



UNIVERSIDADE FEDERAL RURAL DE PERNAMBUCO (UFRPE)
CENTRO DE ENSINO DE GRADUAÇÃO EM EXATAS E NATUREZA (CEGEN)
DEPARTAMENTO DE FÍSICA
RELATÓRIO FINAL - PIC

ERIKE SIMON COSTA CATIVO DO NASCIMENTO

INVESTIGAÇÃO DA DISTRIBUIÇÃO ESPACIAL DE
VULNERABILIDADES SOCIOECONÔMICAS SOBRE DINÂMICAS
EPIDÊMICAS

RECIFE/PE
2022

Sumário

	Páginas
1 Identificação	3
2 Introdução	4
3 Objetivos	5
3.1 Objetivo geral	5
3.2 Objetivos específicos	5
4 Metodologia	5
4.1 Gráfos	6
4.2 Métricas e arquiteturas de redes complexas	6
4.2.1 Grafos aleatórios	8
4.2.2 Redes de mundo pequeno	9
4.2.3 Redes livre de escala	10
5 Resultados e discussões	15
6 Considerações finais	23
7 Atividades desenvolvidas pelo aluno	23
8 Referências bibliográficas	24

1 Identificação

Discente:

- **Nome** Erike Simon Costa Cativo do Nascimento
- **Telefone** (81) 99936-1508
- **E-mail** erike.fisica@gmail.com

Orientador:

- **Nome** Pedro Hugo de Figueirêdo
- **Telefone** (81) 98809-8502
- **E-mail** phugof@gmail.com

Instituição de Vínculo:

- **Nome** Universidade Federal Rural de Pernambuco
- **Endereço** Rua Dom Manuel de Medeiros, s/n, Dois Irmãos - CEP: 52171-900 - Recife/PE
- **Telefone** (81) 3320-6001
- **E-mail** reitoria@ufrpe.br

Curso: Licenciatura Plena em Física

Departamento: Física

Período: Janeiro/2022 - Abril/2022

2 Introdução

Sistemas Complexos são constituídos por um grande conjunto de elementos que interagem entre si exibindo propriedades coletivas emergentes. Tais propriedades decorrem principalmente da relação não-linear entre as partes, ou seja, em um sistema complexo, o todo é mais que a soma das partes. Devido a grande versatilidade de aplicações no mundo real, é por natureza um tema multidisciplinar, envolvendo várias áreas do conhecimento, como física, biologia, matemática, computação e sociologia.

Podemos investigar algumas características dos sistemas complexos através do estudo de grafos, ferramenta através da qual redes complexas podem ser representadas. Grafos apresentam uma estrutura topográfica não trivial, composto por um conjunto de vértices (nós) que são interligados por meio de arestas [01]. Pode-se considerar que o estudo das redes teve sua origem na Teoria dos Grafos por volta de 1736, quando Euler solucionou o famoso problema das sete pontes de Königsberg. Portanto, vários aspectos dos sistemas do mundo real podem ser representados por meio de redes complexas. Dentre as diversas possibilidades, é possível estudar a topologia da *World Wide Web (WWW)*, redes de transporte, neurais, tróficas (*food webs*), metabólicas, sociais, entre outras.

Entre os diversos modelos de redes utilizados na tentativa de caracterização dos sistemas do mundo real, uma das principais são as redes aleatorias [06]. Porém, observa-se que a topologia e a evolução das redes reais apresentam propriedades bastante sofisticadas e distintas das redes aleatórias. Podemos dizer que as redes complexas não seguem um padrão regular e modelar suas peculiaridades torna-se um processo delicado. Fica evidente então a importância da caracterização da anatomia desses sistemas, uma vez que sua estrutura afeta a sua função. Por exemplo, a topologia das redes sociais interfere diretamente na disseminação de informações e espalhamento de infecções, a estrutura dos sistemas de transmissão elétrica de uma cidade afeta sua estabilidade e robustez, as malhas aéreas dependendo das suas conexões entre aeroportos, podem tornar os vôos mais ou menos eficientes, interferindo nos custos de passagem e no tempo de viagem, entre vários outros exemplos que podem ser citados. Segundo E. O. Wilson [02], “O maior desafio hoje, não apenas em biologia celular e ecologia, mas em todas as áreas da ciência, é a descrição precisa e completa sistemas complexos. Os cientistas quebraram muitos tipos de sistemas. Eles acham que conhecem a maioria dos elementos e forças. A próxima tarefa é remontá-los, pelo menos em modelos matemáticos que capturam as propriedades chave de todos os conjuntos.” Nota-se então, que existem diversas complicações inerentes no entendimento das redes. Algumas delas são a complexidade estrutural, evolução da rede, diversidade de conexões, complexidade dinâmica, diversidade dos tipos de nós e meta-complicações [03]. Entender o papel dessas complicações e peculiaridades para cada tipo de rede é de suma importância para um entendimento o mais próximo do real possível dos sistemas complexos.

3 Objetivos

3.1 Objetivo geral

Investigar via simulações computacionais como propriedades estruturais de redes complexas influenciam em dinâmicas epidêmicas, buscando entender o comportamento inicial da pandemia de COVID-19 através da rede de transporte público da região metropolitana do Recife – PE, desenvolvendo um algoritmo computacional genérico que reproduza redes do tipo livre de escala, carregando consigo características mais genuínas dos sistemas complexos do mundo real.

3.2 Objetivos específicos

- Formação do aluno na área de sistemas complexos, abordagem de ampla aplicação em fenômenos físicos e interdisciplinares.
- Formação de uma cultura científica com apresentação dos resultados do trabalho científico, bem como na elaboração de artigos e relatórios.
- Contribuir para a compreensão dos possíveis regimes de propagação epidêmica e sua dependência em relação a topologia de redes complexas.

4 Metodologia

O ponto de partida para a pesquisa se deu inicialmente a partir da revisão bibliográfica sobre sistemas complexos, grafos, modelos epidêmicos, arquitetura e propriedades de redes complexas. Após esse embasamento teórico e juntamente com o intuito de desenvolver um estudo relacionado à dinâmica epidêmica do SARS-CoV-2, surge a ideia do desenvolvimento de um algoritmo computacional que reproduza redes genéricas do tipo sem escala [04], de maneira que possamos gerar dinâmicas epidêmicas do SARS-CoV-2 nestas redes e observar, a partir das variações dos seus parâmetros, o comportamento da curva de casos infectados para a fase inicial da pandemia de COVID-19, fazendo uma comparação com a curva de casos infectados reais para a região metropolitana do Recife (RMR). Entender a anatomia desse tipo de rede é crucial, pois estudos apontam que a característica livre de escala é um ponto chave no aparecimento de um crescimento em lei de potência no número de casos infectados na fase inicial em diversos processos epidêmicos. Tal comportamento foi observado para diversos países para a COVID-19 [15].

O objetivo inicial da pesquisa foi desenvolver um algoritmo computacional em linguagem *Python* que gere redes sem escala típica. Um segundo objetivo é desenvolver outro algoritmo a partir do primeiro para uma variação do modelo rede sem escala típica (Dorogovtsev - Mendes)

[12] e observar como a estrutura da rede genérica a partir das variações dos parâmetros afeta a responsividade da dinâmica apidêmica.

A seguir, uma breve revisão teórica dos principais conceitos, tipos e métricas sobre grafos e redes complexas.

4.1 Gráfos

Grafos são objetos matemáticos compostos por um conjunto de pontos ou nós (vértices) e linhas que formam ligações entre esses pontos (arestas). Na Figura 4.1, temos representações de grafos genéricos. Podem ser não-dirigidos (ligação mútua entre vértices) ou dirigidos (digrafos) como mostram as figuras 4.1(a) e (b), respectivamente. Além disso, tais ligações podem apresentar pesos diferentes, como na Figura 4.1(c).

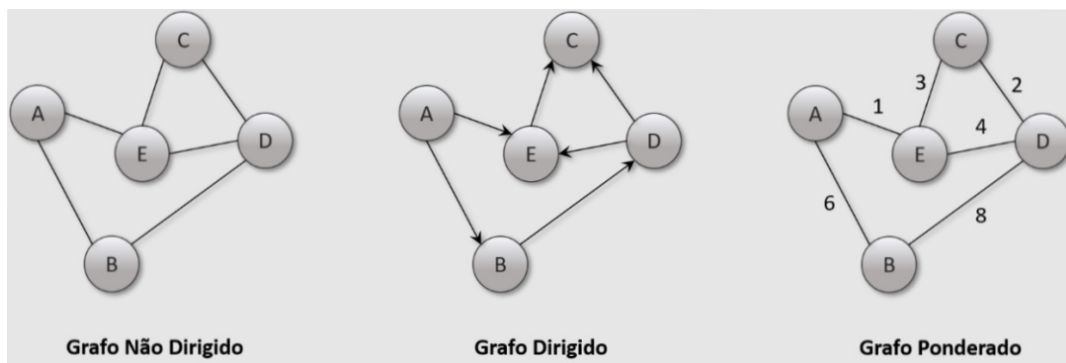


Figura 4.1: (a) Grafo não dirigido. (b) Grafo dirigido. (c) Grafo ponderado.

Fonte: <[http://sao-paulo.bigdataweek.com/2018/08/13/](http://sao-paulo.bigdataweek.com/2018/08/13/uma-visao-geral-do-funcionamento-de-grafos-em-estruturas-de-dados/)

[uma-visao-geral-do-funcionamento-de-grafos-em-estruturas-de-dados/](http://sao-paulo.bigdataweek.com/2018/08/13/uma-visao-geral-do-funcionamento-de-grafos-em-estruturas-de-dados/)>

Sistemas complexos podem ser modelados através da teoria dos grafos. Computacionalmente existem diversas maneiras pela qual grafos podem ser representados. No desenvolvimento do algoritmo para a presente pesquisa, a representação se deu por listas de adjacência, onde uma lista com N elementos (vértices) contém, em cada posição i associada a cada vértice, uma nova lista que indica quais conexões o vértice i tem com os restantes. Por exemplo, temos a seguinte lista com $N=5$ elementos: $[[2,3,5], [1,4], [1,4,5], [2,3], [1,3]]$. O elemento da posição $i=1$ (vértice 1) está conectado aos vértices 2, 3 e 5. O elemento da posição $i=2$ (vértice 2) está conectado aos vértices 1 e 4, e assim por diante.

4.2 Métricas e arquiteturas de redes complexas

Um ponto de partida para o estudo de redes complexas se dá através de suas métricas. Estas nos retornam características estruturais dos tipos de redes que se quer estudar. Começaremos pela métrica mais básica, que é a conectividade k .

Grau ou conectividade de um vértice (k_i): Representa o número de ligações que um dado vértice i faz com seus vizinhos. Para uma rede não-dirigida como na Figura 4.1(a), a conectividade k_i do vértice i é dada simplesmente pelo número de conexões existentes entre ele e seus vizinhos. Caso a rede seja dirigida, cada vértice i irá possuir conectividades de entrada k_{in} e saída k_{out} . Por exemplo, para o vértice E da Figura 4.1(b), temos $k_{E,in} = 2$ e $k_{E,out} = 1$. Logo o número total de conexões para um dado vértice i em uma rede dirigida é dado por:

$$k_i = k_{i,in} + k_{i,out} \quad (1)$$

Podemos então definir a função **distribuição de conectividade $P(k)$** , que nos fornece a probabilidade de que um vértice, escolhido aleatoriamente possua k ligações. Em outras palavras, esta função exibe a fração de vértices de uma rede que apresentam conectividade k . Esta pode ser quantificada por meio de uma função de distribuição cumulativa (Equação 2), onde pk_i é a fração de vértices da rede que possuem conectividade ou grau k_i e P_k é a função cumulativa de distribuição de probabilidades. A Equação 2 é válida para redes não-dirigidas. Já para redes dirigidas, a função de distribuição (Equação 3) é escrita em função de pk_{ij} , com duas variáveis, que representam a fração de vértices que tem, simultaneamente, grau de entrada $k_{i,in}$ e grau de saída $K_{j,out}$.

$$P_k = \sum_i^{\infty} pk_i \quad (2)$$

$$P_k = \sum_{ij}^{\infty} pk_{ij} \quad (3)$$

A depender do tipo de rede, teremos distribuições de conectividades diferentes. Para redes aleatórias segue a distribuição de Poisson, porém, em muitas redes reais a distribuição de graus segue uma Lei de Potência, em que $P_k \sim k^{-\gamma}$, onde γ é um valor constante.

Coefficiente de agrupamento ou clusterização (CA): Mede o grau com que os vértices de um grafo tendem a agrupar-se. Esses agrupamentos intrínsecos também são conhecidos como fenômeno de transitividade. Se determinado vértice i estiver conectado ao nó j , que por sua vez encontra conectado com k , existe uma probabilidade elevada de i também estar conectado com k [05]. Em outras palavras, a transitividade indica a presença de um número elevado de triângulos na rede (conjuntos de três vértices conectados uns aos outros).

É obtido a partir da Equação 4, onde $\#\Delta$ indica o número de triângulos na rede e $\#v$ o número de vértices triplamente conectados (vértices com arestas não-dirigidas para outro par de nós). O fator 3 no numerador refere-se ao fato de que cada triângulo representa três triplas e também para garantir que o coeficiente de aglomeração seja um valor entre zero e um.

$$CA = \frac{3 \times \#\Delta}{\#v} \quad (4)$$

Caminho médio (L): Em teoria dos grafos, a distância é definida em termos de passos, como sendo o número de *edges* (arestas) entre dois vértices quaisquer, medido sobre todos os vértices da rede. Dessa forma, mede a separação física entre dois vértices em um grafo.

4.2.1 Grafos aleatórios

Proposto por Erdős e Rény [06], é o modelo mais simples que uma rede complexa pode assumir. Para um número fixo N de vértices, arestas não-dirigidas (m) são adicionadas aleatoriamente com alguma probabilidade p a um par de vértices. Uma representação de uma rede aleatória pode ser vista na Figura 4.2. Erdős e Rény estudaram como a topologia da rede muda em função de m . Se o número m de arestas é baixo, é provável que a rede seja fragmentada com alguns pequenos *clusters* de vértices, que são chamados de "componentes". A medida que m aumenta, os componentes crescem, ligando-se a vértices isolados e posteriormente a outros componentes. A transição de fase irá ocorrer para $m = N/2$, onde muitos *clusters* se cruzam formando um único componente grande. Para $m > N/2$, seu tamanho escala linearmente com n , com $n \rightarrow \infty$. Além disso, o número máximo de graus de separação entre dois nós cresce lentamente com $\log n$. Como já comentado, a distribuição de conectividade desse tipo de rede prediz uma distribuição de Poisson com limite máximo N . O grau esperado para um vértice qualquer segue a Equação 5, sendo p a probabilidade de conexão entre dois vértices.

$$\langle K \rangle = p(N - 1) \quad (5)$$

Tais grafos tem servido para idealizar arquiteturas acopladas para modelos dinâmicos de rede de genes, ecossistemas, espalhamento de infecções e vírus computacionais.

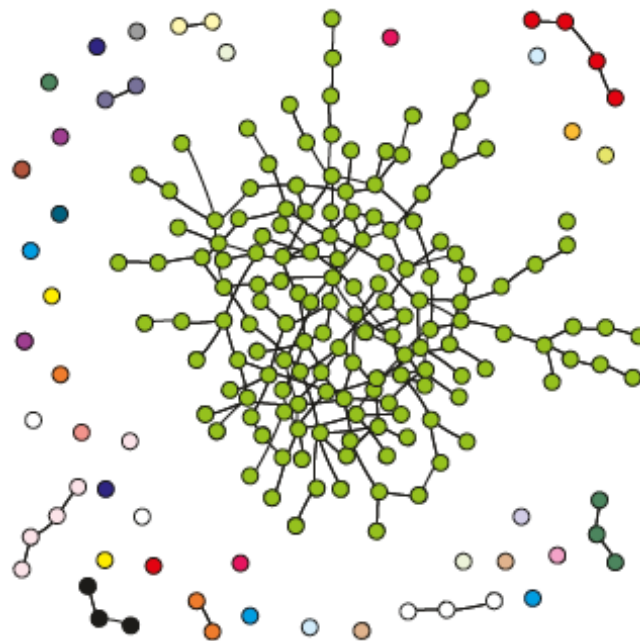


Figura 4.2: Grafo aleatório com um grande componente central e outros menores em torno.

Fonte: STROGATZ, S. H. Exploring Complex Networks. Nature 410, 268 - 276 (2001).

4.2.2 Redes de mundo pequeno

Investigado pela primeira vez como uma questão em sociologia por Milgram [07], o fenômeno de mundo pequeno é o princípio no qual a maioria das pessoas estão conectados por cadeias de conexões curtas de conhecidos entre si. É uma característica de uma série de redes que surgem na natureza e tecnologia. Estudos sobre o fenômeno relevam duas propriedades fundamentais: que tais cadeias de conexões são onipresentes na rede e que indivíduos apenas com informações puramente locais estão aptos a encontrar tais cadeias. A principal característica emergente de redes deste tipo é o baixo valor de caminho médio em contraste com a alta quantidade de vértices presentes na rede.

Watts e Strogatz [08] introduzem o conceito de coeficiente de clusterização para explicar este baixo valor de caminho médio. Imagine que cada vértice numa determinada rede tenha K_v vizinhos mais próximos; Então, a Equação 6 indica o número máximo permitido de arestas M_v que podem existir entre eles. C_v (Equação 7) denota então a fração destas arestas permitidas atualmente, onde E_v é a quantidade de conexões existentes entre os vértices. A Equação 8 define então o coeficiente de clusterização CA para a rede de mundo pequeno como sendo a média de C_v sobre todos os vértices existentes na rede N .

$$M_v = \frac{k_v(k_v - 1)}{2} \quad (6)$$

$$C_v = \frac{E_v}{M_v} = \frac{2E_v}{k_v(k_v - 1)} \quad (7)$$

$$CA = \frac{1}{N} \sum_i^N \frac{2E_i}{k_i(k_i - 1)} \quad (8)$$

O resultado são redes muito mais clusterizadas em relação as componentes gigantes dos grafos aleatórios. As redes de interesse deste tipo devem ter alguns vértices com conexões longas (também chamadas de "pontes"), mas não tanto a ponto do grafo tornar-se desconectado. Assim, exibem sincronidade e poder computacional em comparação a redes do mesmo tipo que sejam do mesmo tamanho. Uma representação pode ser observada na Figura 4.3.

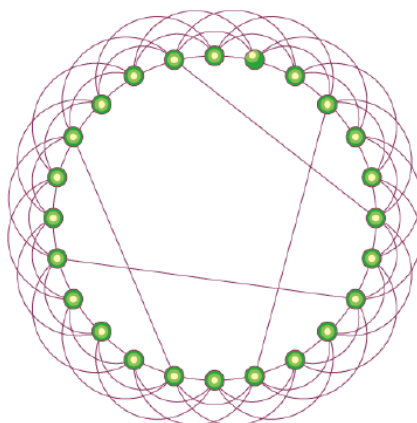


Figura 4.3: Modelo de uma rede de mundo pequeno. O modelo apresenta alta clusterização e algumas poucas conexões de longa distância.

Fonte: STROGATZ, S. H. Exploring Complex Networks. Nature 410, 268 - 276 (2001).

4.2.3 Redes livre de escala

Em busca de uma compreensão mais legítima dos sistemas complexos do mundo real, nos últimos anos, vários estudos têm identificado a presença de estruturas sem escala característica em uma quantidade cada vez maior de sistemas, sejam biológicos (rede de interação de proteínas ou de reações metabólicas), sociais (redes sociais), infraestrutura (redes de estradas, rotas marítimas e rotas aéreas), entre outros. Nestes, alguns poucos vértices possuem uma conectividade elevada em relação a maioria dos vértices presentes da rede. Esse tipo de topologia produz uma característica emergente nesses sistemas: suas distribuições de conectividades decaem em lei de potência. Esse é o ponto em que redes livres de escala diferem das redes aleatórias, que, como já comentado, possuem distribuição de graus do tipo Poissiana. E apesar destas serem amplamente estudadas para descrever e algumas estruturas de redes, falham em reproduzir muitas características observadas em redes reais.

O mecanismo que reproduz esse tipo de comportamento para redes livre de escala surge através dos trabalhos de Barabasi e Albert [10], impondo algumas regras para a evolução e crescimento desse tipo de rede, que será apresentado a seguir.

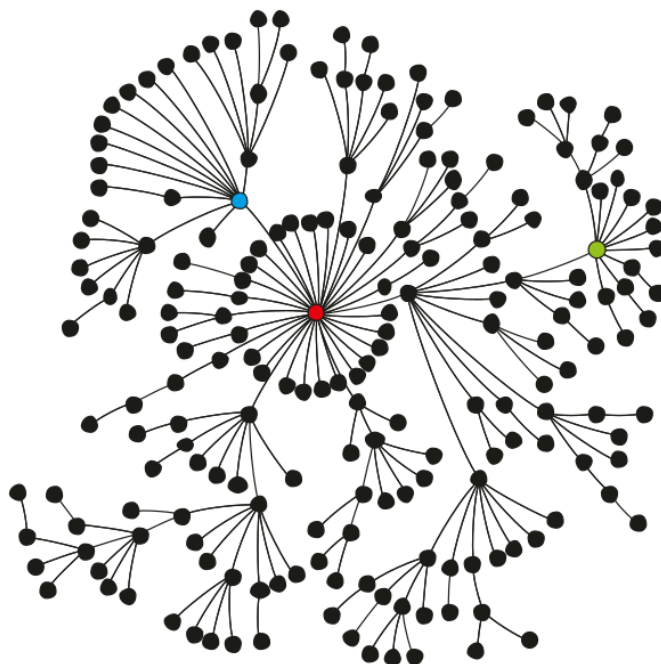


Figura 4.4: Modelo de uma rede livre de escala. Cresce através da adição de novos vértices, que se ligam a outros vértices preexistentes com probabilidade proporcional a conectividade destes.

Fonte: STROGATZ, S. H. Exploring Complex Networks. Nature 410, 268 - 276 (2001).

Modelo Barabasi - Albert :

A partir de estudos estruturais da *World Wide Web* (*www*), Barabási e Albert observaram que as páginas *Web* possuíam uma estrutura particular, de modo que sua distribuição de conectividade seguia uma lei de potência. Para modelar esse problema, estabelecem duas regras essenciais para a construção desse tipo de rede: **crescimento** e **ligação preferencial**. Essas duas regras são o coração do modelo proposto por eles. A Figura 4.4 representa uma rede deste tipo e a Figura 4.5 a distribuição de conectividades para as páginas *Web*. .

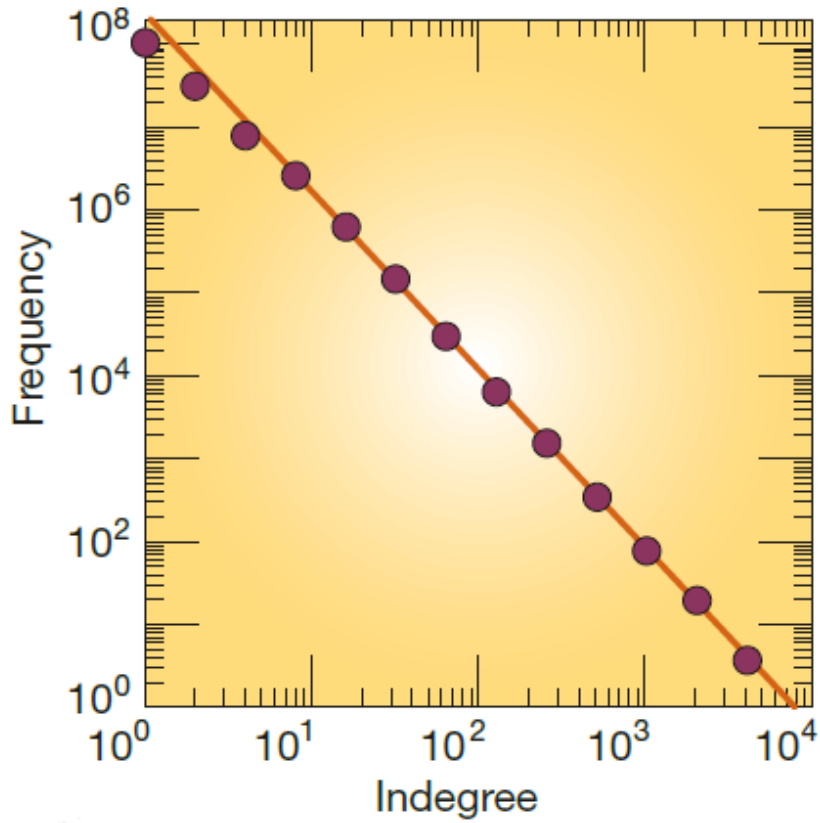


Figura 4.5: Distribuição de conectividades para páginas da *World Wide Web*. O gráfico log-log mostra o número de páginas da *web* que possuem um determinado grau de entrada (número de links de entrada). A cauda do distribuição segue uma lei de potência aproximada com expoente $\gamma \approx 2.2$.

Fonte: STROGATZ, S. H. Exploring Complex Networks. Nature 410, 268 - 276 (2001).

Enquanto nas redes aleatórias o número N total de vértices é fixo e as ligações entre estes são refeitas aleatoriamente, no modelo Barabási - Albert, ao observar que as redes reais não são estáticas, estas crescem anexando novos vértices aleatoriamente a outros previamente existentes (crescimento da rede). Porém, a probabilidade de ligação é proporcional a conectividade k_i do vértice de destino (ligação preferencial), seguindo a expressão descrita pela Equação 9, onde p_{k_i} é a fração de vértices na rede com conectividade de entrada k_i e o somatório no denominador é o grau médio da rede (que é não-dirigida), sendo m o número de conexões no qual o vértice que entra na rede irá realizar (com m constante) [13]; logo, vértices mais conectados tendem a se tornar mais conectados ainda, levando a formação de *hubs* e uma distribuição de graus enviesada com “cauda pesada” ou *heavy tailed*, seguindo uma lei de potência como na Equação 10, com expoente $\gamma = 3$. Os *hubs* representam a diferença estrutural mais notável entre uma rede aleatória e uma rede sem escala típica.

$$\frac{k p_{k_i}}{\sum_i k p_{k_i}} = \frac{K p_{k_i}}{2m} \quad (9)$$

$$P_k \sim k^{-\gamma} \quad (10)$$

A principal vantagem desse tipo de rede é a resistência a falhas aleatórias [11] pelo fato de alguns *hubs* dominarem a topologia. Por outro lado, a desvantagem é que esses mesmos *hubs* são vulneráveis a ataques deliberados.

Modelo Dorogovtsev - Mendes (DM) :

Vários outros modelos posteriores ao de Barabasi e Albert surgiram, os quais abrangem outras variações na ligação preferencial, mudando alguns parâmetros que influenciam sua evolução, de maneira a obter uma representação mais realista dos processos que ocorrem nas redes do mundo real. Dorogovtsev e Mendes [12] em seus estudos, sugerem uma variação do modelo. Evidências sugerem que na *World Wide Web* o grau médio de um vértice aumenta com o tempo, o que significa que o parâmetro m irá aumentar junto com a evolução da rede. Esta variação do modelo considera as seguintes características [13]:

- A conectividade média varia, com m crescendo com o tamanho n da rede como n^a , sendo a uma constante;
- A probabilidade de conexão para um dado vértice preexistente na rede segue a Equação 11 abaixo, onde B é também uma constante:

$$P_{k_i} = k_i + Bn^a \quad (11)$$

- A distribuição de conectividades resultante desse modelo segue uma lei de potência com expoente γ , dado pela seguinte expressão:

$$\gamma = 2 + \frac{B(1+a)}{1-Ba} \quad (12)$$

Desta maneira, o expoente da distribuição de graus γ não é mais fixo como no modelo original proposto por Barabasi e Albert, mas variável em termos dos parâmetros a e B , tornando possível um mecanismo alternativo no qual γ possa ser ajustado de maneira a reproduzir estruturas que são observadas em redes do mundo real.

Modelo Dorogovtsev, Mendes e Samukhin (DMS) :

Outra variação do modelo clássico de Barabasi e Albert também inspirado na *World Wide Web*, é proposto por Dorogovtsev, Mendes e Samukhin [17]. Neste, cada site (vértice) recebe uma atratividade (conectividade) inicial A , de maneira que este parâmetro governe a probabilidade de que vértices recém adicionados venham a receber novos links (conexões). As características desse modelo são as seguintes:

- A cada passo de tempo, um novo vértice é adicionado. Simultaneamente, m novas conexões direcionadas partindo de vértices não especificados são introduzidas;
- A probabilidade de que uma nova conexão aponte para um dado vértice i é:

$$P_i = \frac{A + k_i}{(A + m)N} \quad (13)$$

Onde $A_i = A + k_i$ é a conectividade do vértice i e $A_\Sigma = (A + m)N$, com $a = A/m$;

- A distribuição de conectividades dos vértices segue:

$$P(k, N) = \frac{\Gamma(am + k)}{\Gamma(am)\Gamma(k + 1)} \sum_{k=1}^N \left(\frac{k}{N}\right)^{\frac{am}{1+a}} \left[1 - \left(\frac{k}{N}\right)^{\frac{1}{1+a}}\right]^k \quad (14)$$

onde, no limite para $ma + q \gg 1$, a Equação 14 pode ser expressa como:

$$P(k) \cong (1 + a) \frac{\Gamma[(m + 1)a + 1]}{\Gamma(ma)} (q + ma)^{-(2+a)} \quad (15)$$

- O expoente γ da distribuição é dado por:

$$\gamma = 2 + \frac{A}{m} \quad (16)$$

5 Resultados e discussões

Aqui serão apresentados os resultados obtidos durante toda a pesquisa, inicialmente com a descrição do objetivo inicial (fase 1), que como já mencionado, foi o desenvolvimento de um algoritmo computacional em linguagem *Python* que reproduza redes do tipo sem escala típica (modelo clássico de Barabasi e Albert) e a análise da confiabilidade dos seus resultados. O segundo objetivo (fase 2) que consistia na obtenção de um algoritmo para uma variação do modelo de Barabasi e Albert será descrito posteriormente.

Fase 1 :

O código desenvolvido para o modelo clássico de Barabasi Albert foi baseado no algoritmo de Larissa de Farias Ribeiro [14]. O algoritmo para a construção da rede segue a sequência de passos abaixo:

1. Como condição inicial, a rede possui N_0 vértices conectados igualmente entre si;
2. A cada iteração t um novo vértice com M_0 arestas ($M_0 < N_0$) é adicionado a rede. Este vértice de entrada irá conectar-se a M_0 vértices diferentes preexistentes na rede, com probabilidade proporcional a conectividade k_i destes, dada pela expressão abaixo:

$$P_{k_i} = \frac{k_i}{\sum_j k_j} \quad (17)$$

Onde o denominador é o somatório das conectividades k_j dos vértices restantes na rede;

3. O passo 2 deve ser repetido N vezes (tamanho desejado da rede). Após t iterações, o resultado é uma rede com $N = N_0 + t$ vértices e $M_0(M_0 + 1)/2 + M_0t$ conexões.

A Figura 5.1 ilustra o intervalo de "tempo"(iteraões) de $t_0 = 0$ até $t = 4$. Apesar das M_0 ligaões a serem feitas pelo vértice de entrada a cada iteraão serem aleatórias, vértices mais antigos na rede tendem a serem mais favorecidos a receberem estas ligaões devido a presença da característica de ligaão preferencial, aumentando seus graus de conectividade a medida que a rede cresce. Note que pela equação 17, vértices que possuem maiores conectividades k_i , terão maior probabilidade a receber ligaões. Nesse processo, a grande maioria dos vértices irão possuir conectividades baixas, em contraste com alguns poucos que acabam obtendo conectividades muito elevadas. Surge então, como comentado em seões anteriores, a característica emergente mais importante deste tipo de rede: o aparecimento de *hubs*.

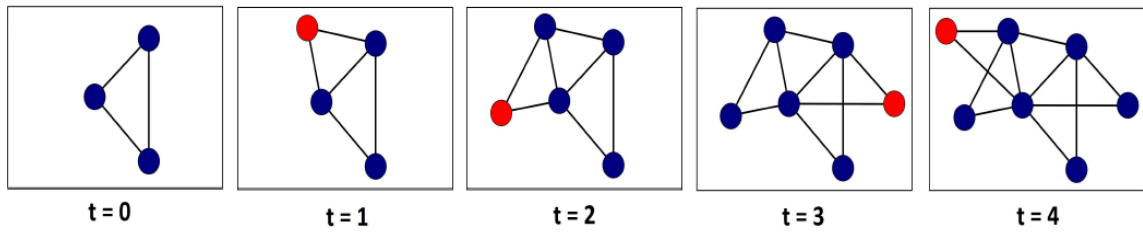


Figura 5.1: Mecanismo de construção da rede genérica ($N_0 = 3$ e $M_0 = 2$).

Fonte: RIBEIRO, L. de F. Redes sem Escala Típica: Visão Geral, Modelos Alternativos e Técnicas Computacionais. [S.l.]: UFRN, Janeiro, 2017.

As simulações realizadas a partir do algoritmo computacional desenvolvido foram feitas para os seguintes parâmetros: $M_0 = 3$, $N_0 = 4$ e $N = 250$, pois a região metropolitana do Recife (RMR) possui 246 zonas de tráfego (ZN), consideradas como vértices na rede genérica desenvolvida no trabalho de Vinicius Mendes Farah [15]. O valor de 250 foi escolhido de maneira que o algoritmo reproduza redes similares a da RMR. O histograma para distribuição de conectividades e a regressão linear feita para esta distribuição (Obtidos para 100 amostras) são mostrados nas Figuras 5.2 e 5.3, respectivamente.

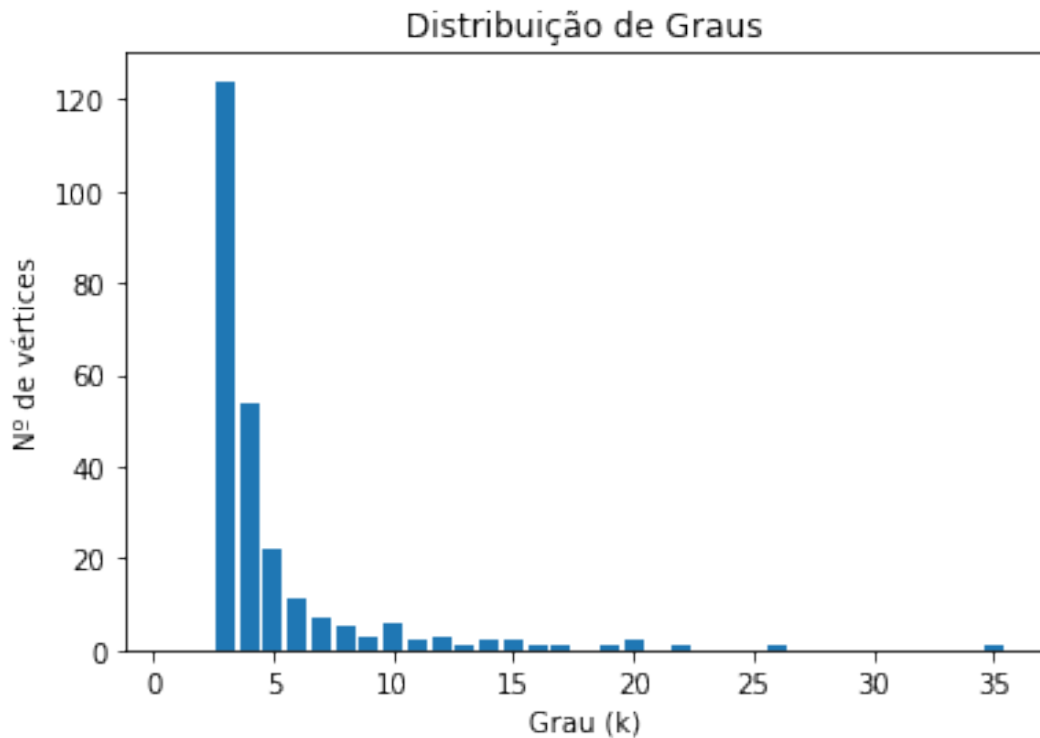


Figura 5.2: Histograma para distribuição de conectividades (graus) para rede genérica obtida.

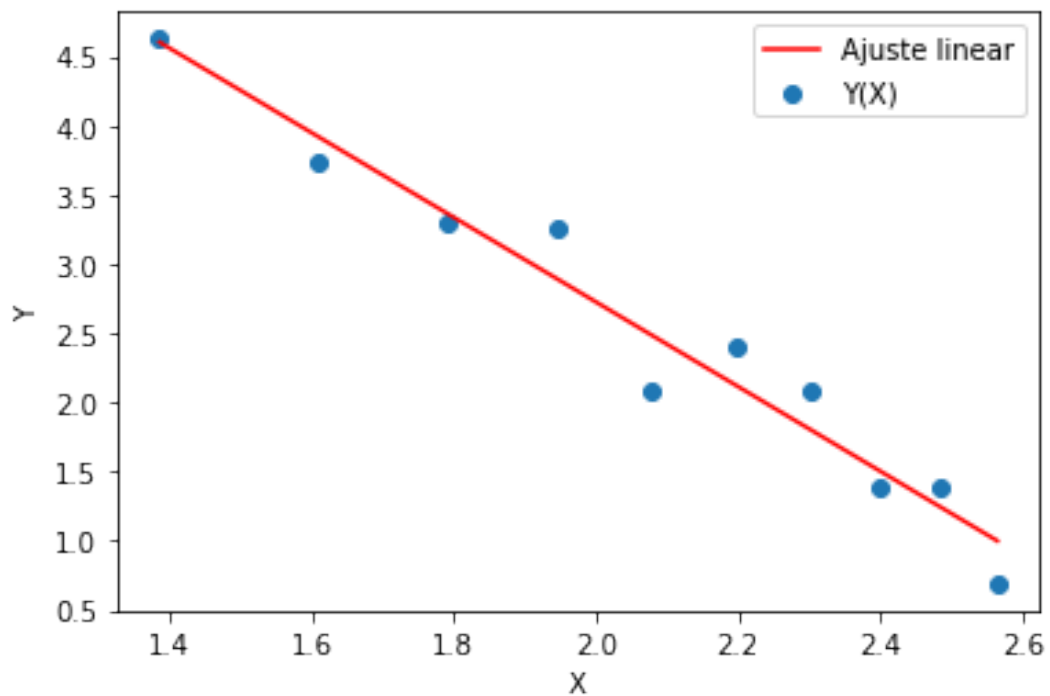


Figura 5.3: Plot $\log - \log$ para Regressão linear da distribuição conectividades (Y número de vértices, X conectividade).

Os resultados obtidos para a simulação foram:

- Valor médio do expoente da regressão linear (distribuição de conectividades em lei de potência) $\gamma \approx 3,1038$;
- Variância $\approx 0,0537$;
- Desvio padrão $\approx 0,2318$;
- Coeficiente de variação $\approx 7,46\%$.

A partir dos resultados, nota-se que o expoente da distribuição de conectividades γ é aproximadamente 3, como prevê o modelo proposto por Barabasi e Albert para a rede do tipo livre de escala clássica, com coeficiente de variação considerado baixo ($< 10\%$). Desta maneira, o objetivo inicial da pesquisa, que era obter um algoritmo confiável que reproduz redes deste tipo foi concluído.

Fase 2 :

Partimos do algoritmo desenvolvido na fase 1 para implementação do modelo de Dorogovtsev e Mendes [12] (descrito na seção 4.2), uma vez que seus resultados foram satisfatórios. A construção da rede segue o algoritmo abaixo:

1. O número m de ligações introduzidos a partir do novo vértice adicionado a cada iterada cresce com o tamanho N da rede com:

$$m(N) = 1 + fN^a \quad (18)$$

Onde o fator de proporcionalidade f garante que o número de ligações m feitas a cada ciclo seja sempre menor que o número de vértices presentes na rede;

2. A cada iteração, o vértice introduzido irá ligar-se a $m(N)$ outros preexistentes com probabilidade de conexão dada pela Equação 11;
3. O passo 2 deve ser repetido N vezes (tamanho desejado da rede).

Para avaliar a confiabilidade do algoritmo, reproduzimos os resultados obtidos por Dorogovtsev e Mendes em [12]. Seus resultados podem observados na curva inferior da Figura 5.4, onde o expoente da distribuição acumulada de conexões de entrada $\gamma = 1,243$ foi obtido para os parâmetros $a = 0,5$ e $B = 0,15$. Para os mesmos parâmetros a e B , tamanho da rede $N = 1000$ vértices, 100 amostras, $M_0 = 2$, $N_0 = 3$ e fator $f = 0,9$, os resultados obtidos pela simulação foram:

- Valor médio do expoente da regressão linear (distribuição de conectividades em lei de potência) $\gamma \approx 1,2457$;
- Variância $\approx 0,0075$;
- Desvio padrão $\approx 0,0868$;
- Coeficiente de variação $\approx 6,97\%$.

A curva obtida pela simulação para a distribuição acumulada pode ser observada na Figura 5.5.

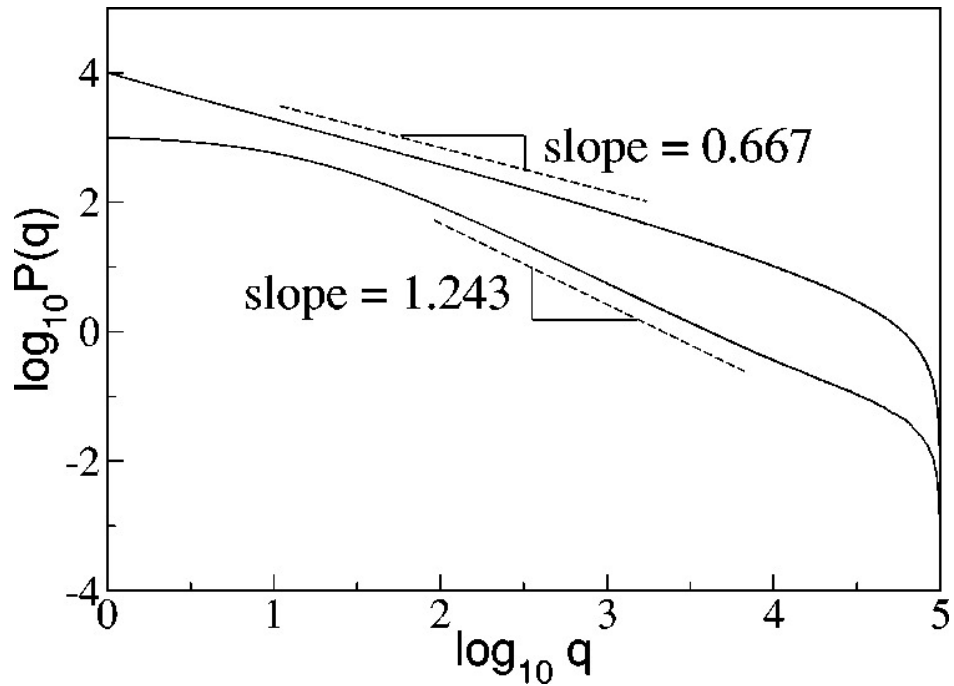


Figura 5.4: Distribuição acumulada de conectividades de entrada obtidos por Dorogovtsev e Mendes para 10.000 páginas da *World Wide Web*.

Fonte: Dorogovtsev, S. N. and Mendes, J. F. F. Effect of the accelerating growth of communications networks on their structure., Phys. Rev. E 63, 025101(R) (2001)

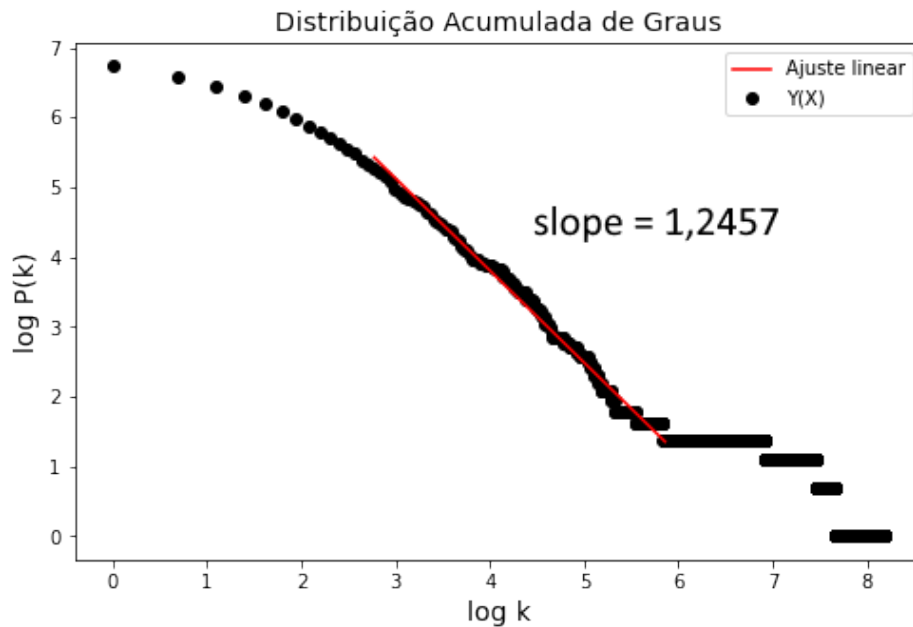


Figura 5.5: Plot $\log - \log$ para Regressão linear da distribuição acumulada de conectividades de entrada para o algoritmo Dorogovtsev-Mendes (DM).

Após garantir a confiabilidade do modelo desenvolvido, o passo seguinte foi obter o expoente da distribuição de conectividades das zonas de tráfego para região metropolitana do Recife ($\gamma = 6,96$) a partir da Equação 12. Os parâmetros utilizados foram $a = 0,99$, $B = 0,72$, tamanho da rede $N = 1000$ vértices, 100 amostras, $M_0 = 2$, $N_0 = 3$ e $f = 0,9$. Os resultados obtidos foram:

- Valor médio do expoente da regressão linear (distribuição de conectividades em lei de potência) $\gamma \approx 4.6953$;
- Variância ≈ 0.0784 ;
- Desvio padrão ≈ 0.2801 ;
- Coeficiente de variação $\approx 5.9659\%$.

Observa-se que o expoente da distribuição γ foi diferente do observado para a RMR de Recife. Devido ao fato de não existir uma expressão analítica para a distribuição de conectividades como função do tamanho da rede (número de vértices), não conhecíamos, portanto, o seu comportamento assintótico. Simulações feitas para averiguar como o expoente γ cresce com o tamanho N da rede podem ser observadas na Figura 5.6. Os parâmetros utilizados foram $a = 0,99$, $B = 0,72$, tamanho da rede variando de $N = 200$ até $N = 1700$ vértices, 100 amostras para cada N , $M_0 = 2$, $N_0 = 3$ e $f = 0,9$.

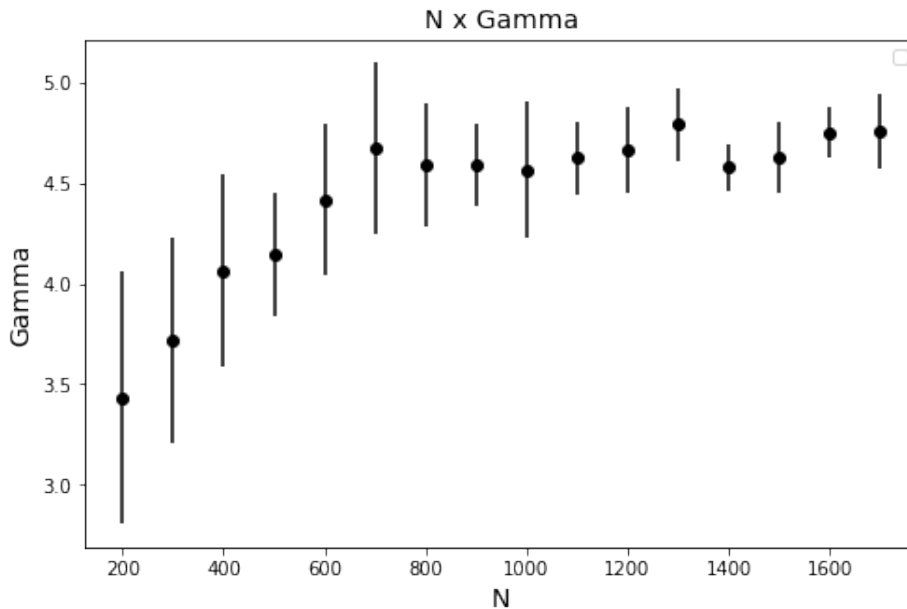


Figura 5.6: Crescimento do expoente γ em função do tamanho N da rede.

Nota-se que há um comportamento assintótico para $\gamma \approx 4,7$. Imagina-se que o expoente deva crescer mais a medida que N seja muito grande.

Diante disso, partimos para o estudo de uma outra variação do modelo clássico de Barabasi e Albert: o modelo DMS [17], descrito na seção 4.2.

A construção da rede segue o seguinte algoritmo:

1. Um novo vértice é adicionado a rede a cada passo de tempo. Simultaneamente, m ligações são feitas partindo de um vértice i (escolhido aleatoriamente);
2. Um novo vértice j é escolhido para receber uma das m ligações com probabilidade de conexão dada pela Equação 13;
3. O passos 1 e 2 devem ser repetidos N vezes (tamanho desejado da rede).

Buscamos, então, entender o comportamento inicial da distribuição de conectividades dada pela Figura 5.7, para o caso onde $m = 1$ (resultados analíticos obtidos por Dorogovtsev, Mendes e Samukhin).

Na Figura 5.8 apresentamos o comportamento das distribuições típicas de conectividades produzidos pelo algoritmo descrito acima. Os parâmetros utilizados foram $a = 1, a = 2, a = 3$ e $a = 4$ para $m = 1, N = 1000$ vértices, $M_0 = 2$ e $N_0 = 3$.

Observa-se que medida que a aumenta (ou seja, a atratividade inicial A aumenta, pois $a = A/m$), há uma maior convergência entre os resultados da simulação e o resultado analítico descrito pela Equação 15. Ou seja, as simulações realizadas acompanham o resultado analítico em um regime de baixa conectividade.

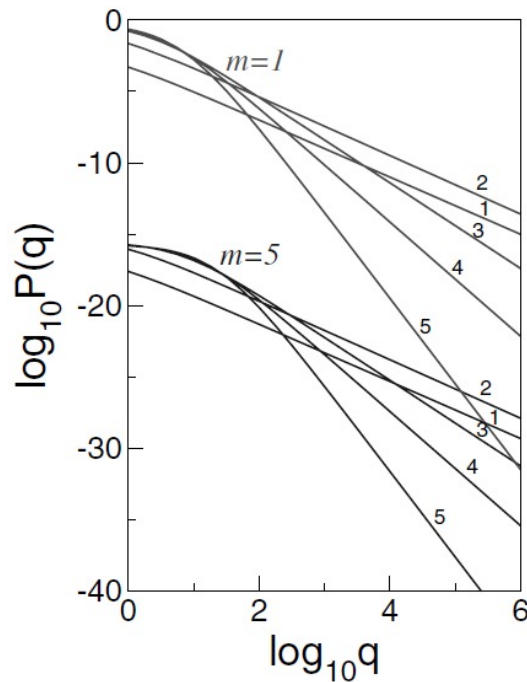


Figura 5.7: Distribuições $\log - \log$ para conexões de entrada para $m = 1$ e $m = 5$ obtidas por Dorogovtsev, Mendes e Samukhin.

Fonte: Dorogovtsev, S. N., Mendes, J. F. F. and Samukhin, A. N. Structure of Growing Networks with Preferential Linking., Phys. Rev. Lett., 85 (2001)

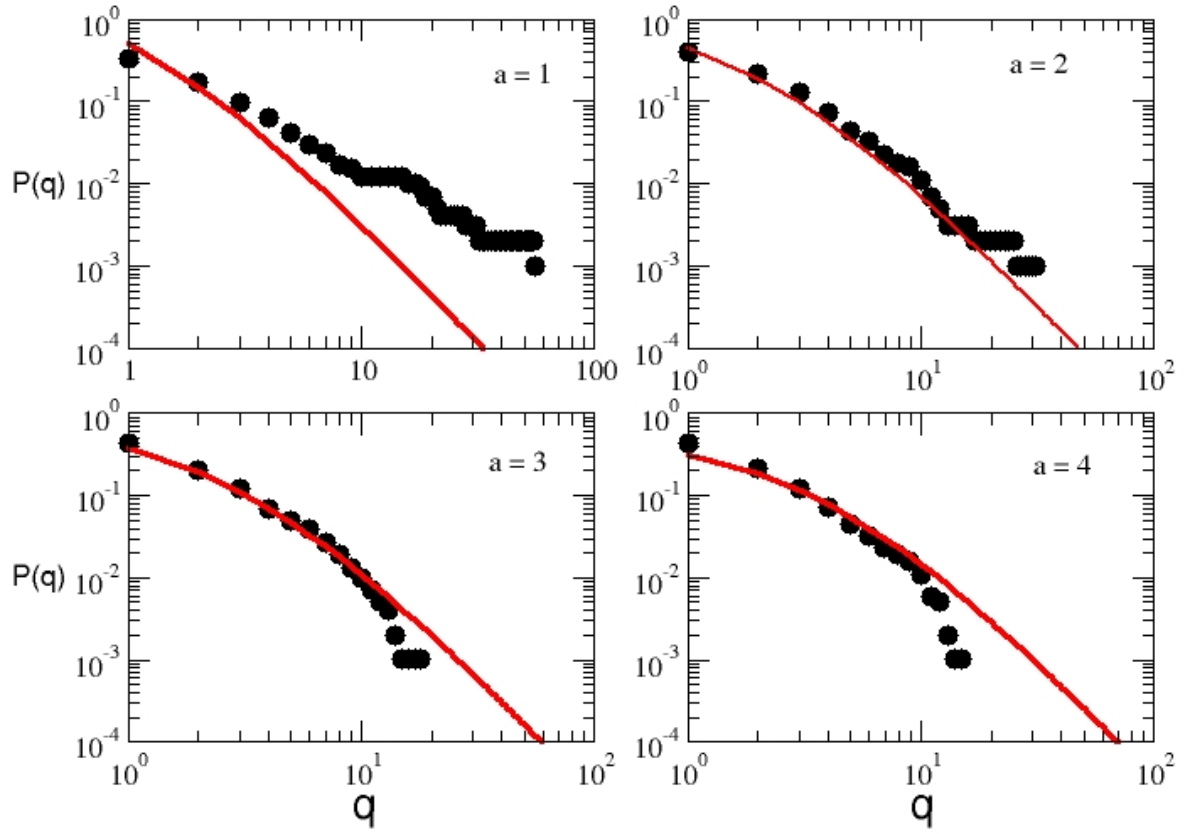


Figura 5.8: Distribuições $\log - \log$ para conexões de entrada para $m = 1$. Em vermelho a distribuição analítica da Equação 15 e em preto as simulações realizadas.

6 Considerações finais

O algoritmo inicial desenvolvido para o modelo de rede livre de escala (fase 1) mostrou-se confiável a partir dos resultados obtidos, e pôde servir como base para a fase 2 da pesquisa, onde o objetivo era buscar entender, a partir do modelo de Dorogovtsev e Mendes (DM) e Dorogovtsev, Mendes e Samukhin (DMS), como processos epidêmicos se comportam em redes livre de escala a depender de sua estrutura. Apesar da confiabilidade dos valores obtidos através do algoritmo desenvolvido na reprodução dos resultados de Dorogovtsev e Mendes (modelo DM) para o expoente de distribuição de conectividades obtidos por eles, e para os resultados do comportamento inicial no regime de baixa conectividade do modelo DMS, não tivemos tempo hábil para a implementação da dinâmica epidêmica numa rede compatível com a RMR de Recife. Com isso, acreditamos que as redes produzidas por esse algoritmo podem ser potencialmente utilizadas na investigação de processos epidêmicos. Todos os algoritmos desenvolvidos podem ser encontrados no GitHub: <https://github.com/Erike-Simon/PIC-UFRPE>.

7 Atividades desenvolvidas pelo aluno

Em fevereiro de 2022, se dá início ao semestre letivo denominado 2021.1, definido assim devido a influência dos períodos letivos excepcionais (PLE's) durante o pandemia de COVID-19. Durante o período de janeiro à setembro de 2022, o discente, cursou as disciplinas de Física Moderna I e Biofísica, já em regime híbrido, e vem cursando no semestre letivo 2021.2 as disciplinas de Física Moderna II, Eletrônica e Física, Música e Educação Sonora. Vêm desenvolvendo embasamento teórico através da leitura de artigos relacionados a Sistemas Complexos e suas influências em Processos Epidêmicos. Após um primeiro embasamento, inicia junto com seu orientador, a fase 1 da pesquisa, que consiste no desenvolvimento de um algoritmo computacional em linguagem de programação *Python* que reproduza redes do tipo sem escala, obtendo resultados e fazendo análise destes. Após a conclusão da fase 1, dá-se início a fase 2, consistindo na elaboração do algoritmo computacional para modelos mais generalizados de redes livre de escala. Em paralelo, realiza reuniões semanais presenciais com seu orientador, onde ocorrem apresentações dos resultados parciais da pesquisa, ajustes e planejamento para os próximos passos. Abaixo, seguem as médias obtidas nas disciplinas cursadas no período letivo 2021.1 e as notas das primeiras avaliações realizadas das disciplinas em andamento do período 2021.2.

Tabela 1: Disciplinas (2021.1)

Disciplina	Média
Física Moderna I	7,8
Biofísica	8,0

Tabela 2: Disciplinas (2021.2)

Disciplina	1 VA
Física Moderna II	9,0
Eletrônica	9,2
Física, Mús. e Edu. Sonora	9,3

8 Referências bibliográficas

- [01] Barabasi A.L. *Linked: How everything is connected to everything else and what it means for business, science and everyday life*, Basic Books (AZ), New York (2014).
- [02] Wilson, E. O. *Consilience: The Unity of Knowledge.*, Vintage Books USA, New York (1999).
- [03] Strogatz, S. H. *Exploring Complex Networks*, Nature 410, 268 - 276 (2001).
- [04] Barabasi, A.L. *Statistical mechanics of complex networks*, REVIEWS OF MODERN PHYSICS, VOLUME 74-47, JANUARY 2002.
- [05] Metz, J., Calvo, R., Seno, R. R. M., Roseli, A. F., Liang, Z. *Redes Complexas: conceitos e aplicações*, ISSN - 0103-2569, janeiro 2007.
- [06] Erdős, P. & Rényi, A. *Random Graphs*, Publicationes Mathematicae Debrecen, 6, 290-297 (1959).
- [07] Milgram, S. *The Small-World Problem.*, Psychology Today 1, 61-67 (1967).
- [08] Watts, D. & Strogatz, S. *Collective dynamics of 'small-world' networks*, Nature 406, 845 (1998).
- [09] Kleinberg, J. M. *Navigation in a small world*, Nature 393, 440-442 (2000).
- [10] Barabasi, A. L., Albert, R. *Emergence of scaling in random networks.*, Science. v.286, p. 509-512 (1999).
- [11] Albert, R., Jeong, H. Barabasi, A. L. *Error and attack tolerance of complex networks*, Nature 406, 378-382 (2000).
- [12] Dorogovtsev, S. N. and Mendes, J. F. F. *Effect of the accelerating growth of communications networks on their structure.*, Phys. Rev. E 63, 025101(R) (2001).
- [13] Newman, M. E. J. *The structure and function of complex networks.*, SIAM Review 45, 167-256 (2003).
- [14] Ribeiro, L. F. *Redes sem Escala Típica: Visão Geral, Modelos Alternativos e Técnicas Computacionais.*, UFRN, janeiro, 2017.
- [15] Farah, V. M. *Modelagem de processos epidêmicos na rede de transporte público da Região*

Metropolitana do Recife, UFRPE, setembro, 2021.

[16] Manchein, C., Brugnago, E. L., da Silva, R. M., Mendes, C. F. O., Beims, M. *Strong correlations between power-law growth of COVID-19 in four continents and the inefficiency of soft quarantine strategies.*, API Chaos 30, 041102 (2020).

[17] Dorogovtsev, S. N., Mendes, J. F. F. and Samukhin, A. N. *Structure of Growing Networks with Preferential Linking.*, Phys. Rev. Lett., 85 (2001).