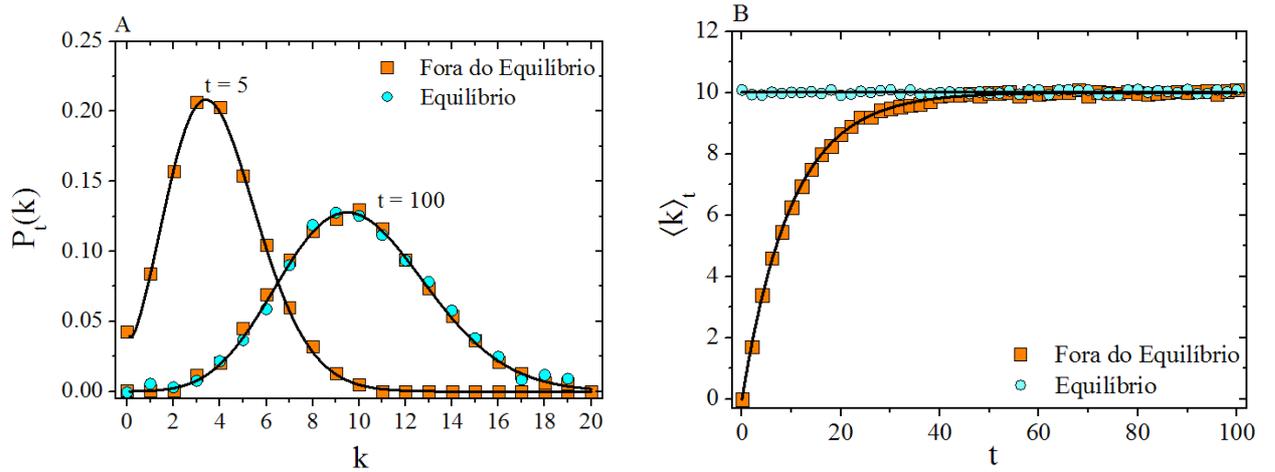


Figura 5.2: (A) Distribuição de grau em dois diferentes tempos ($t = 5$ e $t = 100$). (B) Evolução temporal do grau médio. Os símbolos referem-se as simulações enquanto que as linhas sólidas aos resultados analíticos. Resultados obtidos para $N = 10^4$ indivíduos, $\rho = 1$ e $\sigma = 0.1$. Cada ponto corresponde a uma média de 10^3 realizações.

Tomando os limites $N \rightarrow \infty$ e $\langle k \rangle_t \ll N$ (a exemplo da discussão na subseção 2.3.1), verificamos que

$$P_{k'}(t) \simeq \frac{\langle k \rangle_t^{k'} e^{-\langle k \rangle_t}}{k'!}. \quad (5.7)$$

Logo, a distribuição de grau evolui no tempo como uma Poisson de grau médio $\langle k \rangle_t$, figura 5.4A.

5.2 Epidemia

Consideramos a dinâmica da doença baseada no modelo SIR. Sendo assim, tomamos uma população de tamanho fixo N , cuja dinâmica de interação é dada pela seção anterior. Cada indivíduo da população está dividido em um dos três estados: suscetível, infectado e removido. O número de indivíduos nestes três estados em um dado tempo t é denotado por $S(t)$, $I(t)$ e $R(t)$. Um contato efetivo surge entre cada par de indivíduos de acordo com um processo de Poisson com taxa β . A transmissão da infecção acontece quando um contato efetivo ocorre entre um suscetível e infectado. Um indivíduo infectado permanece neste estado por um dado tempo t que é distribuído exponencialmente com uma taxa γ .

5.2.1 Número Básico de Reprodução

Para compreender a propagação da infecção na população, nós investigamos o que acontece quando um indivíduo infectado é introduzido em uma grande população de in-

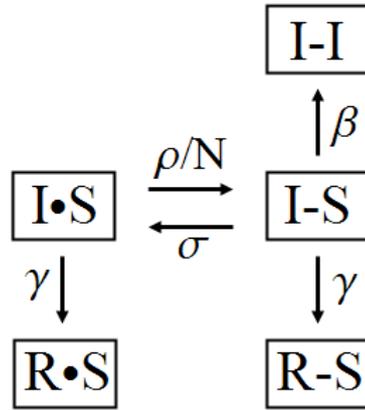


Figura 5.3: Taxas de transição para os estados de uma díade. Em termos dos estados temos: $I \cdot S$ em que primeiro indivíduo é infectado o segundo é suscetível e eles não são parceiros, $R - S$ o primeiro indivíduo é removido o segundo é suscetível e eles são parceiros, etc. A díade inicia com indivíduos infectado e suscetível na configuração $I \cdot S$ ou $I - S$.

divíduos suscetíveis. Denotamos o indivíduo infectados inicial por i . O número básico de reprodução é o número de novos indivíduos infectados por i . Dado que os indivíduos são idênticos, para determina-lo podemos calcular a probabilidade de transmissão da infecção do indivíduo i para algum indivíduo j . A chave para calcular esta probabilidade é considerar a configuração de uma díade (i, j) . Como cada indivíduo pode estar em um dos três possíveis estados da doença e os indivíduos podem ou não podem ser parceiros em um dado tempo t , existem 18 configurações para a díade. Estas configurações serão denotadas por duas letras, indicando o *status* dos dois indivíduos na díade. Um ponto central será usado para indicar que i e j não estão ligados, e um traço para indicar que eles estão ligados. Por exemplo, $S - S$ denota a configuração onde ambos são suscetíveis e parceiros e $I \cdot S$ representa a situação em que i é infectado e j é suscetível e eles não são parceiros. As mudanças de configuração de cada díade são dadas por um processo de Markov. Para determinar R_0 precisamos somente considerar a história de vida de cada díade da primeira infecção a remoção deste caso. Os estados e as correspondentes transições para este processo são mostrados na figura 5.3. Note que este diagrama possui três estados absorventes: $I - I$, $R \cdot S$ e $R - S$.

A díade, em um dado tempo t , inicia no estado $I - S$ ou $I \cdot S$ com probabilidade

$$\begin{aligned} P_{I-S}(t) &= \frac{\langle k \rangle_t}{N} \\ P_{I \cdot S}(t) &= 1 - \frac{\langle k \rangle_t}{N} \end{aligned} \quad (5.8)$$

Se a transmissão ocorre, o par move-se para $I - I$. Agora se a remoção ocorre antes da transmissão, a díade move-se para o estado $R \cdot S$ ou $R - S$, dependendo de quando i e j são parceiros. Para determinar a probabilidade de transmissão de i para j , nós precisamos

calcular a probabilidade de absorção do estado $I - I$. Para realizar este cálculo nós usamos o método padrão para o processo de Markov. Para tanto, temos $A_{I.S}$ e A_{I-S} denotando as probabilidades de absorção do estado $I - I$, iniciando dos estados, respectivamente, $I \cdot S$ e $I - S$. Estas probabilidades satisfazem a equação,

$$\begin{aligned} \left(\gamma + \frac{\rho}{N}\right) A_{I.S} &= \frac{\rho}{N} A_{I-S} \\ (\beta + \gamma + \sigma) A_{I-S} &= \beta + \sigma A_{I.S}. \end{aligned} \quad (5.9)$$

Resolvendo o sistema de equações temos

$$\begin{aligned} A_{I.S} &= \frac{\beta\rho}{N \left[\left(\frac{\rho}{N} + \gamma\right) (\beta + \gamma + \sigma) - \frac{\rho\sigma}{N} \right]} \\ A_{I-S} &= \frac{\beta \left(\frac{\rho}{N} + \gamma\right)}{\left[\left(\frac{\rho}{N} + \gamma\right) (\beta + \gamma + \sigma) - \frac{\rho\sigma}{N} \right]}. \end{aligned} \quad (5.10)$$

Então a probabilidade de transmissão de i para j , no cenário em equilíbrio, é

$$p^{Eq} = A_{I.S} P_{I.S}(\infty) + A_{I-S} P_{I-S}(\infty), \quad (5.11)$$

dado que o grau médio é constante ($\langle k \rangle_\infty$) e os indivíduos idênticos. No outro cenário, fora do equilíbrio, a grau médio evolui no tempo. Desta maneira, para contabilizar a probabilidade de transmissão realizamos a média

$$p^{Nq} = \frac{1}{t_E} \int_0^{t_E} (A_{I.S} P_{I.S}(t) + A_{I-S} P_{I-S}(t)) dt \quad (5.12)$$

onde t_E é o tempo de extinção da epidemia ($I(t_E) = 0$).

Investigaremos no que segue o comportamento de t_E . Inicialmente consideramos $R_0^{Nq} > 1$, sobre esta condição a infecção se extinguirá devido a dois mecanismos sobre os indivíduos infectados: (i) remoção (probabilidade γ) e (ii) isolamento (probabilidade σ). Como os mecanismos são independentes, podemos afirmar que a extinção de um dado indivíduo infectado ocorre com uma probabilidade $\sigma\gamma$. Como os infectados são idênticos, podemos supor que t_E é inversamente proporcional a esta probabilidade. No entanto, perceba que se o infectado em questão foi apenas isolado e não removido, ele terá ρ novas conexões, em média, por passo de tempo. Além disso, no sentimento de superestimar t_E , podemos afirmar que o número de infectados em sequência, até cessar a infecção, é dado pelo número básico de reprodução do modelo padrão, $R_0 = \rho\beta/\sigma\gamma$. Feito estas considerações, afirmamos que

$$t_E \sim \frac{\rho^2\beta}{(\sigma\gamma)^2}. \quad (5.13)$$

Por fim, o número básico de reprodução para o cenário fora do equilíbrio, no limite de grandes populações, é dado por

$$R_0^{Nq} = \lim_{N \rightarrow \infty} N p^{Nq} \sim \frac{\rho\beta}{\sigma\gamma} \frac{\left(e^{-\frac{\rho\beta}{\sigma\gamma^2}} - 1 \right) \frac{\sigma\gamma^3}{\rho\beta} + (1 + \sigma)(\gamma + \sigma)}{\beta + \sigma + \gamma} \quad (5.14)$$

onde para tanto assumimos o tempo de extinção igual ao seu limite superior. De maneira similar, para o outro cenário dinâmico, encontramos

$$R_0^{Eq} = \lim_{N \rightarrow \infty} N p^{Eq} = \frac{\rho\beta}{\sigma\gamma} \frac{(1+\sigma)(\sigma+\gamma)}{\beta+\sigma+\gamma}. \quad (5.15)$$

5.2.2 Probabilidade da Epidemia

Em relação ao cenário fora de equilíbrio, para avaliar a extinção, contabilizamos em nossas simulações o número de realizações em que o número assintótico de removidos é superior a 5% do tamanho da população. Em verdade, esta caracterização é o complementar da probabilidade de extinção, o qual denominamos de probabilidade da epidemia. A figura ??A mostra a probabilidade da epidemia em função do número básico de reprodução do modelo padrão. Esta figura evidencia uma clara dependência com a relação entre as taxas σ e γ . Percebemos, por um lado, que se o tempo característico de propagação da infecção ($1/\gamma$) é muito superior ao tempo característico da infecção ($1/\sigma$) o resultado recai sobre a probabilidade de epidemia do modelo SIR padrão ($1 - 1/R_0$). Por outro lado, se a relação entre os tempos característico é o contrário, a probabilidade da epidemia vai a zero. Em verdade, um estudo detalhado destes resultados (tentativa e erro) nos permitiu elaborar uma forma funcional para o mesmo (linha sólida contínua na figura), ou seja,

$$P_{epi}^{Nq}(R_0) \simeq 1 - \left[\frac{1}{1 + \alpha (R_0 - 1)} \right]^{(R_0)^{-\eta}} \quad (5.16)$$

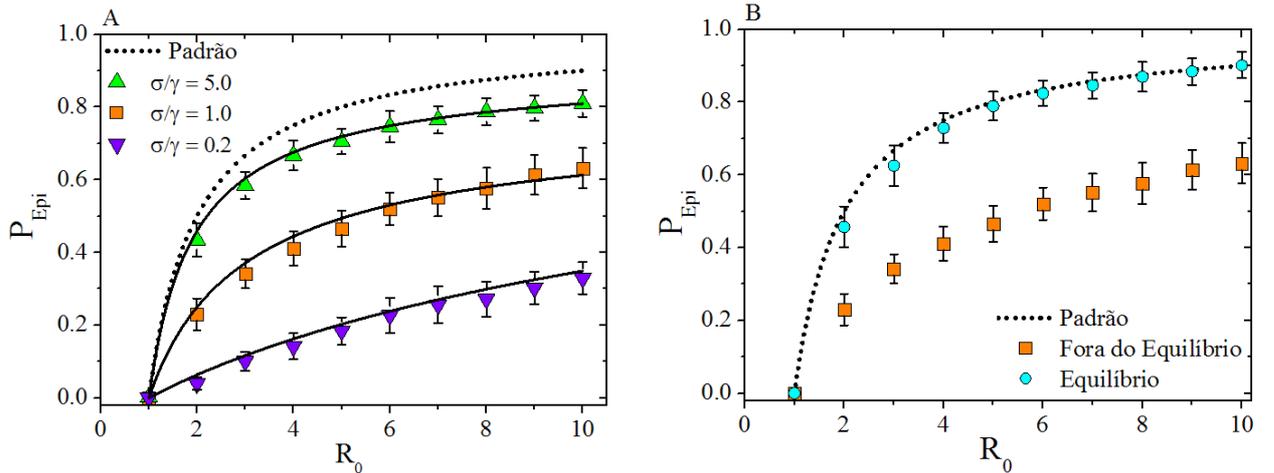


Figura 5.4: Probabilidade da epidemia (P_{epi}) versus o número básico de reprodução do modelo SIR padrão ($R_0 = \rho\beta/\sigma\gamma$). (A) Cenário fora do equilíbrio. (B) Cenário em equilíbrio dinâmico ($\sigma/\gamma = 1$). Os símbolos referem-se as simulações enquanto que as linhas sólidas a forma funcional estipulada. Resultados obtidos para $N = 10^4$ indivíduos. Cada ponto corresponde a uma média de 10^3 realizações.

onde

$$\alpha = \left(\frac{\sigma}{2\gamma} e^{-\frac{\sigma}{\gamma}} \right)^{e^{-\frac{\sigma}{\gamma}}} \left(1 + e^{-\frac{\sigma}{\gamma}} \right)^{-1}, \quad (5.17)$$

$$\eta = \frac{\sigma}{3\gamma} e^{-\frac{\sigma}{2\gamma}} \quad (5.18)$$

e $R_0 = \frac{\rho}{\sigma} \frac{\beta}{\gamma}$ (modelo SIR padrão). Note que a expressão está em acordo com nossas observações, isto é,

$$P_{epi}^{Nq}(R_0) \simeq \begin{cases} 1 - \frac{1}{R_0}, & \sigma \gg \gamma \\ 0, & \sigma \ll \gamma \end{cases} \quad (5.19)$$

De forma similar, concluímos que a probabilidade de epidemia do cenário em equilíbrio é equivalente ao do modelo SIR padrão,

$$P_{epi}^{Eq}(R_0) \simeq P_{epi}^{St}(R_0) = 1 - \frac{1}{R_0} \quad (5.20)$$

veja figura 5.4B. Ainda em relação a esta figura, comparamos as situações fora de equilíbrio e equilíbrio. Fica evidente, para a mesma relação entre σ e γ , que uma epidemia no cenário em equilíbrio dinâmico é mais provável.

5.2.3 Dinâmica da Doença

No capítulo anterior, vimos que a dinâmica dos removidos para o modelo SIR padrão é solução da equação

$$\frac{dR}{d\tau} = 1 - R - (1 - i_0)e^{-R_0 R}. \quad (5.21)$$

onde $\tau = \gamma t$. No contexto desta formulação determinística, podemos observar que a fração de removidos R , no limite de $i_0 \rightarrow 0$, evolui exclusivamente como função do número básico de reprodução R_0 .

Diferente da versão determinística (modelo SIR padrão), o problema que propomos trata de um modelo estocástico. Sendo assim, para observar uma epidemia neste cenário, não basta estar acima do limiar ($R_0 > 1$) é necessário também que haja uma probabilidade não nula para o evento em questão. No entanto, uma vez que a probabilidade da epidemia é diferente de zero uma bifurcação na solução é observada. Uma fração das realizações vão para um estado assintótico ($R(\infty)$) bem definido, enquanto que as demais realizações resultam em $R(\tau) \sim 0$. Baseado nesta observação, consideramos que a média sobre as realizações em que resultam numa epidemia (média pesada) deve evoluir de maneira similar ao que ocorre no modelo determinístico. Para tanto, dado que conhecemos como evoluem as gerações dos infectados, em média, propomos a seguinte substituição

$$R_0 \rightarrow \begin{cases} R_0^{Eq}, & \text{Equilíbrio} \\ R_0^{Nq} \left(1 - e^{-\frac{\sigma}{\gamma}\tau} \right), & \text{Fora do equilíbrio.} \end{cases} \quad (5.22)$$

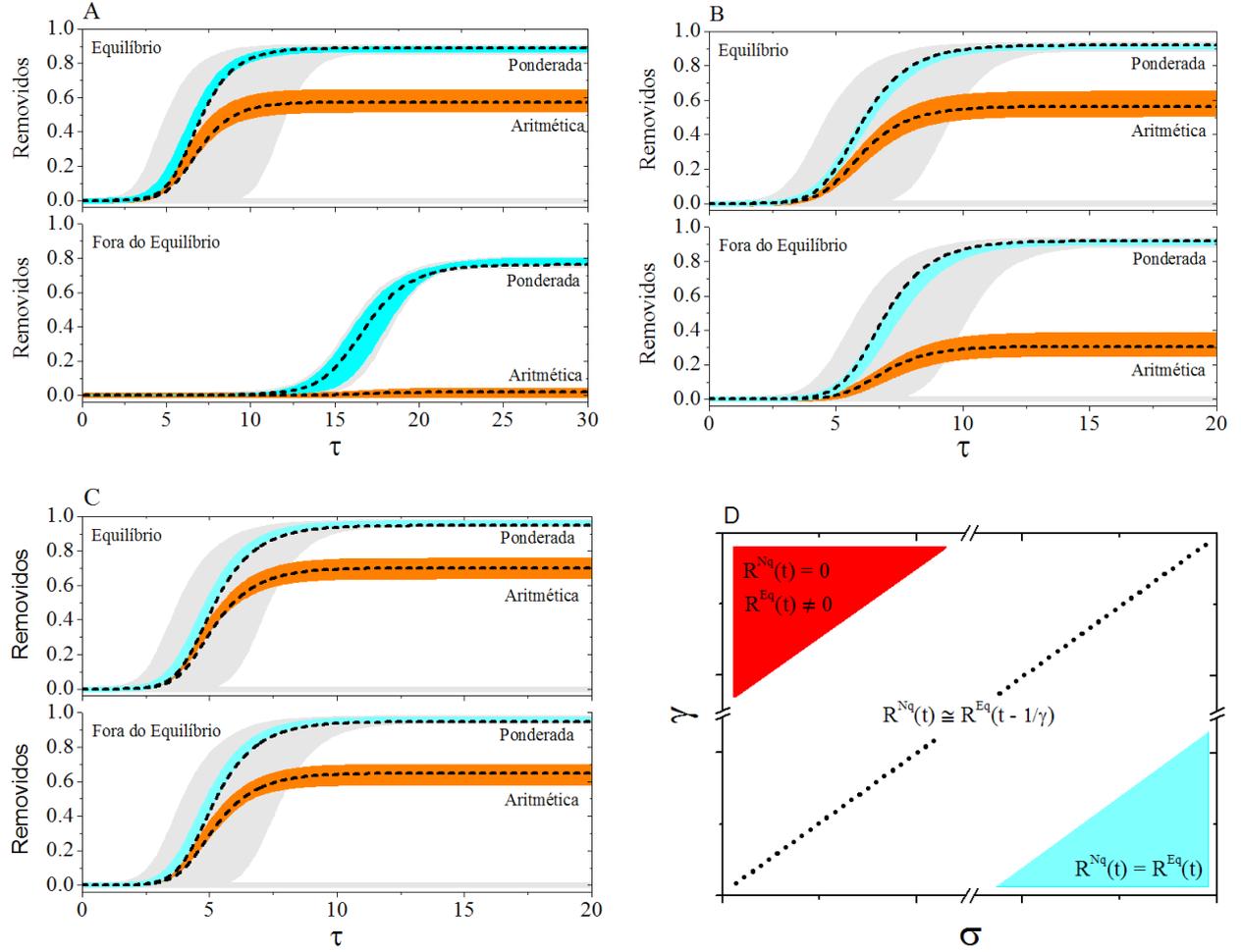


Figura 5.5: Evolução temporal dos removidos. Resultados para (A) $\sigma/\gamma = 0.1$ e $\beta = 0.3$ ($R_0^{Nq} = 1.91$ e $R_0^{Eq} = 2.59$), (B) $\sigma/\gamma = 1.0$ e $\beta = 0.03$ ($R_0^{Nq} = 2.83$ e $R_0^{Eq} = 2.86$) e (C) $\sigma/\gamma = 10$ e $\beta = 0.003$ ($R_0^{Nq} = 3.21$ e $R_0^{Eq} = 3.21$). (D) Representação pictórica do comportamento da dinâmica dos removidos frente a relação entre γ e σ . Resultados obtidos para $N = 10^4$ e $\rho = 1$ sobre 10^3 realizações independentes em 10^2 grafos diferentes (fixo ρ e σ). Cada curva na região sombreada (cinza) refere-se a uma realização. As curvas denominadas de ponderada ou aritmética são uma média de 10^3 realizações independentes fixa um grafo. A ponderada trata-se de uma média sobre as realizações em que $R(\infty) > 0.05$ enquanto que a aritmética é uma média sobre todas as realizações. As linhas pontilhadas referem-se a solução numérica das equações (5.23) e (5.24).

Assim, a dinâmica dos removidos no modelo estocástico possui uma aproximação determinística, dada por

$$\begin{aligned} \frac{dR^{Eq}}{d\tau} &= 1 - R^{Eq} - (1 - i_0)e^{-R_0^{Eq}R^{Eq}} \\ \frac{dR^{Nq}}{d\tau} &= 1 - R^{Nq} - (1 - i_0)e^{-R_0^{Nq}\left(1 - e^{-\frac{\sigma}{\gamma}\tau}\right)R^{Nq}}. \end{aligned} \quad (5.23)$$

Além disso, usando da probabilidade da epidemia, podemos determinar a média sobre todas

as realizações (média aritmética) como

$$\langle R^x(\tau) \rangle = P_{epi}^x R^x(\tau) \quad (5.24)$$

onde o índice x refere-se aos cenários fora do equilíbrio (Nq) e equilíbrio (Eq).

As figuras 5.5(A-C) apresentam uma boa correspondência entre a aproximação determinística e a médias sobre as inúmeras realizações para diferentes relações entre σ e γ . Além disso, comparando os cenários, fora do equilíbrio e equilíbrio, para as diferentes relações entre σ e γ observamos o que segue:

- $\sigma \ll \gamma$

Sobre está relação, para o cenário fora de equilíbrio os indivíduos estão completamente isolados. Portanto,

$$R_0^{Nq} = 0 \quad (5.25)$$

Sendo assim, não há epidemia neste cenário, $R^{Nq}(\infty) = 0$. Já para o cenário de equilíbrio dinâmico, resulta

$$R_0^{Eq} = \frac{\rho\beta}{\sigma(\beta + \gamma)}. \quad (5.26)$$

Logo, surge uma epidemia de tamanho

$$1 - R^{Eq}(\infty) = (1 - i_0)e^{-R_0^{Eq}R^{Eq}(\infty)}. \quad (5.27)$$

- $\sigma \sim \gamma$

Desta condição, temos que

$$R_0^{Nq} \simeq R_0^{Eq} = \frac{\rho\beta(1 + \sigma)(\sigma + \gamma)}{\sigma\gamma(\beta + \sigma + \gamma)} \quad (5.28)$$

dado que

$$\left(e^{-\frac{\rho^2\beta}{\sigma\gamma^2}} - 1 \right) \frac{\sigma\gamma^3}{\rho^2\beta} \sim 0. \quad (5.29)$$

Adicionalmente, no contexto da dependência com a dinâmica de interação, no cenário fora de equilíbrio, obtemos

$$\langle k \rangle_\tau \simeq \langle k \rangle_\infty (1 - e^{-\tau}). \quad (5.30)$$

Podemos verificar deste resultado que para poucos passos de tempo o número médio de contatos satura, $\langle k \rangle_\tau \simeq \langle k \rangle_\infty$. Estas observações são suficientes para supormos que o comportamento da dinâmica da epidemia entres os cenários fora de equilíbrio e equilíbrio sejam similares a menos de um tempo de retardo ϕ . Para tanto, escrevemos que

$$R^{Nq}(\tau) \simeq R^{Eq}(\tau - \phi). \quad (5.31)$$

Numericamente, nós obtemos $\phi = 1$ (relembramos que $\tau = \gamma t$). Sendo assim, concluímos que $R^{Nq}(t) \simeq R^{Eq}(t - 1/\gamma)$, veja figura 5.5B.

- $\sigma \gg \gamma$

Dada esta condição, concluímos que

$$R_0^{Nq} = R_0^{Eq} = \frac{\rho\beta(1 + \sigma)}{\gamma(\beta + \sigma)} \quad (5.32)$$

e

$$1 - e^{-\frac{\sigma}{\gamma}\tau} \rightarrow 1. \quad (5.33)$$

Com efeito, a dinâmica da epidemia é igual em ambos os cenários, $R^{Nq}(\tau) = R^{Eq}(\tau)$, veja figura 5.5C.

A figura 5.5D sumariza estes resultados por intermédio de um diagrama γ versus σ .

Capítulo 6

Dinâmica de Epidemias em uma Rede Social Real

Como discutimos no capítulo 2 as redes complexas, têm sido recentemente estudadas no contexto da transmissão de doenças. 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 artificiais (*small-world*, *scale-free*, etc), 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 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 artificiais. Os resultados que apresentaremos são todos dados em termos de simulações computacionais. Contudo, nas seções que antecedem a estes resultados, fazemos referência do modelo computacional que estamos utilizando.

6.1 Redes

A fim de fazer previsão significativa sobre a propagação de doenças, é hoje reconhecido que o conhecimento da estrutura das redes é fundamental. A descrição acima é válida para as doenças em geral, mas é particularmente importante para as doenças sexualmente transmissíveis (DST). Até recentemente, as informações sobre estruturas da rede sexual eram praticamente nulas, fato que fez Barabasi [3] afirmar que *obter um mapa da rede sexual seria*

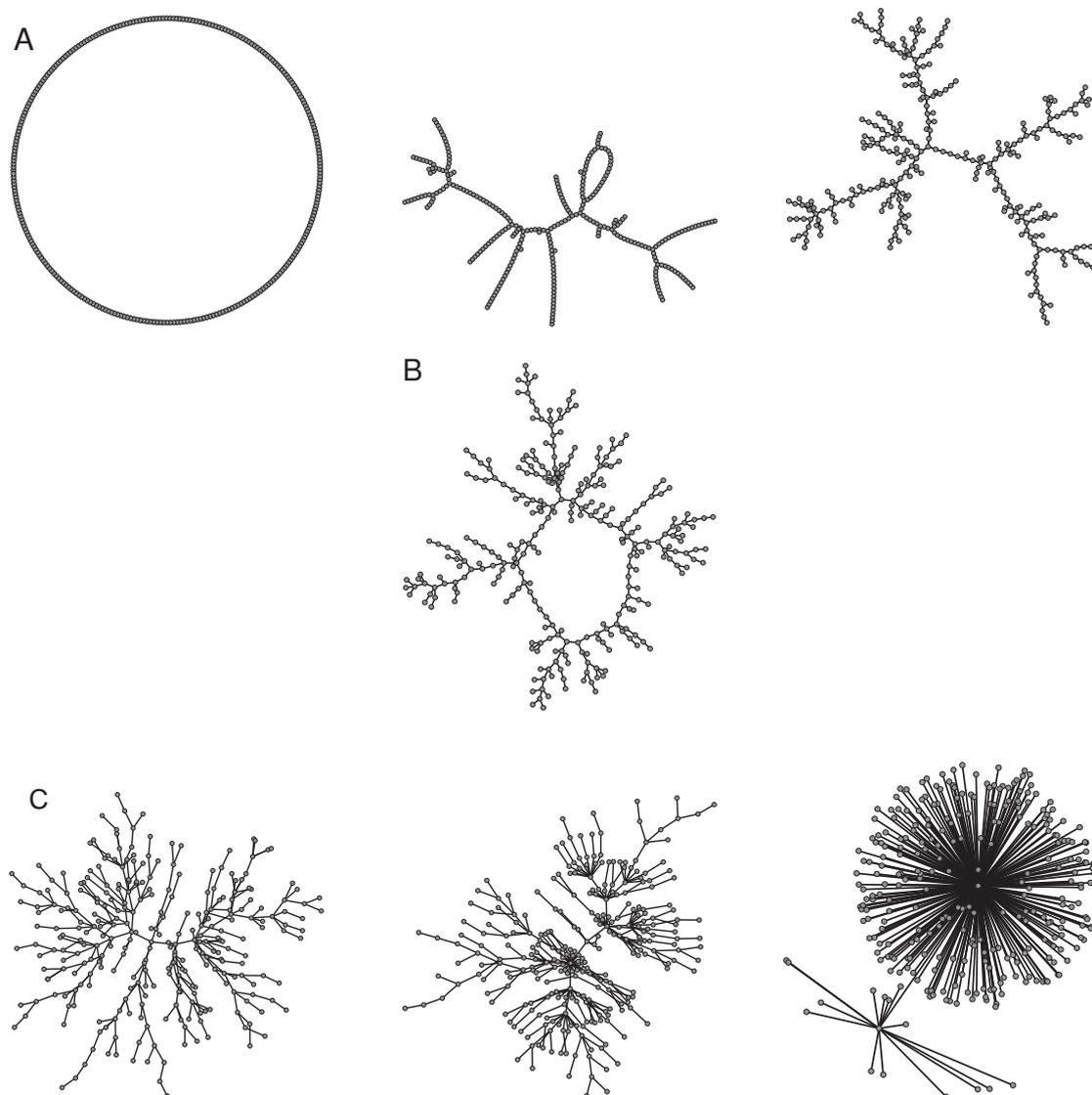


Figura 6.1: Redes utilizadas no trabalho: (A) Watts-Strogatz: da esquerda para direita, $p = 0$ (anel), $p = 0.1$ e $p = 1$ (aleatória); (B) Romântica; (C) Krapivsky-Redner: da esquerda para direita, $r = 0.1$, $r = 0.5$ e $r = 1$. Todas tem o mesmo número de vértice, $N = 288$, e o mesmo grau médio, $\langle k \rangle = 2$. Figuras geradas com PAJEK [28].

quase impossível. No entanto, em 2004 a primeira imagem de uma rede sexual, quase completa, foi publicada (Bearman et al [29]). Esta rede, chamada pelo autores “rede romântica e sexual” surgiu a partir de dados extraídos *National Longitudinal Study of Adolescent Health*. Os dados foram obtidos após entrevistas domiciliares, abrangendo um período de 18 meses, onde alunos foram convidados a identificar os seus parceiros sexuais ou românticas em uma lista dos alunos que frequentam suas escola. “*Jefferson High School*” é o nome falso de uma escola de ensino médio em “*Jefferson City*”, uma cidade do centro-oeste dos Estados Unidos, como uma homogênea e isolada comunidade. Participaram da pesquisa 90% dos alunos, um

total de 873 completaram o questionário. Depois dos dados coletados, uma rede de 573 vértices surgiu, a partir do qual a característica mais notável é uma componente gigante envolvendo 288 alunos, o que é mostrado na figura 6.1B. A componente gigante é resultado da superposição de todos os contatos entre os alunos durante o período de 18 meses.

Várias topologias de redes, construídas por algoritmos computacionais, redes artificiais, tem sido exploradas no contexto da transmissão de doenças na literatura. 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 DST entre as redes artificiais e real. Abordaremos, no que segue, dois tipos de redes artificiais, a qual denominamos de aleatórias e lei de potência. As redes aleatórias, figura 6.1A, são implementadas de acordo com a construção de Watts-Strogatz (WS), para detalhes veja a subseção 2.3.2. Para as redes lei de potência, figura 6.1C, usamos o modelo de redirecionamento de Krapivsky e Redner [30] (KR). Na construção de Krapivsky-Redner, em cada passo de tempo um novo vértice é adicionado ao sistema e um vértice i , adicionado nos passos anteriores, é selecionado aleatoriamente como uma possível ligação. Com probabilidade $1 - r$ uma aresta é criada entre o novo vértice e o vértice i ; porém, com probabilidade r a aresta é redirecionada para vértice antecessor j do vértice i . Este modelo conduz a uma distribuição com lei de potência com expoente $\gamma = 1 + \frac{1}{r}$, que pode ser ajustado para qualquer valor maior que 2. Na subseção 2.3.3 apresentamos alguns resultados para o caso particular $r = 0.5$. Para a rede romântica, mapeamos manualmente a representação gráfica da componente gigante publicada no artigo do Bearman e comparamos com Pajek [28] que nossa transcrição da a mesma representação gráfica que a original. Todas as redes, artificiais e real, são bipartidas, com o mesmo número de indivíduos de sexo masculino (145) e feminino (143). Além disso, não consideramos ligações entre indivíduos do mesmo sexo.¹

6.2 Simulações Computacionais

As epidemias nas redes são estudadas por meio de simulações computacionais do modelo SIR. Para cada rede específica e realização — na qual cada vértice representa um indivíduo (bipartido: masculino ou feminino) e com arestas que representam a interação romântica (relação sexual) — uma pessoa é escolhida aleatoriamente para ser o primeiro infectado, deixando o resto da população no estado suscetível. Este é o estado inicial: um vértice no estado I e os outros 287 vértices no estado S . A simulação desenrola da

¹Na verdade, a rede romântica tem três ligações homossexuais. Dado o pequeno número de contatos entre indivíduos do mesmo sexo, decidimos considerar um rede heterossexual. Do ponto de vista das redes, significa que o coeficiente de clusterização é nulo.

seguinte forma: em cada passo de tempo (que pode representar um dia ou uma semana), cada indivíduo escolhe um parceiro a partir da lista de indivíduos conectados com ele/ela. Das três possíveis situações de interação do par, ou seja, ambos suscetível, ambos infectados e apenas um infectado; o último é o único que pode alterar o estado do par. Se este for o caso, o indivíduo suscetível vai alterar o estado com probabilidade (por unidade de tempo) β , esse processo é semelhante a relação sexual de dois indivíduos em que um deles tem uma DST. Um passo de tempo da simulação é concluída depois que todos os vértices tiveram a oportunidade de contactar um dos seus parceiros. Assim, supomos que a atividade sexual é homogênea entre a população (o que não é necessariamente verdade) e independe do número de parceiros. Iremos discutir outras opções mais tarde. Uma vez infectado, o indivíduo permanece neste estado por um período fixo τ , sendo potencialmente contagioso a sua vizinhança neste intervalo de tempo. Após esse período, o indivíduo infectado é removido (ou se torna permanente imune). A evolução progride até que nenhum indivíduo infectado permaneça na população. Para uma dada situação, os parâmetros da doença β e τ são constantes e têm o mesmo valor para cada indivíduo da população.

Dado este esquema, a propagação de uma DST hipotética pode ser estudado nas redes apresentadas anteriormente (redes romântica e artificiais). O passo de tempo, como mencionamos antes, pode representar um dia, uma semana ou outro período de tempo. Se for um dia, significa que cada indivíduo tem pelo menos uma relação sexual por dia. No entanto, esse período de tempo poderia ser de uma semana que pode subestimar o perigo da epidemia. De qualquer maneira, não é nosso objetivo calibrar de maneira precisa os parâmetros de uma doença específica ou fazer uma previsão de um surto real, mas a comparação dos resultados de uma topologia para outra.

6.3 Resultados

Apresentamos, no que segue, os resultados de uma DST hipotética nas diferentes redes. Na primeira subseção, apresentamos os resultados em termos da dinâmica da doença: máximo número de infectado I_{max} , tempo do pico de infecção t_p , tamanho final da epidemia, $R(\infty)$ (veja figura 6.2) e probabilidade da epidemia. Todas as medidas foram obtidas pelo cálculo das realizações que terminaram com mais de 5% de indivíduos removidos. Como efeito, é a razão entre o número dessas realizações e o total que usamos para definir a probabilidade de observar uma epidemia. Ainda em relação a rede romântica, estudamos o comportamento em resposta a três variantes. O primeiro tem a ver com o efeito de algum indivíduo, pertencente a rede, ter um contato com indivíduo infectado externo. Realizamos isto por meio de um parâmetro que denominado de “campo externo”. Este parâmetro re-

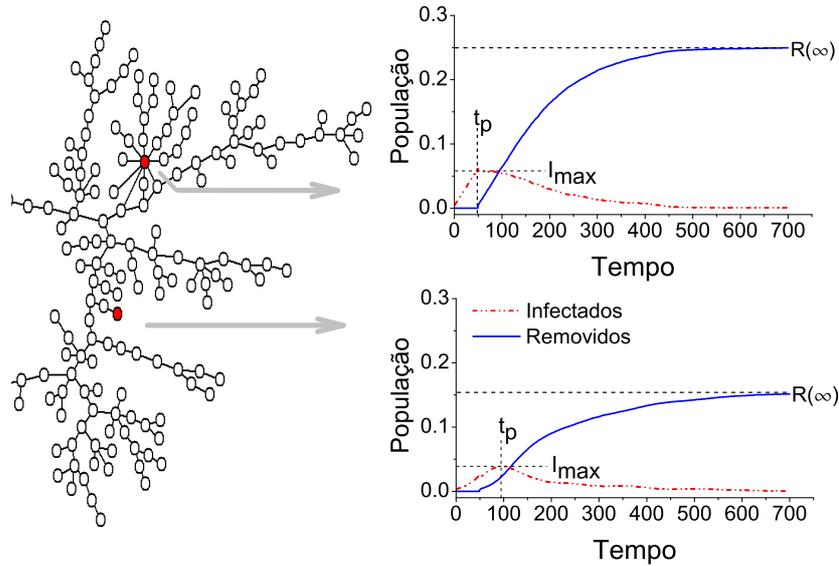


Figura 6.2: Efeito do infectado inicial, vértice vermelho, no tamanho final da epidemia $R(\infty)$, máximo número de infectados I_{max} e tempo do pico de infecção t_p .

presenta a probabilidade de um indivíduo escolhido ao acaso interagir com outro indivíduo infectado externo. Observe que isto é equivalente a probabilidade de um indivíduo ser espontaneamente infectado. A segunda característica que exploraremos é o que chamamos de “promiscuidade”. Ela está relacionada com o número de parceiros do indivíduo inicialmente infectado. Por fim, investigamos a influência da variabilidade dos contatos por unidade de tempo. Denominamos estas características de “regras de interação”. Na segunda subseção, mostramos os correspondentes resultados quando consideramos a formação dinâmica dos contatos.

6.3.1 Rede Estática

Rede Romântica vs Redes Artificiais

Os autores da “Jefferson school” analisaram muitas medidas típicas de redes na rede romântica e concluíram que ela não pode ser classificada como qualquer rede artificial. Não é completamente aleatória, *small-world* ou *scale-free*. Em vez disso, ela pode emprestar algumas características de todas elas. Nesta etapa do trabalho estamos interessados na comparação da rede romântica com as redes artificiais a partir da perspectiva da dinâmica da epidemia. São estas peculiares diferenças entre as redes romântica e artificiais, que nós sabemos da análise da rede, o suficiente para produzir diferenças significativas na propagação de doenças? Ou existe uma rede artificial que apresenta resultados da epidemia que são

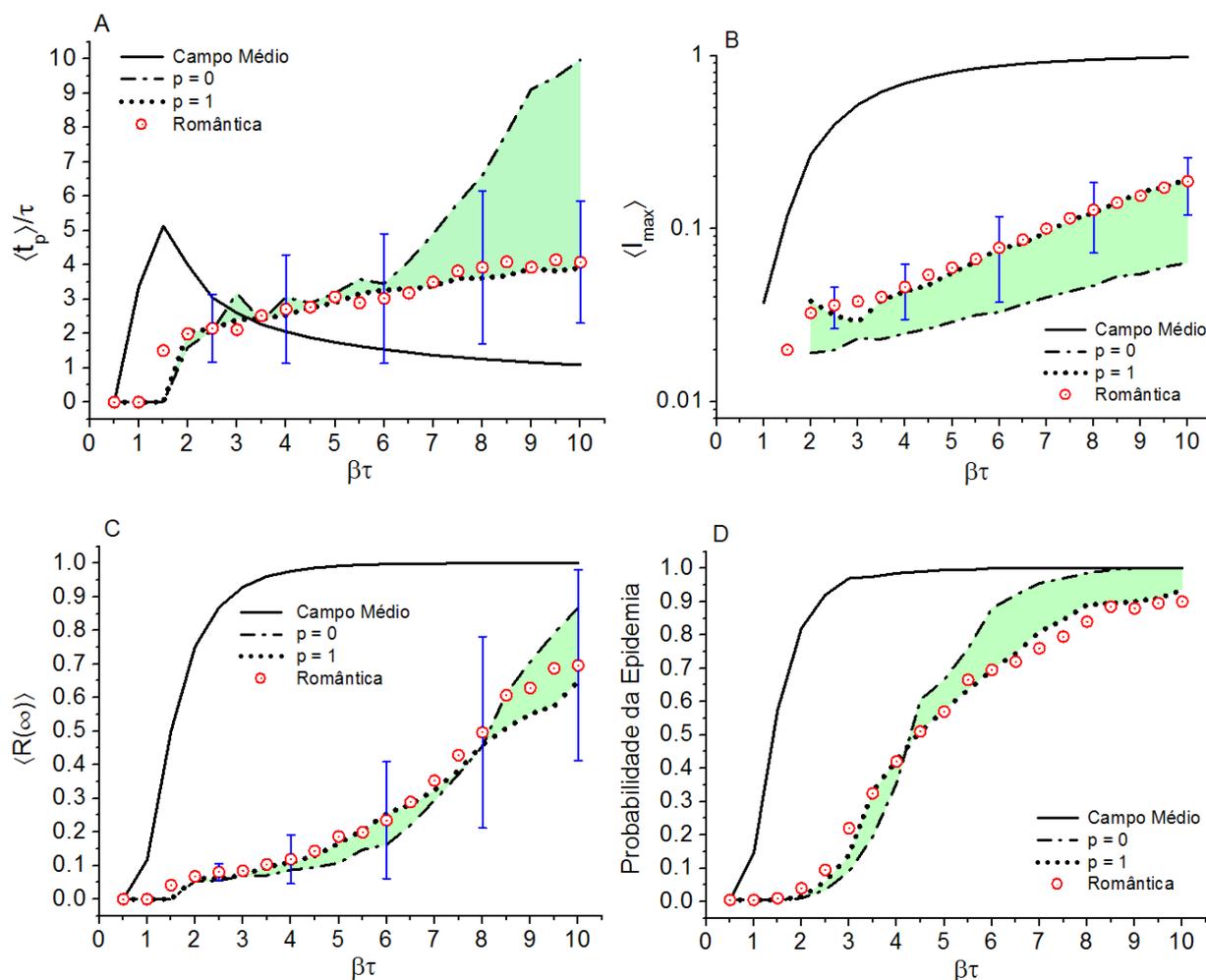


Figura 6.3: Simulação da epidemia em diferentes topologias: Watts-Strogatz, Romântica e campo médio. Resultados para (A) tempo do pico da epidemia, (B) número máximo de infectados, (C) número total de removidos e (D) probabilidade da epidemia. Cada ponto das curvas é o resultado de uma média sobre 200 realizações independentes em 10 grafos diferentes (exceto para a rede romântica que é somente um grafo). Barra de erros, exibidas somente em alguns pontos, correspondem ao desvio padrão sobre as 200 realizações no caso da rede Romântica. A região sombreada ajuda a visualizar a variação dos resultados devido a desordem progressiva nas redes W-S.

essencialmente equivalentes ao que poderia ter acontecido na rede romântica? Esta é a questão central desta contribuição, porque ela nos diz se o modelo em questão é razoável para reproduzir uma rede real. Mais audaciosamente ela pode nos dar uma base para decidir se o modelo se aproxima do problema real.

Em uma perspectiva geral, nós observamos que, exceto para a rede KR com $r \approx 1$ (a rede mais heterogênea), muitas topologias resultam em epidemias similares desde que $\beta\tau < 4$. Isto pode ser concluído da comparação entre as redes Românticas e WS, como pode ser visto na Fig.6.3, e a KR para $r < 1$, como mostra a Fig.???. Além disso, se nós

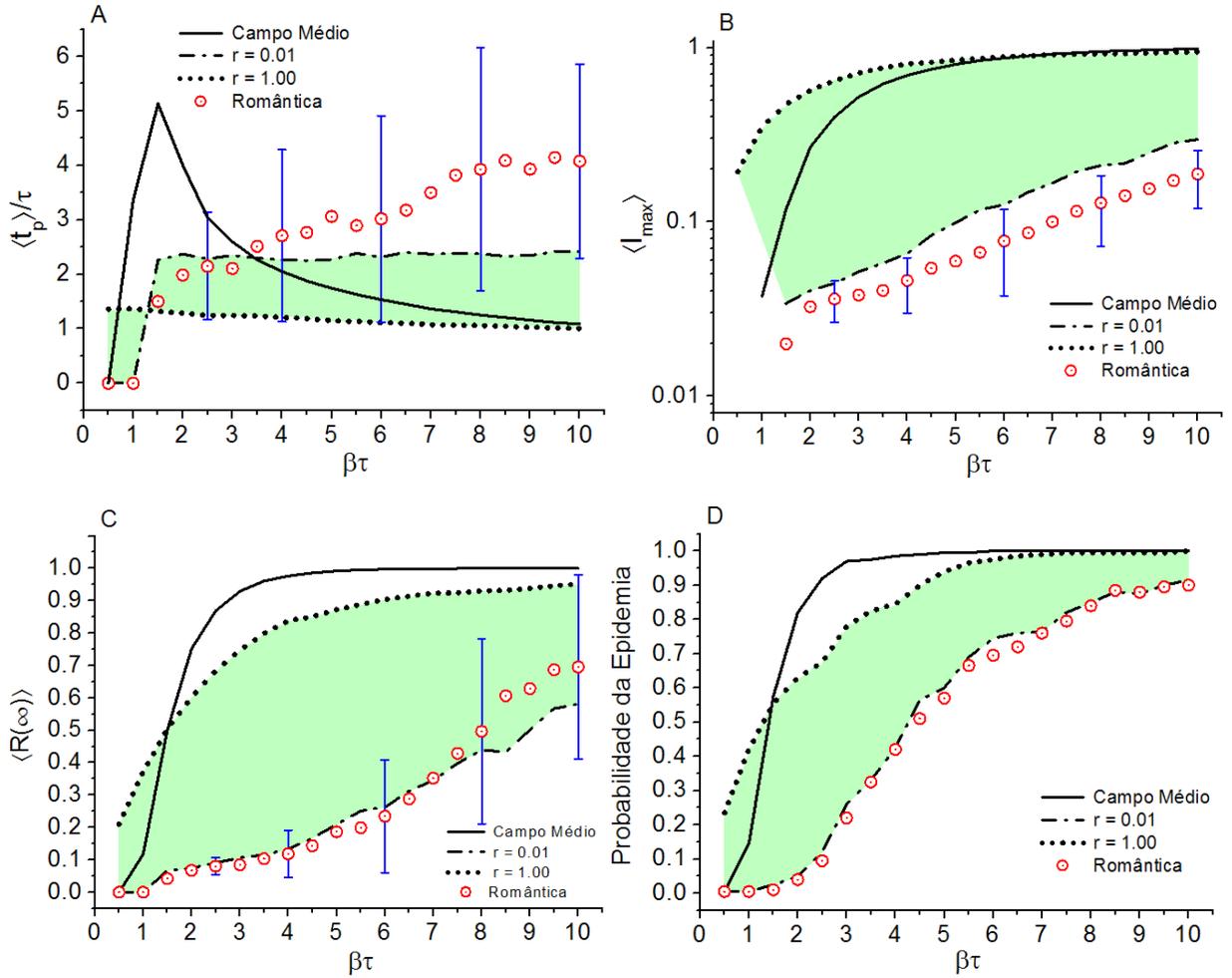


Figura 6.4: Epidemia em diferentes topologias: Krapivsky-Redner, Romântica e campo médio. Resultados para (A) tempo do pico da epidemia, (B) número máximo de infectados, (C) número total de removidos e (D) probabilidade da epidemia. Número de realizações, média por ponto, e barra de erros como na Fig. 6.3. A região sombreada ajuda a visualizar a variação dos resultados devido a heterogeneidade progressiva nas redes K-R.

consideramos um pequeno grupo de topologias, a WS como grande p (≈ 1) e a KR como pequeno r (≈ 0), o comportamento dos resultados, especialmente $R(\infty)$ e a probabilidade da epidemia (Fig. 6.3, 6.4C, D) são similares a rede Romântica, $\beta\tau < 8$. Considerando realizações individuais, as barras de erro exibidas nas Fig. 6.3A-C and Fig. 6.4A-C correspondem a rede Romântica e dão uma ideia da dispersão observada associada ao pequeno tamanho dela.

Além disso, nós observamos que o cenário de campo médio é o pior possível no ponto de vista da epidemia. Qualquer outra situação, ou seja, qualquer outra estrutura de rede em que nem todos vértices são potencialmente conectados, produz um quadro menos desfavorável. A exceção é a rede KR com $r = 1$ (uma rede “*scale-free*” para um sistema infinito) mas para valores de $\beta\tau < 1$.

Comparando os resultados entre as diferentes redes, modelos e Romântica, é possível ver que nenhum dos modelos propostos fita exatamente a rede real. No entanto, dentro das incertezas das flutuações associadas com as rede pequenas (288 vértices) consideradas aqui, uma epidemia na rede aleatória obtida com $p \approx 1$ é estatisticamente indistinguível para uma epidemia na rede Romântica. Por outro lado, a topologia que diverge mais da rede Romântica (e de todas as outras) é a rede KR com $r = 1$. Este particular valor resulta em uma rede como uma distribuição de grau dada por uma lei de potência com expoente $\gamma = 2$ (*scale-free*) no limite de grandes tamanhos. No presente caso, trata-se de uma rede pequena e resulta em uma topologia de estrela, muito diferente da topologia da rede romântica.

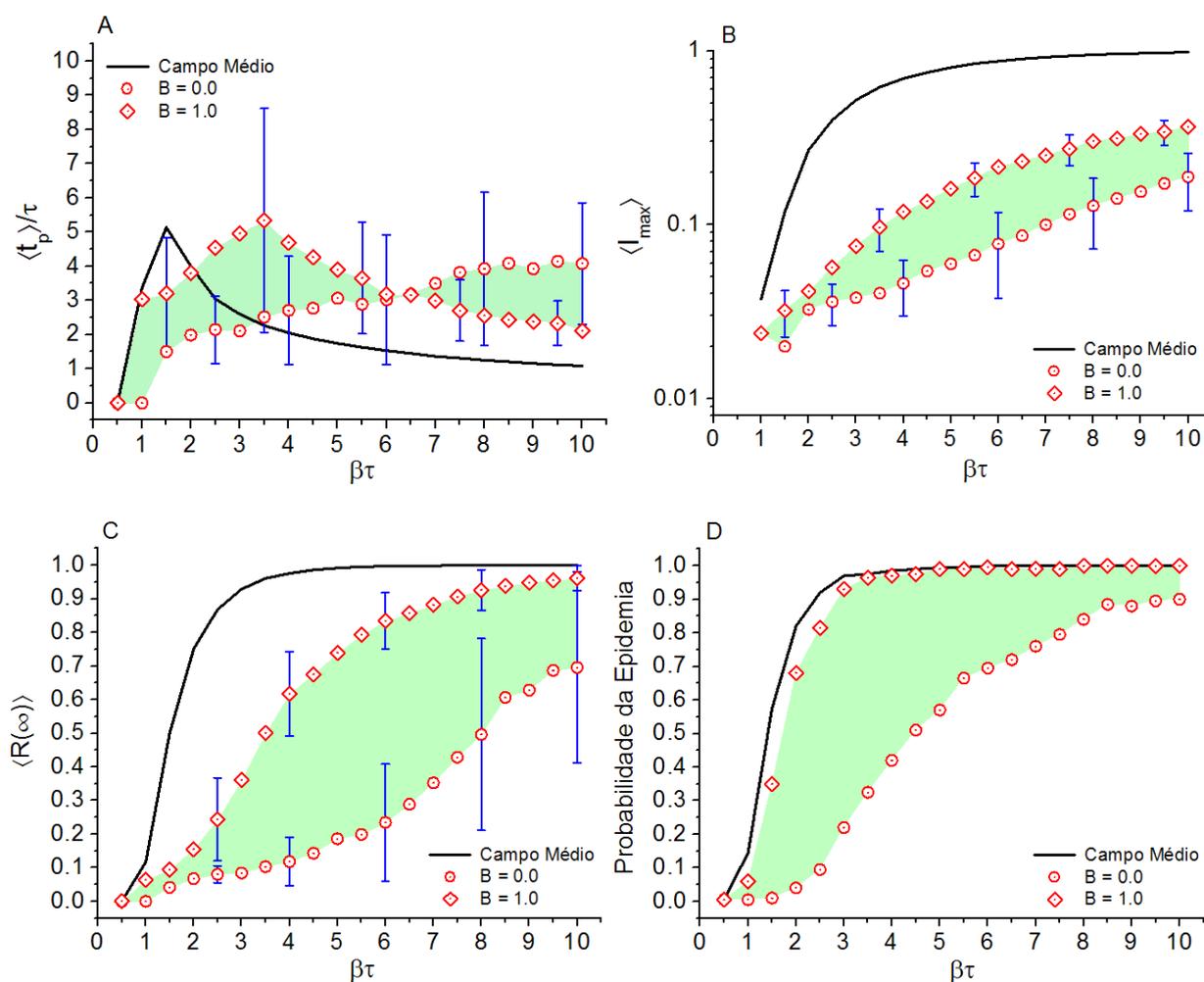


Figura 6.5: Efeito do “campo externo”, B , aplicado na rede Romântica. Resultados para (A) tempo do pico da epidemia, (B) número máximo de infectados, (C) número total de removidos e (D) probabilidade da epidemia. Número de realizações, média por ponto, e barra de erros como na Fig. 6.3. A região sombreada ajuda a visualizar a variação dos resultados com o aumento progressivo do parâmetro B .

Campo Externo

Mostramos, no que segue, os resultados das propriedades epidêmicas quando um “campo externo” é aplicado. O campo externo B representa a probabilidade de algum sujeito dentro da rede romântica interagir com alguém fora dela, que supomos já infectado. Então tomamos um indivíduo, escolhido aleatoriamente dentro da população, se torna infectado com probabilidade β como usual. Obviamente, $B = 0$ é equivalente a dinâmica normal sem campo externo, enquanto que $B = 1$ significa que a cada tempo alguém encontra um infectado. Portanto, podemos dizer que $B\beta$ é a probabilidade de uma espontânea infecção ocorrer na rede romântica.

Como podemos ver na figura 6.5, o efeito do campo externo faz o surto da infecção — número total de indivíduos infectados como função de $B\beta$ (curva de infecção) — próximo do que é esperado para o modelo SIR a campo médio. Isto sugere que o efeito prático do campo externo é fazer que a rede se aproxime da topologia em que todos os vértices estão conectados. Outra interpretação deste efeito é o seguinte: B é a probabilidade de fazer uma temporária conexão entre um indivíduo aleatoriamente escolhido da rede e um infectado, que dentro de uma certa perspectiva pode ser um sujeito da própria rede. Isto não é equivalente a ter todos os indivíduos conectados, mas a rede romântica aproxima deste caso.

Promiscuidade

Passamos agora para a influência que o infectado inicial tem nas propriedades epidêmicas (I_{max} , t_p e $R(\infty)$). Em particular, temos interesse em ver de que modo o grau do vértice do sujeito inicialmente infectado influencia na epidemia. A razão disto é que alguém que tem muitas ligações deverá ter maior potencial de infecção que alguém, por exemplo, que tenha somente uma ligação. Em ordem para testar esta influência, definimos os indivíduos inicialmente infectados de acordo com o grau e , conseqüentemente, do mínimo valor $k_i = 1$ ao máximo valor $k_i = 9$ existentes na rede romântica. Quando $k_i = 9$ existe somente uma opção para o infectado inicial. Dado que somente um sujeito na rede romântica tem este grau, consideramos este caso como promiscuidade máxima. Explorando outros valores de k_i , temos mais opções para o indivíduo inicialmente infectado. Devido ao fato que a distribuição de grau sobe. Neste casos, realizaremos uma média sobre as opções iniciais. Em qualquer caso, realizaremos uma média sobre muitas realizações.

Os resultados, apresentados na Fig. 6.6, são um pouco sutil para interpretar. Olhando a Fig. 6.6B,C (I_{max} e $R(\infty)$, respectivamente), nós somos tentados a concluir que a influência de quem é o infectado inicial, se é mais ou menos conectado, é irrelevante – o qual admitimos não ser óbvio. Porém, as outras duas figuras Fig. 6.6A,D), representando o tempo do pico

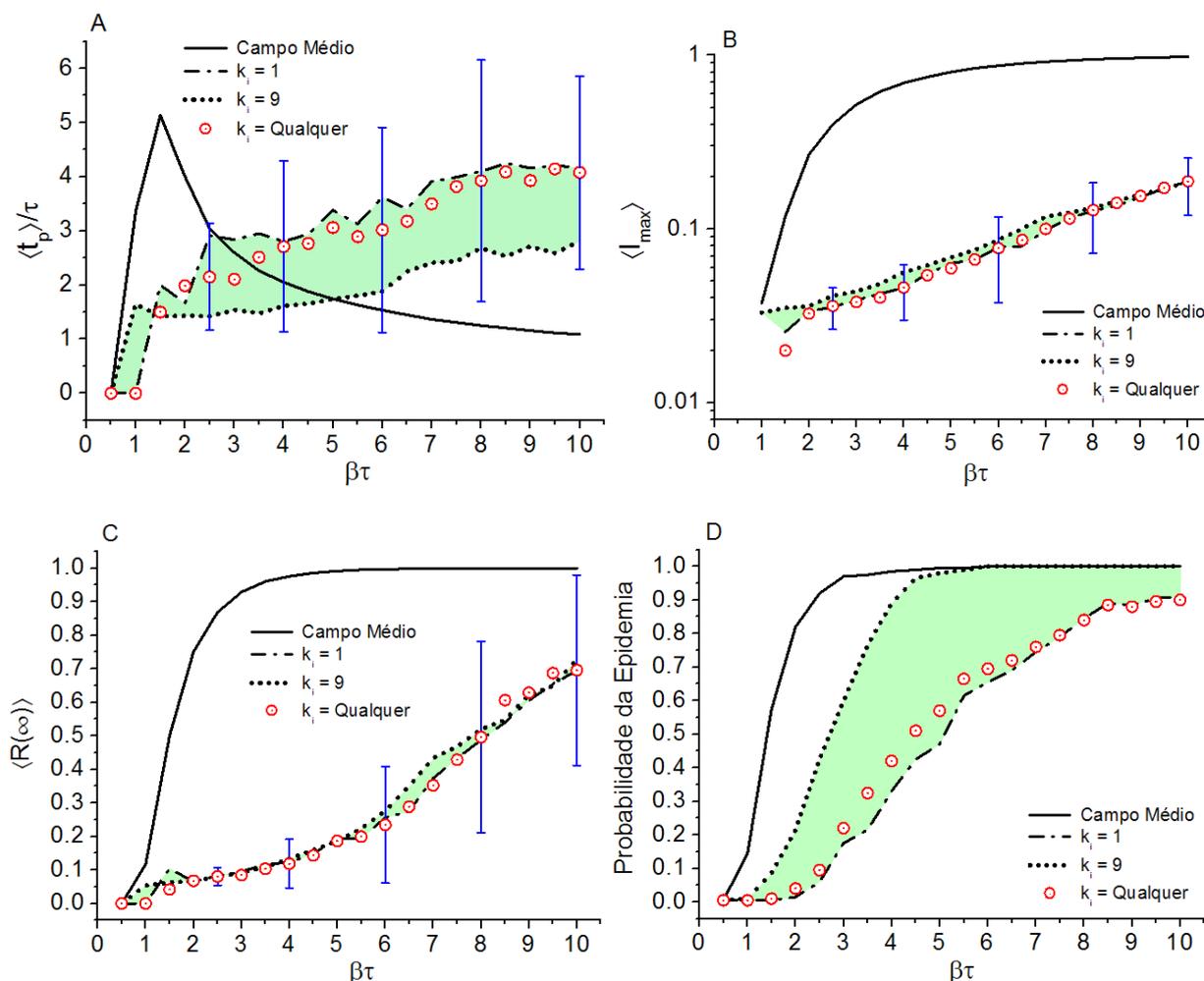


Figura 6.6: Efeito do grau do sujeito escolhido como infectado inicial na rede Romântica. Resultados para (A) tempo do pico da epidemia, (B) número máximo de infectados, (C) número total de removidos e (D) probabilidade da epidemia. Número de realizações, média por ponto, e barra de erros como na Fig. 6.3. A região sombreada ajuda a visualizar a variação dos resultados como resposta a variação do grau do indivíduo inicialmente infectado.

de infecção e a probabilidade de epidemia, mostram uma clara divergência de baixo ($k_i = 1$) para alto ($k_i = 9$) grau do indivíduo inicialmente infectado. Em $\beta\tau = 4$ por exemplo, nós claramente vemos na Fig. 6.6C que, independente de quem estava inicialmente infectado, o número médio de sujeitos removidos no final é em torno de 10%. Apesar disso, Fig. 6.6D indica que no caso de baixo grau a epidemia ocorre apenas com probabilidade de 30%, contrastando com os 90% observados no caso de alto grau. Portanto, como esperado, o grau do indivíduo infectado inicialmente é um fator chave no resultado da propagação da doença na rede Romântica.

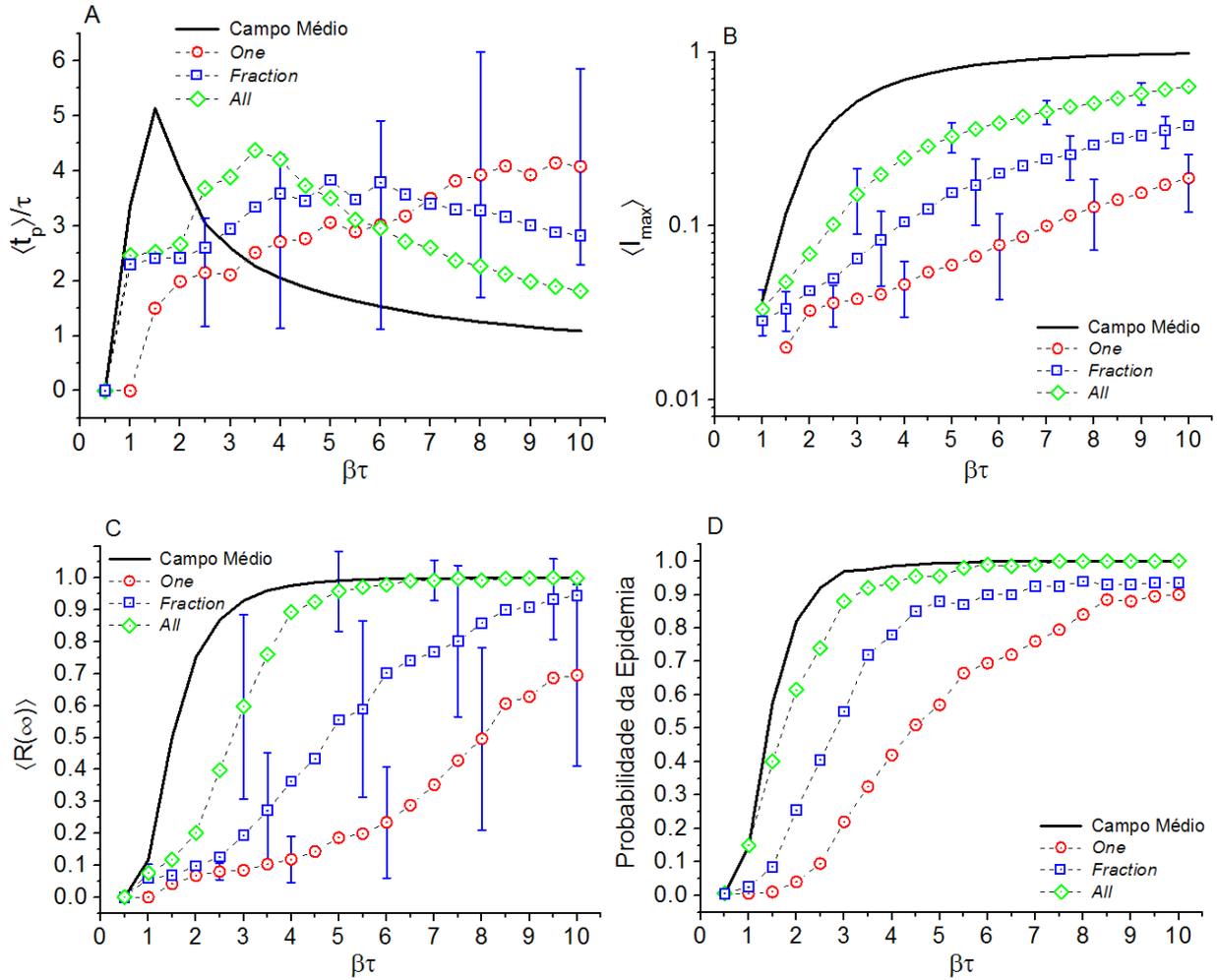


Figura 6.7: Efeito das diferentes regras de interação. Resultados para (A) tempo do pico da epidemia, (B) número máximo de infectados, (C) número total de removidos e (D) probabilidade da epidemia. Número de realizações, média por ponto, e barra de erros como na Fig. 6.3.

Regras

Antes de ir para a rede dinâmica faremos menção a influência da variabilidade das interações, o qual denominamos de regras de transmissão: *one*, *fraction* e *all*. Na real rede dinâmica a dinâmica da frequência das interações não é distribuída uniformemente, ou seja, um indivíduo pode visitar um ou mais de seus contatos durante um período fixo de tempo. Para capturar estas características no modelo da epidemia nos realizamos diferentes possibilidades: somente uma interação por sujeito por passo de tempo, escolhido aleatoriamente de todas as possíveis conexões (regra *one*). Assumimos, como outra possibilidade, que cada vértice toma um subconjunto aleatório das suas ligações e todos selecionados interagem simultaneamente (regra *fraction*). Por fim, o caso extremo da interação com toda a vizinhança (regra *all*). Mostramos na próxima figura como estas regras afetam os resultados da dinâmica da epidemia na rede.

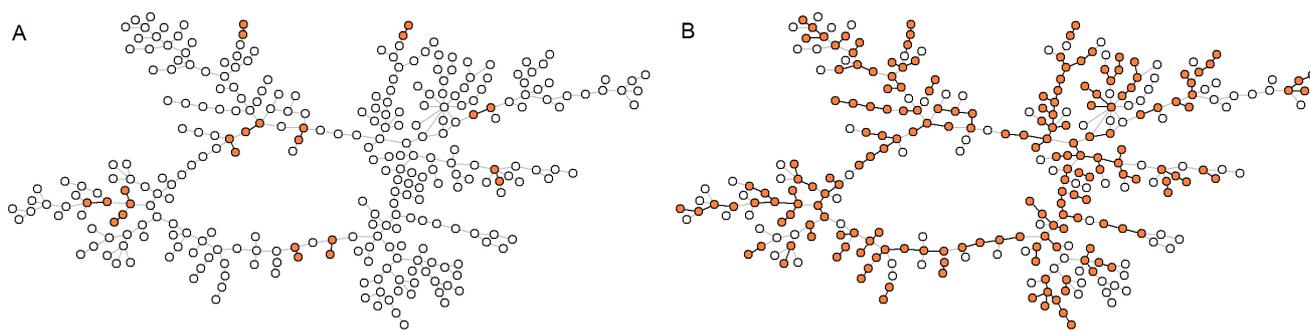


Figura 6.8: Instantes da rede dados os *frames* no tempo (A) $t_1 = 10$ and (B) $t_2 = 34$. Os vértices hachurados referem-se aos ativos nos referidos tempos.

As três medidas mostradas na Fig. 6.7 apontam para a mesma tendência geral: da regra *one* a regra *all* nós gradualmente nos aproximamos do caso de todos conectados. Nós observamos que este caso limite é claramente um limite acadêmico. Observe que o extremo *all* está abaixo deste limite, o que se deve ao fato dos sujeitos na rede romântica não estarem totalmente conectados.

6.3.2 Rede Dinâmica

A mais relevante característica da rede romântica é a componente gigante, formada por 288 estudantes. Porém, como foi dito antes, a rede estática é o resultado da superposição de todas as ligações ao longo dos 18 meses de pesquisa. Baseado no filme feito por um dos autores do estudo, James Moody ², queremos ver como evolui uma hipotética infecção tendo em conta a sequência temporal. Para esta suposição, tomamos *gif* animada do site mencionado. Tal animação está partida em 45 *frames*. Cada *frame* corresponde a um período de 12 dias, que por simplicidade assumimos uniformemente distribuído no tempo. Os vértices e as ligações entre eles foram identificados e convertidos em dados de entrada para a simulação do modelo de epidemia. Cada *frame* tomada individualmente é muito diferente da superposição de todas elas (componente gigante estática). Para dar uma ideia de como são, mostramos na figura 6.8 duas *frames* no tempo $t_1 = 10$ e $t_2 = 34$ que representa tempos de baixa e alta atividade, respectivamente, na rede. Não é difícil ver que *clusters*, vértices ativos em uma determinada janela, deverão ter um tremendo impacto na propagação da doença. Diferente do caso estático, não dispõem da conexão para todo período de tempo.

Antes de apresentarmos os resultados para epidemia no cenário dinâmico, mostraremos algumas relevantes medidas da rede dinâmica. A figura 6.9A exibe a distribuição de

²Dinâmica de criação de parceiros como apresentado em www.soc.duke.edu/~jmoody77/NetMovies/rom_flip.html

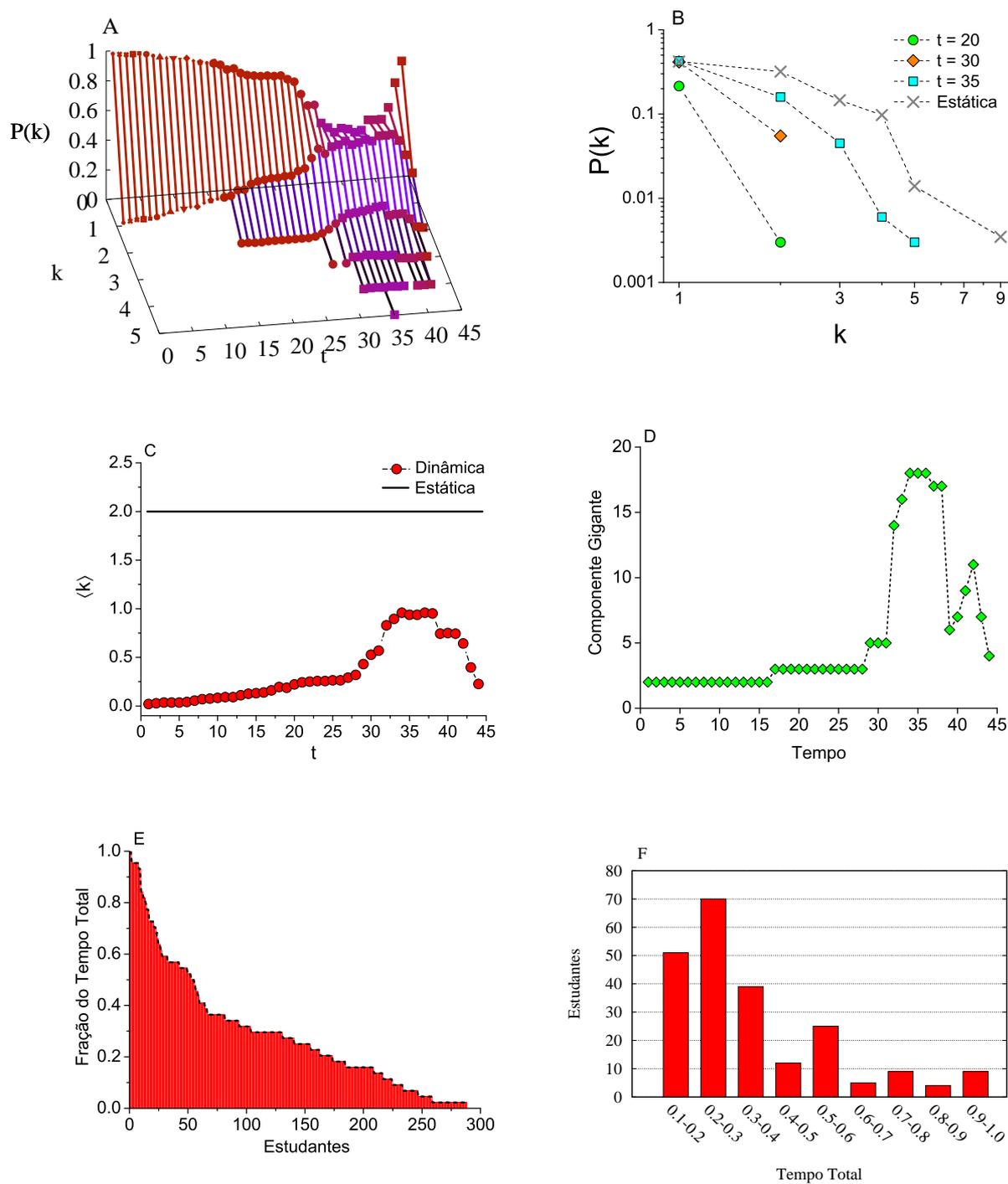


Figura 6.9: (A) Evolução da distribuição de grau da componente gigante da rede romântica, quando os *frames* são tomados separadamente. (B) Distribuição de grau em três tempos distintos ($t = 20, 30, 35$) comparado com a distribuição de grau estática. (C) Grau médio $\langle k \rangle$ em função do tempo. (D) Tamanho da componente gigante em função do tempo. (E) Fração do tempo de interação por estudante. (F) Distribuição dos tempos de interação, como fração do tempo total, por estudante.

grau em diferentes tempos. Por simples inspeção na figura, vemos que muitas medidas usuais no contexto de redes (grau, clusterização, etc) serão distintas se aplicado a um indivíduo de uma dada *frame* do que aplicado a mesmo indivíduo na rede estática. Podemos observar, por exemplo, na figura 6.9B onde comparamos a distribuição de grau em três diferentes tempos com a distribuição estática (equivalente a superposição de todas as *frames*). Mais drástico é o efeito sobre o grau médio (figura 6.9C) onde vemos que a dinâmica leva a um pico de valor $\langle k \rangle = 1$, mas na maior parte do tempo é muito menor do que valor estático $\langle k \rangle = 2$. Provavelmente o mais importante efeito na medida — em razão dos diferentes valores — é que em todas as *frames* muitos indivíduos não tem contato ($k_i = 0$). O efeito da propagação da doença é claro, eles podem não ser infectados durante estes intervalos de tempo. Conforme o tempo passa, destacamos: para $t > 15$ a interação é somente de pares, com muitos indivíduos isolados ao mesmo tempo; o intervalo de tempo $30 < t < 40$ é onde aumenta a atividade, chegando ao máximo em $t = 35$; para $t > 40$ a distribuição de grau retorna a baixos valores. Outra importante questão relacionada com a dinâmica de formação das ligações é a fração do tempo que um indivíduo participa da atividade romântica, que é representada na figura 6.9E. Desta figura vemos uma espécie de lei de Pareto para o tempo de participação: uma fração pequena de estudantes, $\sim 3\%$, interage 90% do tempo total, enquanto que a maioria, $\sim 70\%$ dos estudantes, participa 40% do tempo total. Colocando junto as medidas da dinâmica obtidas da rede romântica, ou seja, a distribuição de grau como função do tempo e a distribuição dos tempos de interação (Fig. 6.9F), é evidente que a propagação da doença que depende destas ligações será fortemente afetada pela dinâmica. Mas como? Enquanto que é claro que fazer tratamento analítico gerais não é possível. Isto se deve em parte, ao fato que não existe uma forma funcional para sequência de como são criadas e destruídas as ligações. É também claro que dinâmica da rede será sempre resultará em uma epidemia menos grave. Não somente existem períodos em que vemos pequenas componentes isoladas (nesses momentos a infecção não pode saltar de um componente para a outras), mas a janela de tempo em que eles estão conectados (e não todos ao mesmo tempo) são relativamente curtos. No contexto da propagação da doença, é mais difícil se propagar quando as ligações não são estáticas. Em geral o tamanho final da epidemia $R(\infty)$ dependerá da relação entre duas escalas de tempo: o período de infecção e o típico tempo de interação entre os indivíduos da rede, veja capítulo anterior.

No que segue apresentaremos os resultados das simulações da epidemia na rede dinâmica. A rede dinâmica é, de acordo com a *gif* animada, é uma sucessão de fragmentos estáticos (subconjuntos da rede romântica estática) que duram 12 passos (isto pode representar 12 dias). O período é resultado do tempo total de relaxação do programa *Jefferson School* ($t_f = 540$ dias) e o número de *frames* (45) da *gif* animada. Notavelmente, para

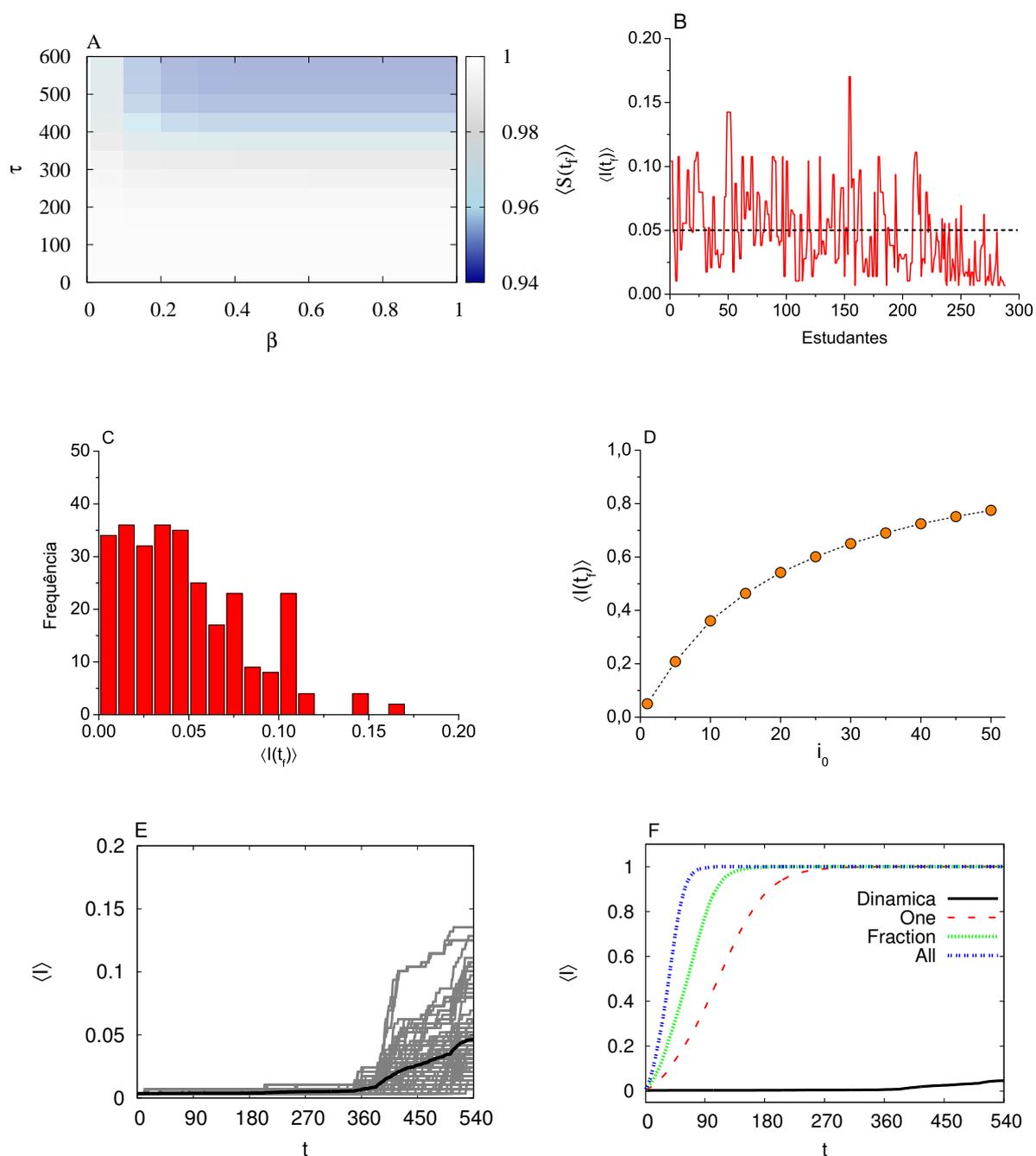


Figura 6.10: Modelo sustentável-Infetado (SI) resulta na “rede romântica dinâmica”: (A) Fração final de sujeitos não infectados $S(t_f)$ como função do tempo de infecção τ e probabilidade de infecção β (região branca implica sem infecção); (B) Distribuição do tamanho da infecção como função da sujeitos infectados iniciais; (C) Histograma da distribuição; (D) Tamanho final da infecção versus o número de infectado inicial; (E) Fração de infectados versus o tempo (diferentes realizações em cinza e média em preto) e (F) Número médio de infectados em função do tempo, comparação entre as redes dinâmica e estática com diferentes regras para a dinâmica.

o intervalo de valores $\beta\tau$ explorados o tamanho final da epidemia é sempre muito menor do que os correspondentes na rede estática. Independente de $\beta\tau$ para $\beta\tau > 1$. Logo, não observaremos epidemia neste cenário. Na realidade, podemos ver na figura 6.10A que apenas para tempos de infecção τ próximos ao tempo total da dinâmica existem alguns efeitos sobre a população de suscetíveis, mas muito fracos. Então, é interessante considerar um caso extremo de epidemia, isto é, uma dinâmica da epidemia SI. Em outras palavras, uma doença com tempo de infecção infinito³. Desta maneira, o relevante tempo é a duração da dinâmica. Fazendo o caso SI extremo tomamos $\beta = 1$, que significa que cada contato entre indivíduos infectado e suscetível muda o estado do último para infectado. Nesta configuração o surto final de infectados, considerando todos os possíveis estudantes como infectado inicial não é maior que 5% (veja a figura 6.10C). Só para se ter a situação para muitos extremos, a figura 6.10D exhibe o número de infectados finais como função do número de infectados iniciais, vemos que com mais de 50 sobre 288 infectados iniciais (17%) a proporção final de infectados é menor do que 80%. Juntando estes resultados, torna-se evidente que a regra chave da formação dinâmica dos pares traz um diferente cenário para todas as perspectivas dos resultados em modelos estáticos de redes. Na rede dinâmica o número total de indivíduos infectados é fortemente afetado pela condição inicial como pode ser visto na figura 6.10E. Neste caso, quem é o infectado inicial é muito importante, mas não devido a sua pessoal atividade, devido a sequência particular em que os indivíduos interagem entre si. O importante é toda a sequência, não o particular sujeito.

Para a dinâmica vemos que o número de sujeitos infectados cresce substancialmente para $t > 360$. Isto porque a distribuição de grau muda significamente neste intervalo de tempo (figura 6.9A). Visto que a partir deste tempo um expressivo número de indivíduos aumenta o número de seus contatos. Este é um sinal que a dinâmica social aumenta significamente. A figura 6.10F mostra a comparação entre os cenários estático e dinâmico da rede romântica relacionado com a propagação de uma doença SI. A diferença é notável. A rede estática, em qualquer regra de transmissão, é mais eficiente para a propagação da doença do que a versão dinâmica. Portanto, a inclusão da formação dinâmica de ligações traz uma boa notícia para o campo da epidemia. O fato que todas as ligações podem ser observadas na rede estática não são avaliados em todos os tempos e, conseqüentemente, fazendo uma grande diferença. Em outras palavras, em diferentes tempos muitos subgrafos estão desconectados da componente gigante. Isto porque escolhemos o caso mais extremo para a epidemia, dinâmica SI, em que uma vez infectado o indivíduo permanece neste estado por toda a vida (ou em termos práticos, por um tempo muito maior que o observado).

³Algumas doenças, por exemplo a infecção do HIV, tem esta espécie de dinâmica onde o τ é igual a expectativa de vida

Capítulo 7

Conclusões

Como foi visto neste trabalho, investigamos o comportamento de um modelo de epidemias baseado em uma dinâmica em que os indivíduos de uma dada população podem ser classificados como Suscetíveis-Infetados-Removidos (SIR). Dentre as inúmeras possibilidades de investigação com este modelo, escolhemos explorar a solução algébrica na aproximação de campo médio, os efeitos na propagação da epidemia frente a formação dinâmica de contatos e a difusão de uma doença contagiosa em uma rede realística.

Na primeira contribuição, apresentamos uma solução algébrica para o modelo SIR padrão (aproximação de campo médio). Em síntese, o estado da doença (S, I, R) possui uma forma paramétrica em termos de uma dada função. Esta função, em particular, é responsável pela dinâmica das classes. Em geral, dado uma pequena fração de infectados iniciais, esta solução é válida ao longo de toda a dinâmica. Além disso, sugere se manter como uma boa aproximação para um vasto intervalo do número básico de reprodução R_0 . Consequentemente, decorre desta solução um importante resultado: uma expressão analítica para o tempo do pico de infecção. Este resultado aliado as expressões de R_0 , número assintótico de removidos $R(\infty)$ e número máximo de infectados I_{max} (apresentadas no artigo original de Kermack-McKendrick) representam de forma algébrica as principais quantidades do modelo SIR.

Na segunda contribuição, inicialmente apresentamos um modelo de interação dinâmico entre indivíduos em que as probabilidades de formação e dissolução das ligações são constantes. Nós observamos que neste cenário dois possíveis regimes de interação são possíveis: equilíbrio e fora do equilíbrio. Enquanto que no regime de equilíbrio o grau médio permanece constante (o número médio de contatos permanece constante, não os contatos) no fora de equilíbrio o grau médio evolui no tempo de forma exponencial. Em particular, esta forma exponencial evolui até um assintótico que coincide com grau médio do regime de equilíbrio, dadas as mesma taxas de interação. Em outras palavras, para um tempo suficientemente

grande os regimes são equivalentes.

Na sequência, investigamos a propagação de uma hipotética doença cuja dinâmica é do tipo SIR nestes regimes. Para tanto, nós estudamos duas quantidades; número básico de reprodução (R_0) e a probabilidade de ocorrência da epidemia (P_{epi}). Além disso, a relação destas quantidades com a dinâmica da epidemia. Em verdade, em mais detalhes, observamos que tanto R_0 como P_{epi} dependem da relação entre os tempos característicos de propagação da doença ($1/\gamma$) e formação da rede ($1/\sigma$), ou simplesmente das probabilidades de recuperação (γ) e remoção (σ). Ao considerar $\sigma \gg \gamma$ (tempo de formação da rede muito menor do que o tempo de propagação da doença), concluímos que tanto R_0 como P_{epi} são equivalentes nos regimes de equilíbrio e fora de equilíbrio. Em particular, iguais aos resultados típicos de campo médio (modelo SIR padrão). Com efeito, concluímos que a dinâmica da epidemia neste cenário não se difere dos resultados típicos de campo médio (descrito por Kermack-McKendrick). No outro extremo das relações entre os tempos característicos, $\sigma \ll \gamma$, observamos que as quantidades investigadas são diferentes entre os regimes. Desta maneira, nós obtemos que para o regime fora de equilíbrio $R_0 = P_{epi} = 0$, ou seja, sobre esta relação entre tempos não há epidemia. Já para o outro regime, equilíbrio, dependendo dos valores dos parâmetros tipos da rede (ρ e σ) e da doença (β e γ), nós constatamos que $R_0 > 1$ e $P_{epi} > 0$. Estas observações indicam que uma epidemia é possível neste cenário. Já entre estes extremos $\sigma \sim \gamma$, verificamos novamente uma diferença do comportamento da epidemia entre os regimes. Por um lado, o número básico de reprodução é próximo entre a situação fora de equilíbrio e equilíbrio. Por outro, a probabilidade da epidemia no regime fora do equilíbrio é inferior daquela relacionada com situação de equilíbrio. Em síntese, em termos da dinâmica da epidemia, concluímos que ambos os regimes são equivalentes a menos de um tempo de retardo, tempo este associado a probabilidade de remoção.

Na terceira contribuição, nós apresentamos o problema da propagação de uma doença sexualmente transmissível em uma rede real (chamada de “rede romântica”) em dois aspectos. Primeiro, comparamos a epidemia nesta rede com a epidemia em redes *small world* e *scale free*. De interesse nesta comparação, a relação entre a rede real e as topologias geradas. O segundo aspecto, o efeito da dinâmica de formação de *links* e como eles podem afetar a dinâmica da epidemia.

Em relação à rede estática e a comparação entre as redes *small world* e *scale free*, constatamos que a máxima incidência I_{max} e o tamanho final $R(\infty)$ são sensíveis à topologia para valores relativamente grande de $\beta\tau$ (> 5). Para pequenos valores de $\beta\tau$ (< 4) é difícil distinguir as diferentes topologias, exceto para redes com alta heterogeneidade, em termos do surto da epidemia. Ainda mais, se considerarmos a ampla dispersão dos resultados devido ao pequeno tamanho das redes (barras das figuras 6.3 e 6.4). Redes homogêneas, como a

Watts-Strogatz com $p \simeq 1$ ou Krapvisky-Redner com $r \simeq 0$ são estatisticamente similares entre eles e com a rede romântica em termos do tamanho final da epidemia. Isto é válido, quando $\beta\tau$ não é muito grande e é, devido ao peculiar valor do grau médio ($\langle k \rangle = 2$), que faz com que todas as estruturas sejam do tipo anel ou árvore, como se pode ver através da figura 6.1. A principal diferença entre essas realizações artificiais e a real é o seu anel. No entanto, é um único anel e relativamente grande. Entre todas as redes estáticas artificiais testadas neste contribuição, podemos concluir que a rede Watt-Strogatz com $p \simeq 1$ (rede aleatória com $\langle k \rangle = 2$) é indistinguível da rede romântica em termos de propagação da doença.

Posteriormente, considerando o campo externo, a promiscuidade, e as regras de interação, nossos resultados sugerem que a dinâmica da epidemia é afetada por estas três variantes. Em relação ao campo externo, o efeito de que os indivíduos estão em contato com indivíduos infectados externos tem é a de tornar a rede mais perto de uma rede em que todos os vértices compartilham ligação (campo médio). No que diz respeito a promiscuidade sua importância é confirmada na rede romântica. Em $\beta\tau = 4$, por exemplo, a probabilidade de uma epidemia envolvendo mais que 5% dos indivíduos triplica quando mudamos o grau do infectado inicial (menor para o maior). Além disso, quanto maior o grau do infectado inicial menor o tempo para atingir o pico de infecção. Nossos resultados com as regras de interações indicam que a propagação de uma epidemia não apenas depende do número disponível de contatos, mas também da frequência dos contatos. Observamos ainda que o acréscimo na frequência dos contatos aproxima a dinâmica da epidemia aos resultados típicos de campo médio.

Indo para a situação da rede dinâmica, em que consideramos a sequência temporal das interações entre os sujeitos, chegamos a conclusões muito interessantes. Nós mostramos, baseado em simulações, que a rede dinâmica sempre causará epidemias em que o tamanho final é pequeno. Porque tanto há períodos em que vemos apenas pequenas componentes isoladas como existem janelas de tempo relativamente curtas em que elas são ligadas. Em outras palavras a relação de tempos, diretamente influencia o curso da epidemia. Claramente, a propagação da doença, que depende das conexões, será fortemente afetada pelo processo dinâmico. Sendo assim, a distribuição dos tempos de interação, ausente na rede estática, é crucial para a propagação da doença. Ao testar uma situação de epidemia hipotética extrema, o modelo Suscetível-Infectado (SI) com máxima probabilidade de infecção, nós obtemos uma fração de infectados em torno de 15% no pior cenário (iniciando a infecção com o indivíduo mais infectado). Porém, a fração média de infectados, sobre todas as possíveis escolhas para o infectado inicial, não é maior do que 5%. Juntando isto, este resultado indica o papel chave da dinâmica de formação dos pares, o que traz um cenário diferente do anterior resultado em redes estáticas. Notavelmente, na rede dinâmica o número total de indivíduos

infectados é fortemente afetada pela condição inicial. Neste caso, quem é o infectado inicial é importante, devido a sequência particular em que os indivíduos se conectam entre eles. Assim, é toda a sequência que importa, não exclusivamente o número de parceiros. Além disso, comparando as figuras 6.10E e 6.9C vemos uma clara correlação entre elas. Então, a dinâmica da componente gigante está relacionada com a evolução do fração de indivíduos infectados. Como em uma rede aleatória, a componente gigante segue uma relação fechada com o grau médio, como pode ser comparado nas figuras 6.9C e 6.9D. Isto nos leva a concluir que a fração de pessoas infectadas está relacionada com a dinâmica do grau médio, o que pode ser importante para a saúde pública. Adicionalmente, nós observamos que a rede estática, em qualquer regra de transmissão, é muito menos eficiente para a propagação da doença que a sua versão dinâmica. O efeito da formação dinâmica das conexões está completamente claro: uma doença é influenciada pela estrutura das ligações que, por sua vez, não estão disponíveis o tempo todo, mas em janelas de tempo muito específicas. Assim, o pior cenário é sempre a rede estática, quando todas as conexões estão disponíveis o tempo todo.

Como perspectiva natural para a continuidade do trabalho, propomos analisar os efeitos na propagação de uma epidemia em um cenário dinâmico em que as probabilidades de formação e dissolução das interações não seja constante. Por exemplo, seria interessante adicionar uma dependência temporal neste parâmetro de formação dos pares, visto que as interações sociais bem como a intensidade delas sofrem influência do tempo. Outra possibilidade seria estender as medidas realizadas para a rede romântica para outras redes realísticas.

Referências Bibliográficas

- [1] Lechevalier, H., Bacteriological reviews **36** (1972) 135.
- [2] Kermack, W. and McKendrick, A., Bulletin of Mathematical Biology **53** (1991) 33.
- [3] Barabasi, A.-L., Linked: How everything is connected to everything else and what it means, 2003.
- [4] Newman, M. E. J. and Park, J., Phys. Rev. E **68** (2003) 036122.
- [5] Pastor-Satorras, R. and Vespignani, A., Phys. Rev. Lett. **86** (2001) 3200.
- [6] Risau-Gusman, S., arXiv/1004.1378 (2010).
- [7] Albert, R. and Barabási, A.-L., Rev. Mod. Phys. **74** (2002) 47.
- [8] Milgram, S., Psychology Today **2** (1967) 60.
- [9] Erdos, P. and Renyi, A., Publ. Math. Debrecen **6** (1959) 290.
- [10] Granovetter, M., The American Journal of Sociology **78** (1973) 1360.
- [11] Watts, D. J. and Strogatz, S. H., Nature **393** (1998) 440.
- [12] Barabási, A.-L. and Albert, R., Science **286** (1999) 509.
- [13] Newman, M., SIAM Review **45** (2003) 167.
- [14] Costa, L. d. F., Rodrigues, F. A., Travieso, G., and Boas, P. R. V., Advances in Physics **56** (2007) 167.
- [15] Barrat, A. and Weigt, M., The European Physical Journal B - Condensed Matter and Complex Systems **13** (2000) 547.
- [16] Newman, M. E. J., Moore, C., and Watts, D. J., Phys. Rev. Lett. **84** (2000) 3201.
- [17] Chen, F., Chen, Z., Wang, X., and Yuan, Z., Communications in Nonlinear Science and Numerical Simulation **13** (2008) 1405.

- [18] Wald, A., *The Annals of Mathematical Statistics* **15** (1944) 283.
- [19] Fronczak, A., Fronczak, P., and Holyst, J., *Phys. Rev. E* **68** (2003) 046126.
- [20] Abbey, H., *Hum. Biol.* **3** (1952) 201.
- [21] Watson, H. W. and Galton, F., *Journal of the Anthropological Institute of Great Britain*. **4** (1875) 138.
- [22] Draief, M., *Physica A: Statistical Mechanics and its Applications* **363** (2006) 120.
- [23] Newman, M. E. J., *Phys. Rev. E* **66** (2002) 016128.
- [24] Newman, M. E. J., Strogatz, S. H., and Watts, D. J., *Phys. Rev. E* **64** (2001) 026118.
- [25] Volz, E. and Meyers, L. A., (2007).
- [26] Read, J. M., Eames, K. T., and Edmunds, W. J., *Journal of The Royal Society Interface* **5** (2008) 1001.
- [27] Volz, E. and Meyers, L. A., *Proceedings of the Royal Society B: Biological Sciences* **274** (2007) 2925.
- [28] Batagelj, V. and Mrvar, A., *Pajek – program for large network analysis*, 1998.
- [29] Bearman, P. S., Moody, J., and Stovel, K., *American Journal of Sociology* **110** (2004) 44.
- [30] Krapivsky, P. L. and Redner, S., *Phys. Rev. E* **63** (2001) 0066123.