



UNIVERSIDADE FEDERAL DO CEARÁ
CENTRO DE CIÊNCIAS
DEPARTAMENTO DE ESTATÍSTICA E MATEMÁTICA APLICADA
PROGRAMA DE PÓS-GRADUAÇÃO EM MODELAGEM E MÉTODOS
QUANTITATIVOS

CARLOS MIGUEL MOREIRA GONÇALVES

OTIMIZAÇÃO DE VACINAS EM REDES COMPLEXAS

FORTALEZA
2023

CARLOS MIGUEL MOREIRA GONÇALVES

OTIMIZAÇÃO DE VACINAS EM REDES COMPLEXAS

Qualificação de Mestrado apresentada ao Programa de Pós-Graduação em Modelagem e Métodos Quantitativos da Universidade Federal do Ceará, como requisito parcial para a obtenção do Título de Mestre em Modelagem e Métodos Quantitativos. Área de Concentração: Modelagem e Métodos Quantitativos.

Orientador: Prof. Dr. Leandro Chaves Rêgo.

Coorientador: Prof. Dr. Pablo Ignacio Fierens.

Dados Internacionais de Catalogação na Publicação
Universidade Federal do Ceará
Biblioteca Universitária

- S185r Gonçalves, Carlos Miguel Moreira .
 Otimização de Vacinas em Redes Complexas / Carlos Miguel Moreira Gonçalves. -
 2023.
 44 f. : il.
- Qualificação (mestrado) - Universidade Federal do Ceará, Centro de Ciências, PRO-
 GRAMA DE PÓS-GRADUAÇÃO EM MODELAGEM E MÉTODOS QUANTITATIVOS,
 Fortaleza, 2023.
 Orientação: Prof. Dr. Leandro Chaves Rêgo.
 Coorientação: Prof. Dr. Pablo Ignacio Fierens.
1. Redes; 2. Covid-19; 3. Palavra-chave3; 4. Palavra-chave4. I. Título.

CDD 530

CARLOS MIGUEL MOREIRA GONÇALVES

OTIMIZAÇÃO DE VACINAS EM REDES COMPLEXAS

Qualificação de Mestrado apresentada ao Programa de Pós-Graduação em Modelagem e Métodos Quantitativos da Universidade Federal do Ceará, como requisito parcial para a obtenção do Título de Mestre em Modelagem e Métodos Quantitativos. Área de Concentração: Modelagem e Métodos Quantitativos.

Aprovada em 07/03/2013.

BANCA EXAMINADORA

Prof. Dr. Leandro Chaves Rêgo (Orientador)
Universidade Federal do Ceará (UFC)

Prof. Dr. Pablo Ignacio Fierens (Coorientador)
Instituto Tecnológico de Buenos Aires (ITBA)

Prof. Dr. Sicrano de Tal
Universidade Federal do Piauí (UFPI)

“The properties of softness and hardness and darkness and clearness do not reside in the carbon atoms; they reside in the interconnections between the carbon atoms, or at least arise because of the interconnections between the carbon atoms.” - Nicholas Christakis

AGRADECIMIENTOS

A beltaro de tal bla bla bla bla bla bla bla blabla bla blabla bla blabla bla blabla
bla blabla bla bla bla bla bla bla bla bla bla blabla bla blabla bla blabla bla blabla
bla bla blabla bla blabla bla blabla.

A fulano de tal blabla bla blabla bla bla blabla bla blabla bla bla blabla bla blabla
bla bla blabla bla blabla bla bla blabla bla blabla bla blabla.

RESUMO

A pandemia do Covid-19 foi um episódio adverso na história da humanidade causando um estrago na economia e na saúde mental de várias pessoas. Isso se deve muito ao perigo da doença antes da produção de vacinas e por causa de um isolamento forçado dentro de casa. Apesar de necessário, o isolamento social contribuiu para o agravamento desse cenário, colocar todas as pessoas em isolamento criou um ambiente de baixa produtividade, seja no Mercado de Trabalho e na Educação, por exemplo. Para evitar a quebra total, as instituições optaram pela utilização dos chamados *home-office*, contudo ninguém estava preparado para uma mudança tão drástica na forma de viver. Será que o isolamento total de todos é necessário?

Nesse trabalho, iremos estudar a remoção de indivíduos em redes levando em conta a sua importância na topologia e também considerando a taxa de mortalidade pelo Covid-19 de cada pessoa advinda da sua idade.

Palavras-chave: Redes; Covid-19; Palavra-chave3; Palavra-chave4.

ABSTRACT

Write your abstract here bla bla bla bla bla bla bla bla blabla bla blabla bla blabla bla blabla bla blabla bla bla.

Keywords: Keyword1; Keyword2; Keyword3; Keyword4.

LISTA DE FIGURAS

| | |
|--|----|
| Figura 1 – Pontes de Königsberg | 15 |
| Figura 2 – Ilustração de Redes e seus Parâmetros | 17 |
| Figura 3 – Ilustração de Redes e suas Métricas Globais | 19 |
| Figura 4 – Ilustração do Modelo | 23 |
| Figura 5 – Funcionamento do Modelo de Configuração, escolhemos dois sítios i e j ale- atoriamente e os conectamos. Fonte: Elaborado pelo autor | 23 |
| Figura 6 – Modelo SI | 25 |
| Figura 7 – Número de Contatos por mês | 27 |
| Figura 8 – Gráfico do Número de Contatos por Idade | 28 |
| Figura 9 – Gráfico do Número de Contatos por Idade | 29 |
| Figura 10 – Fração média de ligações entre faixas etárias | 30 |
| Figura 11 – Proporção de cada faixa etária no Banco de dados. | 31 |
| Figura 12 – Ilustração do Modelo de Manzo | 32 |
| Figura 13 – Funcionamento do Modelo de Configuração, escolhemos dois sítios i e j ale- atoriamente e os conectamos. O Algoritmo pode gerar várias topologias de redes, porém ainda limitadas pelas quantidades $\{k_i\}$ de graus impostas pelo MC. Fonte: Elaborado pelo autor | 32 |
| Figura 14 – Ilustração do Modelo SEIHARD | 36 |
| Figura 15 – Esquematização do Modelo SEIHARD de contágio. Os blocos em vermelho significam quem pode contaminar outros, os blocos em azul escuro quem não é infectado e em azul claro quem está isolado da rede. | 36 |

LISTA DE GRÁFICOS

LISTA DE TABELAS

| | |
|--|----|
| Tabela 1 – Tabela do resultado de Manzo | 33 |
| Tabela 2 – Nessa tabela são apresentados resultados advindos do autor para a rede francesa feitas em 100 simulações de redes. Apresentando as características da rede e em parêntese o erro associado a cada valor. Fonte: Elaborada pelo autor. | 33 |
| Tabela 3 – Tabela do meu resultado. | 34 |
| Tabela 4 – Nessa tabela são apresentados resultados advindos do meu código na qual foram feitas 1000 simulações de redes para a rede francesa que se aproximam bastante do apresentado no artigo junto com os desvios padrões entre parênteses. Os que não bateram certo estão na faixa de erro mostrada anteriormente. Fonte: Elaborada pelo autor. | 34 |
| Tabela 5 – Tabela de parâmetros do modelo. | 36 |
| Tabela 6 – Nessa tabela são apresentados os valores de α^i , que é a probabilidade de se tornar sintomático pela idade. Fonte: Elaborada pelo autor. | 37 |
| Tabela 7 – Nessa tabela são apresentados os valores de ψ^i , que é a probabilidade de um infectado ser hospitalizado. Fonte: Elaborada pelo autor. | 37 |
| Tabela 8 – Nessa tabela são apresentados os valores de τ^i , que é a probabilidade de um hospitalizado morrer. Esses dados foram coletados a partir do Open Data SUS do ano de 2020. Fonte: Elaborada pelo autor. | 37 |

LISTA DE SIGLAS

| | |
|-------|---|
| ABNT | Associação Brasileira de Normas Técnicas |
| CNPq | Conselho Nacional de Desenvolvimento Científico e Tecnológico |
| TESTE | bla bla bla bla bla bla blabla bla blabla bla blabla bla blabla bla bla bla bla bla bla bla blabla bla blabla bla blabla blablabla bla blabla bla bla |

LISTA DE SÍMBOLOS

| | |
|-----------------|--------------------------------|
| $\eta_{\mu\nu}$ | Métrica do espaço de Minkowsky |
| $g_{\mu\nu}$ | Métrica do espaço curvo |
| x | bla bla |

SUMÁRIO

| | | |
|--------------|---|-----------|
| 1 | INTRODUÇÃO | 14 |
| 1.1 | Contextualização da Problemática | 14 |
| 2 | INTRODUÇÃO À TEORIA DE REDES | 15 |
| 2.1 | Conceitos Fundamentais | 16 |
| 2.2 | Métricas de Redes | 17 |
| 2.2.1 | <i>Métricas Globais de Redes</i> | <i>17</i> |
| 2.2.2 | <i>Métricas de Centralidade</i> | <i>19</i> |
| 2.3 | Modelos de Formação de Redes | 21 |
| 2.3.1 | <i>Rede de Erdős-Rényi</i> | <i>21</i> |
| 2.3.2 | <i>Modelo de Configuração</i> | <i>22</i> |
| 2.4 | Modelos de Infecção ou Difusão | 23 |
| 2.4.1 | <i>Modelo SI</i> | <i>24</i> |
| 2.4.2 | <i>Modelo SIS</i> | <i>25</i> |
| 2.4.3 | <i>Modelo SIR</i> | <i>25</i> |
| 2.4.4 | <i>Modelos em Redes</i> | <i>25</i> |
| 3 | METODOLOGIA | 26 |
| 3.1 | Análise dos Dados | 26 |
| 3.2 | Modelo de Redes | 32 |
| 3.3 | Modelo de Infecção | 35 |
| 4 | CONCLUSÃO | 38 |
| | REFERÊNCIAS | 39 |
| | APÊNDICE A – TESTANDO O APÊNDICE | 42 |
| | APÊNDICE B – TESTE2 | 43 |
| | ANEXO A – TESTANDO O ANEXO | 44 |

1 INTRODUÇÃO

Desde as mais antigas civilizações humanas, elas têm enfrentado o problema de propagação de doenças em larga escala [1]. Em 430 a.C aconteceu a epidemia chamada Peste de Atenas que foi responsável pela morte de cerca de 1/3 da população de Atenas. Em 541 houve a primeira pandemia chamada Praga de Justiniano que ocorreu no mediterrâneo, em 1347 aconteceu a mais devastadora pandemia na história da humanidade, a Peste Negra. Mais recentemente tivemos a Gripe Suína em 2009 e recentemente a Covid-19 e a Sétima Pandemia da Cólera. Nesse sentido o estudo de infecções se tornou cada vez mais necessário para cientistas seja para entender como uma infecção afeta o nosso corpo, seja para modelar a propagação dela.

Outrossim, com a expansão da humanidade nos últimos anos a partir do comércio, desmatamento e turismo facilitou a interação entre humanos e entre humanos e animais. Isso favoreceu uma maior propagação de doenças entre as civilizações [2]. Essa propagação tem 4 classificações possíveis de acordo com a taxa de contágio e a sua área de atuação [3].

- **Endemia** significa que uma infecção tem taxa de contágio controlada e previsível que atua desde uma cidade até um continente;
- **Surto** expressa um aumento repentino na ocorrência de casos da doença em pequenas áreas;
- **Epidemia** é um surto em grande escala;
- **Pandemia** é uma epidemia em escala mundial.

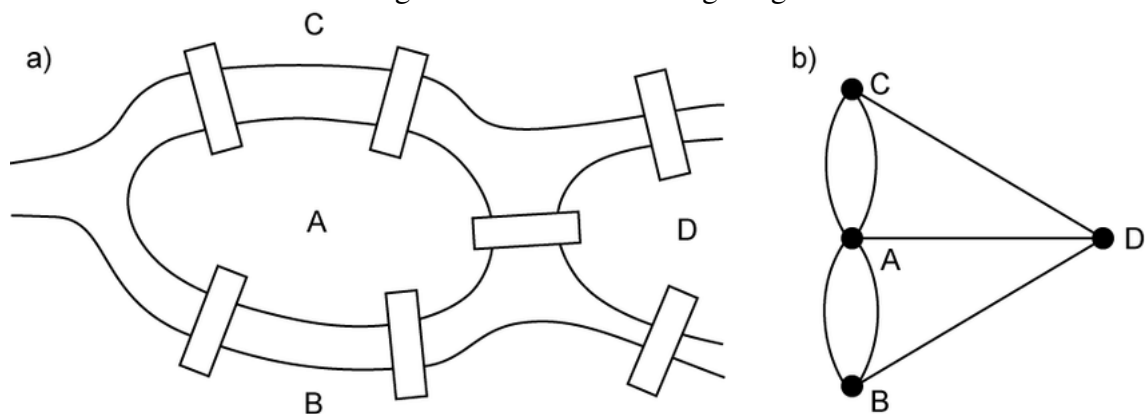
1.1 Contextualização da Problemática

2 INTRODUÇÃO À TEORIA DE REDES

A nossa principal ferramenta para modelar o nosso problema surgiu em 1736. Nessa época existia a cidade de Königsberg (atual Kaliningrado), nela passava o Rio Prególia que separava a cidade em 4 partes. Para se caminhar livremente pelas 4 regiões foram construídas 7 pontes que ligavam cada região. Isso gerava uma dúvida intrigante entre os moradores: seria possível sair de uma região e voltar para ela passando por todas pontes apenas uma vez?

Leonhard Euler [4] se interessou pelo assunto e tentou resolver esse problema. Para isso ele criou uma estrutura chamada Grafo que era composto por pontos (nós) que representavam as regiões e algo ligando entre eles (ligações) que representavam e ignorou toda a forma geométrica de cada um desses objetos. Euler percebeu que para que haja esse caminho é necessário e suficiente que todos os nós tenham um número de ligações pares. Pois para haver uma solução deve existir um caminho de ida e um de volta para cada vértice, como isso não é verdade para a cidade de Königsberg, então o problema não tem solução.

Figura 1 – Pontes de Königsberg



Fonte: Boguslawski, Pawel [5].

A prova de Euler nos mostra que quando queremos modelar algum problema não é necessário considerar todas as variáveis existentes neles, mas o essencial para a solução. Nesse caso o mais importante foi esquematizar o problema usando Grafos e, principalmente, analisar uma estrutura intrínseca ao Grafo.

Com a solução do problema surge a área da matemática chamada Grafos que é a base matemática para o que chamamos em redes. Essa nomenclatura varia de área para área, na física, por exemplo, são sinônimos enquanto que na computação Grafos estão relacionados à problemas de fluxo [6, 7] e Redes são utilizados para problemas visando a estrutura do Grafo e suas interações [8, 9]. Nesse trabalho usaremos os dois como sinônimo.

2.1 Conceitos Fundamentais

Um Grafo $G(\mathcal{N}, \mathcal{L})$ é uma dupla na qual $\mathcal{N} = \{0, 1, 2, \dots, i, \dots\}$ é um conjunto não vazio de elementos chamados de vértices, ou nós, e \mathcal{L} é um conjunto não vazio de pares não ordenados de elementos de \mathcal{N} . Dessa forma podemos definir $N = |\mathcal{N}|$ que mensura a quantidade de vértice que há em G e $L = |\mathcal{L}|$ que mensura o número de arestas. Essas quantidades também são chamadas de ordem e tamanho [7], respectivamente, que medem a cardinalidade de \mathcal{N} e \mathcal{L} .

A partir dessa definição de redes podemos definir a Matriz de Adjacência, ela vai guardar a informação de quem está conectado com quem. Seja uma matriz A quadrada de tamanho $N \times N$, cada elemento da matriz segue a seguinte regra:

$$A_{i,j} = \begin{cases} 1 & \text{se } i \text{ e } j \text{ estão conectados} \\ 0 & \text{caso contrário.} \end{cases}$$

A partir dela surgem características importantes de redes. Se $A_{i,j} = A_{j,i} \forall i, j \in \mathcal{N}$ então a rede é chamada de não direcionada, se $a_{i,j} \neq a_{j,i} \forall i, j \in \mathcal{N}$ ela é chamada de direcionada. Isso pode ter diferentes interpretações a depender do contexto inserido. Por exemplo na rede de amigos do Facebook, se o usuário A é amigo de B, então B é amigo de A. No caso do Instagram se A segue B não necessariamente B segue A. No primeiro caso é natural modelarmos utilizando redes não direcionadas, enquanto que no segundo usamos redes direcionadas.

Uma outra forma de guardamos a informação de quem está conectado com quem é a partir da lista de vizinhos. Dois vértices i e j são vizinhos se existe uma ligação entre i e j , assim denotamos $v(i)$ o conjunto de vizinhos do vértice i para redes não direcionadas. Em redes direcionadas agora temos dois tipos de vizinhos (i, j) e (k, i) , portanto denotaremos $v_{in}(i)$ todos os vizinhos que têm ligação que chega em i e $v_{out}(i)$ todos os vizinhos que têm ligação que sai em i .

Por fim, podemos adicionar pesos às nossas ligações, isso têm várias interpretações nos problemas: em uma rede de canos de esgoto, o fluxo que passa em um cano pode ser um peso e em redes de Facebook o nível de amizade pode ser um peso das ligações entre usuários, dentre outros exemplos. Ou podemos adicionar aos vértices como em redes sociais se quisermos estudar a propagação de um vírus a idade ou se ela faz parte de um grupo de risco ou não são parâmetros importantes a serem analisados. Formalmente, um Grafo ponderado $G(\mathcal{N}, \mathcal{L}, \omega)$ é formado por um conjunto de Nós \mathcal{N} , um conjunto de arestas \mathcal{L} e um mapeamento $\omega : \mathcal{L} \mapsto \mathbb{R}$ e/ou $\omega : \mathcal{N} \mapsto \mathbb{R}$. Nesse sentido podemos definir a matrix de pesos W que apresenta uma definição parecida da Matriz de Adjacência, seja T um valor limiar então:

$$w_{i,j} = \begin{cases} \omega(i, j) & \text{se } i \text{ e } j \text{ estão conectados} \\ 0 & \text{caso contrário.} \end{cases}$$

Figura 2 – Ilustração de Redes e seus Parâmetros

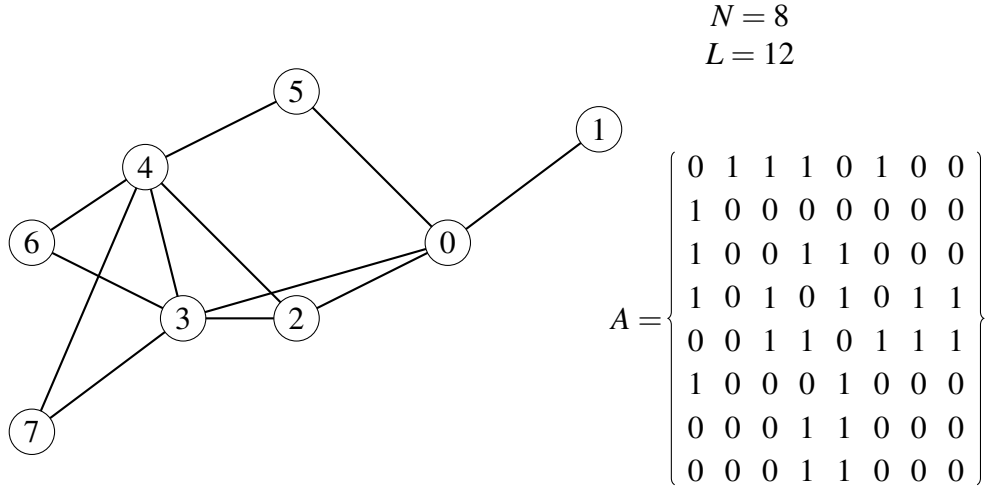


Ilustração de uma rede, das suas quantidades de ligações, de nós e da Matriz de Adjacência.

Fonte: Elaborado pelo autor

2.2 Métricas de Redes

Como discutido anteriormente é muito importante estudarmos a estrutura da rede, pois elas vão governar o comportamento da rede. Para entendermos melhor essa estrutura existem várias métricas [10] relacionadas à topologia da rede podendo ser classificadas como globais ou locais. As métricas globais são aquelas que estão intrinsecamente ligadas ao Grafo, enquanto que as locais estão ligadas à cada vértice da rede.

2.2.1 Métricas Globais de Redes

A primeira delas a ser analisada é a medida de distância entre um vértice e outro. Seja dois vértices $i, j \in \mathcal{N}$ dizemos que existe um caminho entre i e j se existe uma sequência de vértices $v_1, v_2, \dots, v_k | (v_s, v_{s+1}) \in \mathcal{L} \forall 1 \leq s < k$ e $v_1 = i, v_k = j$. Existem vários caminhos que podem existir entre i e j , portanto a nossa distância geodésica entre dois vértices $l(i, j)$ é definida como a quantidade de arestas do menor caminho entre i, j .

Agora que definimos qual é a distância em redes que usaremos, outras métricas surgem a partir dessa definição. Por exemplo temos o menor caminho médio de toda a rede:

$$\langle l \rangle = \sum_{\substack{i, j \\ i \neq j}} \frac{l(i, j)}{N \cdot (N - 1)}. \quad (2.1)$$

Esse valor nos dá a informação sobre o quanto em média é necessário caminhar na rede de um nó para outro. Outro valor importante é o chamado diâmetro d que é a maior distância geodésica da rede

$$d = \max_{i, j \in \mathcal{N}} l_{i, j} \quad (2.2)$$

Esses valores são bastante importantes para estudos em redes de pequeno mundo [11] na qual discutiremos nas próximas sessões.

Uma outra métrica a ser avaliada é quantas conexões cada nó contém. Definiremos o grau k_i do nó i como sendo o número de conexões (ou vizinhos) que o nó i possui para redes não direcionadas. Para redes direcionadas teremos outra divisão entre k_i^{in} e k_i^{out} para os sítios que pertencem a $v_{in}(i)$ e $v_{out}(i)$. Podemos achar esses valores a partir da Matriz de Adjacência:

$$k_i^{in} = \sum_j A_{ij} \quad k_i^{out} = \sum_j A_{ji}. \quad (2.3)$$

Além disso podemos também definir o grau médio $\langle k \rangle$ da rede esse valor tem bastante importância quantitativa quando queremos trabalhar com distribuições de graus.

$$\langle k \rangle = \sum_{i \in G} \frac{k_i}{N} = 2 \frac{L}{N} \quad (2.4)$$

Outras duas métricas importantes são a densidade $\rho(G)$ e a reciprocidade $rc(G)$. A primeira mede a razão entre a quantidade de arestas dentro do grafo G pela quantidade total de arestas que o grafo pode suportar. Já a segunda nos apresenta a fração de arestas que existem em ambas as direções, no caso trivial de redes não direcionadas esse valor é igual a 1.

$$\rho(G) = \begin{cases} \frac{2L}{N \cdot (N-1)} & \text{se a rede for não direcionada, ou} \\ \frac{L}{N \cdot (N-1)} & \text{se a rede for direcionada.} \end{cases}$$

$$rc(G) = \frac{\sum_{i,j} A_{i,j} A_{j,i}}{\sum_{i,j} A_{i,j}} \quad (2.5)$$

Por fim temos uma métrica para analisar o Agrupamento de uma Rede, ou seja quanto que os vértices estão unidos. Essa métrica será definida a partir da probabilidade de um vizinho do nó i estar conectado com outro vizinho de i . Existe a definição desse valor a partir da Matriz de Adjacência pela 2.6, porém é mais compreensível pela 2.7

$$C(G) = \frac{\sum_{(i,j,k): i \neq j \neq k} A(i,j)A(j,k)A(k,j)}{\sum_{(i,j,k): i \neq j \neq k} A(i,j)A(j,k)}, \quad (2.6)$$

$$C(G) = 3 \frac{\#Triangulos}{\#Triades}. \quad (2.7)$$

Em que $\#$ significa número de alguma coisa, no caso estamos contando o número de triângulos e o número de tríades que aparecem na rede. Essa é a definição do que chamamos de Agrupamento Global, podemos também definir o Agrupamento Local 2.8 e pro ele o Agrupamento Médio 2.9.

$$C(i) = 2 \frac{\#Triangulos(i)}{\#Triades(i)}. \quad (2.8)$$

$$\langle C \rangle = \frac{\sum_i C(i)}{N} \quad (2.9)$$

Figura 3 – Ilustração de Redes e suas Métricas Globais

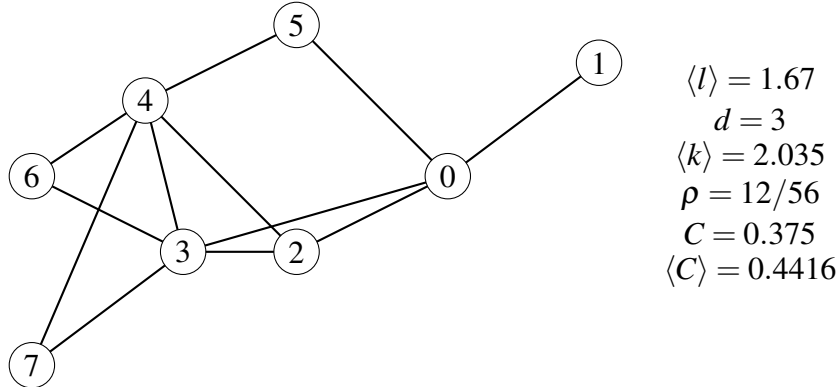


Ilustração de uma rede e sua respectivas Métricas Globais: menor caminho médio, diâmetro, grau médio, densidade, Agrupamento, Agrupamento Médio.

Fonte: Elaborado pelo autor

2.2.2 Métricas de Centralidade

Ao estudar redes estamos interessados em entender suas estruturas e propriedades, isso já é possível dadas as métricas que apresentamos anteriormente. Contudo, essas métricas globais nos trazem informações sobre a rede como um todo e as propriedades que ela apresenta podem estar diluídas na rede como um todo ou concentrada em sítios específicos. Portanto estudar os sítios em si e a sua importância perante a rede é necessário, para isso usaremos as chamadas Métricas de Centralidade.

Centralidade nesse contexto está ligado à importância, quando falamos de Métricas de Centralidade queremos estudar a importância topológica de cada nó perante a rede. Mas em que sentido seria essa importância? Depende do problema, talvez seja necessário estudar qual o autor mais importante em uma rede de citações de artigos ou qual pessoa devemos retirar de uma rede social para inibir a propagação de *fake-news* para ambas existe uma métrica apropriada para atingir tais objetivos.

A primeira métrica a ser avaliada é propriamente o número de conexões de cada sítio, o grau k_i . Ao avaliarmos essa métrica consideramos o sítio central aquele que tem o maior k_i , ou seja estamos atrás daquele sítio que apresenta mais ligações.

A segunda métrica importante é a Centralidade de Excentricidade (CE). Excentricidade está definido no dicionário como "desvio ou distância para o centro", no nosso caso a CE vai apontar qual sítio é mais central baseado na maior distância que o sítio pode percorrer na rede, sendo ele o mais central aquele que tem a menor distância máxima. Formalmente

$CE(i) \forall i \in \mathcal{N}$ é definida como:

$$CE(i) = \frac{1}{\max_{\forall j \in \mathcal{N}} l_{i,j}} \quad (2.10)$$

Nessa mesma ideia podemos avaliar outras duas métricas parecidas: Centralidade de Proximidade (CP) e Centralidade Harmônica (CH). Em ambas estamos interessados a entender qual é o nó mais central a partir de quanto ele dista dos outros, a diferença entre as duas é que a primeira é calculado pelo inverso da média aritmética e a segunda com a média harmônica.

$$CP(i) = \frac{N-1}{\sum_{j \neq i} l_{i,j}} \quad (2.11)$$

$$CH(i) = \frac{1}{N-1} \sum_{j \neq i} \frac{1}{l_{i,j}} \quad (2.12)$$

Uma generalização da CP é a Centralidade de média- p (C_p), a ideia desta medida de centralidade é usar a noção de média generalizada das distâncias. Ela é definida como:

$$C_p(i) = \begin{cases} \left(\frac{\sum_{j \neq i \in \mathcal{N}} l(i,j)^p}{N-1} \right)^{-\frac{1}{p}} & \text{se } p \neq 0 \\ \left(\prod_{j \neq i \in \mathcal{N}} l(i,j) \right)^{-\frac{1}{N-1}} & \text{se } p = 0 \end{cases} \quad (2.13)$$

Entretanto a importância de um nó não está associado somente a sua distância entre os outros nós da rede, existem aqueles que são importantes por manterem conexões entre nós. Por exemplo, seja a rede mundial aeroportos que seja construída a partir de nós como sendo os aeroportos e as ligações sendo as viagens entre aeroportos. Nessa situação, imagine que tenha um país altamente conectado entre si, contudo ele só tem um aeroporto que liga esse país ao exterior, esse nó é importante não pela distância dele para os outros aeroportos mas porque ele é o principal e único intermédio entre exterior e o país.

Para ser possível detectar e mensurar esse tipo de importância, considere $\mathbb{L}_{u,v}$ o número de mínimos caminhos que vão de u para v e $\mathbb{L}_{u,v}^i$ o número de mínimos caminhos que vão de u para v e passam pelo vértice i . Portanto, podemos definir a Centralidade de Intermediação (CI) como

$$CI(i) = \frac{1}{(N-1)(N-2)} \sum_{u \neq k} \frac{\mathbb{L}_{u,v}^i}{\mathbb{L}_{u,v}} \quad (2.14)$$

Ao lidarmos com uma rede de uma empresa sabemos que o CEO é muito importante para ela, contudo ele tem contato apenas com os chefes de cada região da empresa e esses chefes não precisariam ter contato com o chefe para ter contato entre si. Então onde estaria a informação de que o CEO é o mais importante? A resposta para essa pergunta está no fato de que o CEO têm ligações com pessoas de importância para rede, algo que não é detectado pelas

métricas anteriores Definimos a Centralidade de Autovetor (CA) como:

$$CA(i) = \frac{1}{\lambda} \sum_j^n A_{i,j} CA(j) \quad (2.15)$$

Na qual λ é o maior auto-vetor associado a matriz de Adjacência de Rede.

Quando analisamos sites e a sua importância é que essa está ligada com quantas vezes eles são citados por outros sites. Para quantificar isso a Google criou o *Page Rank* (PR) [12]. As métricas anteriores são definidas em redes tanto direcionadas quanto não direcionadas, contudo esta só tem sentido quando tratamos de redes direcionadas. Podemos então definir o *Page Rank* como:

$$PR(i) = (1 - \alpha) + \alpha \sum_{j \neq i} \frac{A_{j,i} PR(j)}{k_{out}(j)} \quad (2.16)$$

Na qual α é um parâmetro de amortecimento.

2.3 Modelos de Formação de Redes

Na sessão anterior discutimos sobre as métricas que tentam quantificar a estrutura da rede, nesse sentido dada um grafo G conseguimos extrair as suas propriedades. Contudo, será que seria possível fazer o contrário? Dado uma propriedade conseguimos criar uma Rede que contenha essas propriedades? Portanto, nessa sessão explicaremos os principais modelos utilizados para formarmos redes artificiais.

2.3.1 Rede de Erdős-Rényi

Um dos primeiros modelos que conseguiu ter sucesso nesse tocante foi o de Erdős-Rényi que iniciou os estudos de grafos (ou redes) aleatórias[13]. O modelo é considerado estocástico com tempo discreto partindo do pressuposto que as conexões entre os sítios é algo aleatório e conseguimos retirar propriedades importantes da rede. Apesar de simples e poucas dessas propriedades aparecem em redes reais, é interessante estudar esse modelo por motivos históricos e também podemos aprender muito com ele apesar de simples.

Seja $G(\mathcal{N}, \mathcal{L})$ um grafo aleatório que segue o modelo de Erdős-Rényi, que denotaremos pro $G(\mathcal{N}, p)$, nesse modelo temos fixados o conjunto de vértices e temos uma probabilidade p para que cada ligação esteja ligada. Podemos construir a rede da seguinte forma:

1. Crie N vértices;
2. Selecione um vértice i e outro vértice j e tente os conectar com probabilidade p , caso consiga a ligação é criada, caso contrário salvamos que ela não existe;
3. Fazemos 2. até testarmos todas as ligações.

Dada a descrição que acabamos de dar do modelo podemos estudar como suas métricas se comportam. O primeiro parâmetro a ser estudado é o grau. Dado um sítio qualquer da rede que no final terá grau k , portanto existiram k ligações com probabilidade p e $N - 1 - k$ (não conta a si mesmo) ligações com $1 - p$. Portanto, a distribuição de grau será:

$$P(k) = \binom{N-1}{k} p^k (1-p)^{N-1-k} \quad (2.17)$$

A partir dela conseguimos calcular facilmente o grau médio da rede, pois é uma distribuição binomial:

$$\langle k \rangle = \sum_{k=1}^{N-1} P(k)k = \sum_{k=1}^{N-1} k p^k (1-p)^{N-1-k} = (N-1)p \quad (2.18)$$

No caso N muito grande ($N \gg k$) é obtido o resultado que a distribuição tende a uma dde Poisson. Isso deve ser analisado pois em redes reais N tem essa característica, por isso esse modelo também é chamado de Rede Binomial ou Rede de Poisson. Considere o resultado anterior de que $\langle k \rangle = (N-1)p \approx Np$ e que $\binom{N-1}{k} = \frac{\prod_{i=1}^k (N-i)}{k!} \approx \frac{N^k}{k!}$, Portanto no limite $N \rightarrow \infty$:

$$P(k) = \lim_{N \rightarrow \infty} \frac{N^k}{k!} p^k (1-p)^{N-1-k} = \frac{\langle k \rangle^k}{k!} \lim_{N \rightarrow \infty} \left(1 - \frac{\langle k \rangle}{N}\right)^{N-k} = \frac{\langle k \rangle^k}{k!} e^{-\langle k \rangle}. \quad (2.19)$$

Outra métrica a ser avaliada é o agrupamento C da rede. Podemos ver pela definição que demos na Sessão anterior que o agrupamento é a probabilidade de dois nós estarem ligados dado que estão ligados ao mesmo nó, isso é basicamente a probabilidade p :

$$C = p = \frac{\langle k \rangle}{N-1} \quad (2.20)$$

2.3.2 Modelo de Configuração

Diferente do Modelo de Erdős-Rényi podemos encontrar o caso na qual nós sabemos que o nosso grafo tem uma quantidade N de sítios e sabemos qual a sequência de graus $\{k_i\}$ cada sítio i da rede. Para gerar redes assim usamos o chamado Modelo de Configuração (MC). Esse modelo tem uma restrição de que o número de ligações (que pode ser obtido facilmente por 2.4) ser um número par. O modelo segue da seguinte forma:

1. Criam-se os N vértices da rede;
2. Cada vértice i recebe κ_i ligações a serem criadas;
3. Escolhemos o vértice i em ordem decrescente de grau da sequência $\{\kappa_i\}$ (isso serve para facilitar a convergência do modelo);

4. Escolhemos aleatoriamente outro j sítio da rede que não tenha ligação com i e que ainda não tenha criada todas as κ_j ligações e criamos a ligação e reduzimos em 1 o valor de κ_i e κ_j ;
5. Repetimos 3 e 4 até colocarmos todas as ligações na rede.

Algorithm 1 Modelo de Configuração

Require: k

▷ Essa lista está em ordem decrescente de grau.

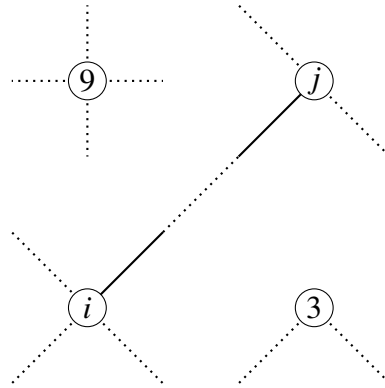
Require: $edges$

```

while ( $i \in \mathcal{N}$ ) and ( $k_i \in k$ ) do      ▷ Os sítios são escolhidos em ordem decrescente de grau.
   $\kappa_i \leftarrow k_i$ 
  while  $j \in \mathcal{N}$  do
    if ( $i \neq j$ ) and ( $k_i \neq 0$ ) then
      if ( $i,j$  not in  $edges$ ) and ( $k_j \neq 0$ ) then
         $edges.insert((i,j))$ 
         $\kappa_i \leftarrow \kappa_i - 1$ 
         $\kappa_j \leftarrow \kappa_j - 1$ 
      end if
    end if
  end while
end while

```

Figura 4 – Ilustração do Modelo

Figura 5 – Funcionamento do Modelo de Configuração, escolhemos dois sítios i e j aleatoriamente e os conectamos. Fonte: Elaborado pelo autor

2.4 Modelos de Infecção ou Difusão

Propagação de doenças, *fake-news*, vírus de internet, rumores são alguns dos temas muito pesquisados no estudo de redes. Entender qual o principal motivo de algo se espalhar em nossa sociedade seja por quais as pessoas que difundem ou pela forma que a rede é construída é essencial para entender a difusão. Além disso, usar de modelagem matemática em é essencial para boa parte desses problemas, pois apesar de termos dados sobre propagação de rumores [14] muitas vezes não é possível coletar dados a partir da experimentação, ou seja não podemos fazer um experimento devido a problemas logísticos ou até morais principalmente quando falamos de saúde pública.

Em vista disso, nessa sessão iremos discutir sobre os modelos de infecção clássicos da literatura. Eles podem aparecer tanto como um modelo em forma de Equação Diferencial ou em Redes. Naquele estamos estudando o nosso modelo de forma macroscópica e considerando que qualquer indivíduo consegue passar a infecção para qualquer outra pessoa enquanto que neste estamos estudando de forma microscópica entendendo como a dinâmica das conexões governam sobre a dinâmica.

2.4.1 Modelo SI

Considere uma população N e que seja $S(0)$ pessoas sejam suscetíveis a uma certa doença e introduzimos $I(0)$ indivíduos infectados e uma probabilidade β de que a doença se espalhe, o que irá acontecer com essa população em um dado tempo t ? Podemos modelar que a variação no tempo da população suscetível $\frac{dS}{dt}$ seja proporcional à probabilidade de encontramos uma pessoa nesse estágio $\frac{S(t)}{N}$ multiplicado pela probabilidade β de ser infectado e considerando que todas pessoas suscetíveis estejam conectadas com todos os infectados temos que multiplicar pelo número de infectados $I(t)$, além disso essa variação tem que ser negativa. Da mesma forma podemos concluir que a variação de infectados em relação ao tempo $\frac{dI}{dt}$ é proporcional a probabilidade de encontrarmos uma pessoa infectada $\frac{I(t)}{N}$ multiplicado pela probabilidade de infecção β e pelo número de pessoas suscetíveis $S(t)$. Gerando assim as seguintes equações diferenciais acopladas:

$$\frac{dS}{dt} = -\beta \frac{S(t)I(t)}{N} \quad (2.21)$$

$$\frac{dI}{dt} = \beta \frac{S(t)I(t)}{N} \quad (2.22)$$

Esse é o modelo mais simples no estudo de infecções na qual um indivíduo só pode estar em dois estágios: suscetível ou infectado. Ao somarmos as duas equações temos:

$$\frac{dS}{dt} + \frac{dI}{dt} = 0 \rightarrow S + I = N = cte \quad (2.23)$$

Essa equação é presente em qualquer outro modelo infecção, é a equação de conservação, isto é todos os indivíduos têm que estar em algum estágio. Além disso substituindo (2.23) em (2.22) temos:

$$\frac{dI}{dt} = \beta \frac{(N - I(t))I(t)}{N} \quad (2.24)$$

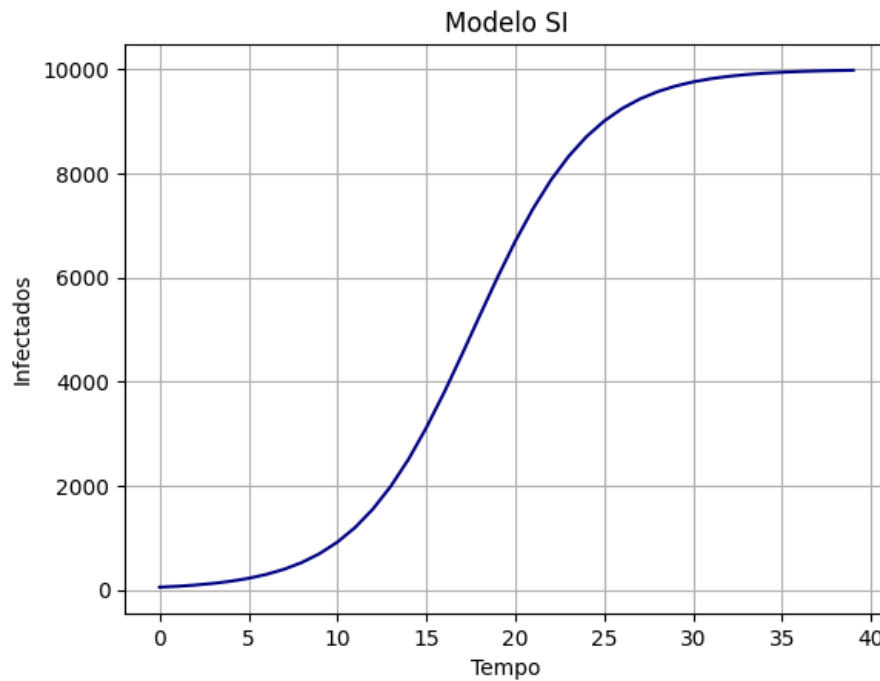
$$\frac{1}{I(t)(N - I(t))} \frac{dI}{dt} = \beta \frac{1}{N} \quad (2.25)$$

Essa equação é facilmente resolvida obtendo:

$$\ln\left(\frac{\frac{N}{I} - 1}{\frac{N}{I(0)} - 1}\right) = -t\beta \rightarrow I(t) = \frac{NI(0)e^{\beta t}}{N - I(0) + I(0)e^{\beta t}} \quad (2.26)$$

Essa é a chamada equação logística que descreve o crescimento do número de infectados em função do tempo nesse modelo. Quando $t \rightarrow 0$ encontramos o valor inicial de infectados e quando $t \rightarrow \infty$ $I(t)$ converge para N pois com o passar do tempo todos os sítios se tornarão infectados o que podemos perceber na Figura 6.

Figura 6 – Modelo SI



Aqui o gráfico de um modelo SI com $N = 1000$ $I(0) = 50$ e $\beta = 0.3$

Fonte: Autor.

Esse simples modelo nos dá uma pequena noção do comportamento exponencial do crescimento da infecção, contudo isso não é real, pois as pessoas conseguem se recuperar da doença, o que nos leva ao próximo modelo.

2.4.2 Modelo SIS

2.4.3 Modelo SIR

2.4.4 Modelos em Redes

3 METODOLOGIA

Por conseguinte é natural a utilização de Redes Complexas como modelagem para estudos da Covid-19. Para tanto precisamos agora de dados sobre redes de contágio que contenham informações relevantes para esse estudo, como idade, tempo de contato, se faz parte de grupo de risco ou não, dentre outros.

Entretanto, não é fácil encontrar um banco de dados tão completo e não enviesado pela recente pandemia. Em 2008, visando o estudo da pandemia de influenza, a Comissão Europeia criou o projeto chamado POLYMOD [15] que tinha o intuito da coleta de dados para entender quais seriam os padrões de contatos da Europa [16]. Apoiado nesse estudo, pesquisas posteriores se inspiraram no molde do POLYMOD [17–27].

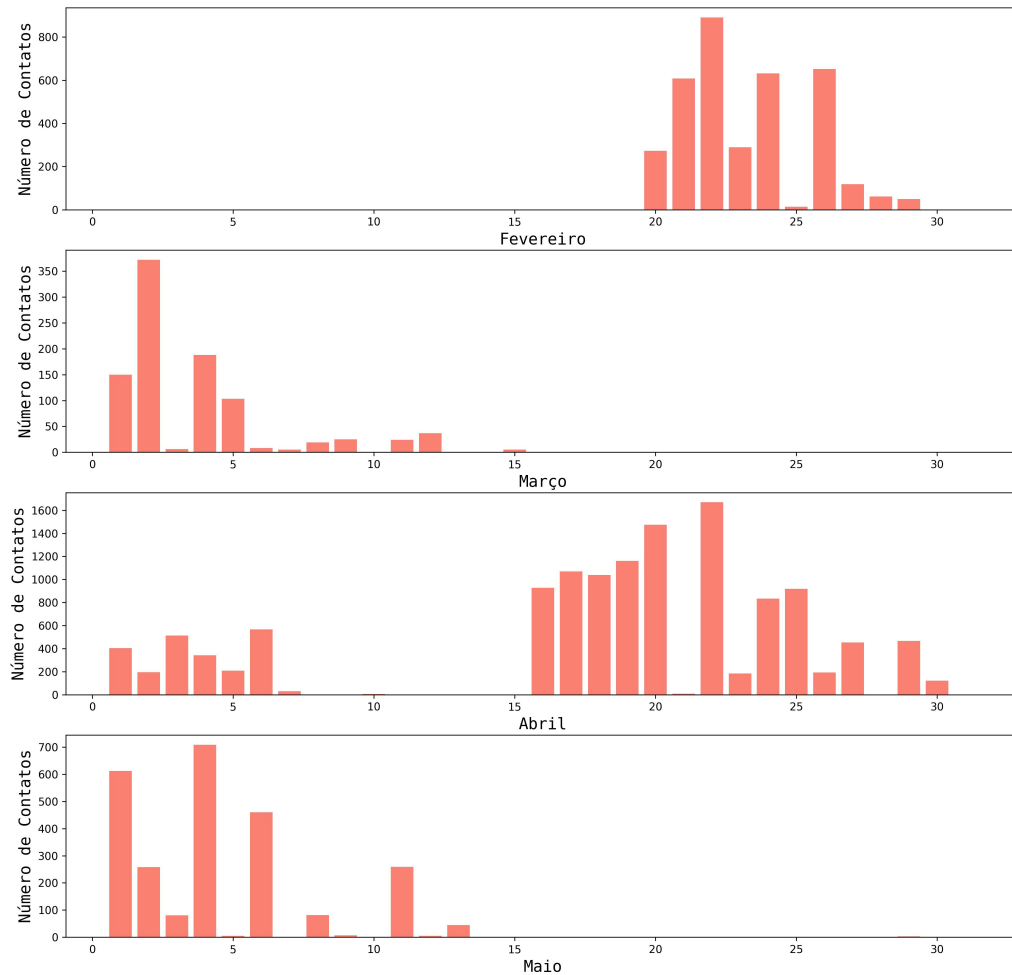
3.1 Análise dos Dados

O foco dessa pesquisa será no banco de dados da França [20] que aconteceu durante o primeiro semestre de 2012. Ele foi coletado a partir de ligações aleatórias da população francesa excluindo territórios ultramarinos. Cada indivíduo que atendesse o telefone receberia um diário de contato para ser preenchido em 2 dias. Apenas uma pessoa por ligação poderia participar, com exceção de crianças e adolescentes para ter uma maior precisão e nesse caso um adulto ficaria responsável por fazer ou auxiliar. Ao final do dia a pessoa teria que preencher o diário anotando sobre informações pessoais, ambiente (escola, trabalho, casa,...) e as pessoas na qual teve contato de curto alcance no dia, a idade aproximada (ou exata) dessas pessoas, duração da conversa, frequência na qual têm contato, dentre outras informações, além disso para facilitar a memorização notificação os participantes foram instruídos a anotarem no máximo 40 contatos. Caso o participante fosse empregado, ele teria que informar qual a média de contatos no seu trabalho, caso fosse maior que 20 então os contatos que seriam anotados no diário seriam os não profissionais.

Toda a pesquisa era para acontecer em dois períodos: Fevereiro a Março e Abril a Maio de 2012. Contudo, no primeiro período foram recrutados menos pessoas do que deveriam e para corrigir isso foi criado um novo período em Abril para completar o primeiro. No total 24.250 pessoas receberam as ligações, 3.977 diários foram enviados, 2033 foram devolvidos e 2029 diários apresentaram contatos próximos. A Figura 7 mostra o número de contatos por dia em cada mês dos períodos da pesquisa.

Figura 7 – Número de Contatos por mês

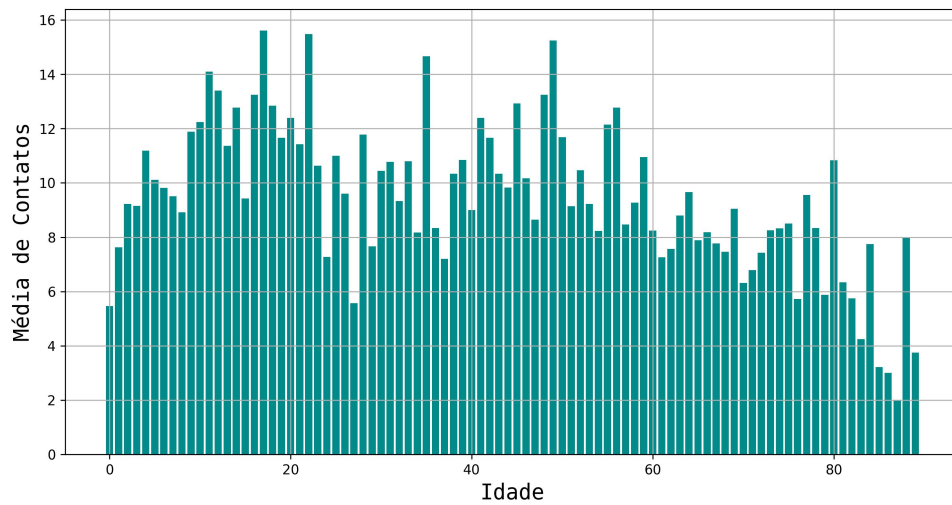
Número de Contatos por Mês



Fonte: Autor.

A partir desses dados vamos coletar alguns resultados deles. Na Figura 8 percebemos que a idade é um fator importante nos seus contatos. Nos primeiros anos de vida do ser humano seus contatos estão limitados a ficar dentro do seu convívio familiar, com o passar dos anos ele consegue ter maior liberdade para sair de casa e ter mais contatos com indivíduos, mas ao chegar no final da vida, depende-se cada vez mais do contato familiar pela limitação da idade.

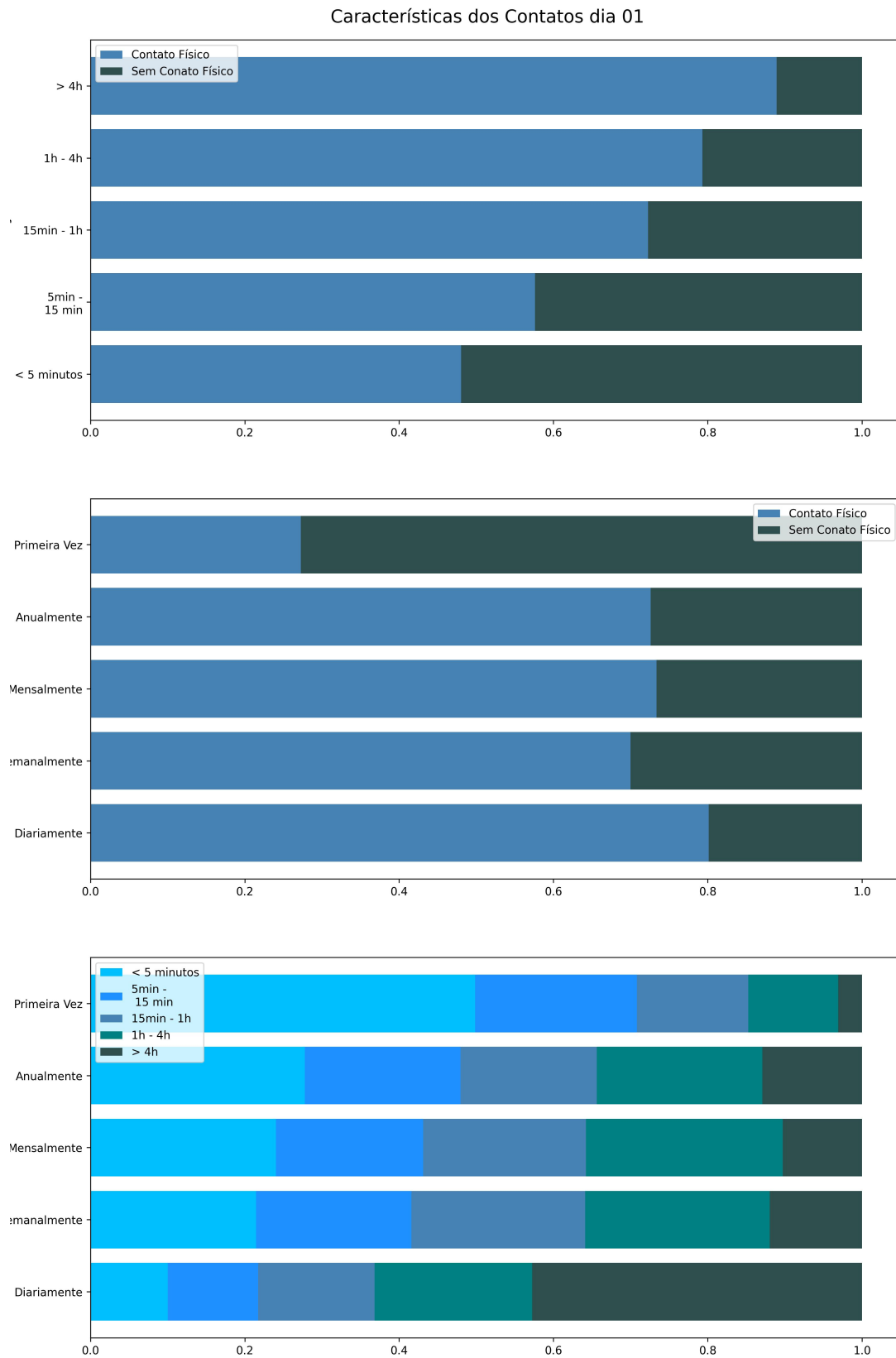
Figura 8 – Gráfico do Número de Contatos por Idade



Fonte: Autor.

Na Figura 9 observamos como os contatos são caracterizados. Quanto mais tempo nós passamos conversando com a pessoa menores as chances de acontecer algum contato físico, apesar de ser muito comum. A frequência com a qual você vê a pessoa também tem importante na probabilidade de acontecer ou não contato físico, contudo fora a primeira contato e o contato diário, praticamente a chance de haver contato físico é a mesma. A última imagem vai ao encontro dessa afirmação, percebe-se que as convivências que acontecem pela primeira vez tendem a durar em torno de minutos, enquanto que os convivências diárias podem facilmente durar horas, já em outras frequências de contatos, não parece haver uma diferenciação clara. À vista disso, há uma tendência a uma supervalorização (em questão de tempo) no contato diário pelas pessoas, há uma desvalorização na primeira vez e não há uma diferença em qualquer outra frequência.

Figura 9 – Gráfico do Número de Contatos por Idade

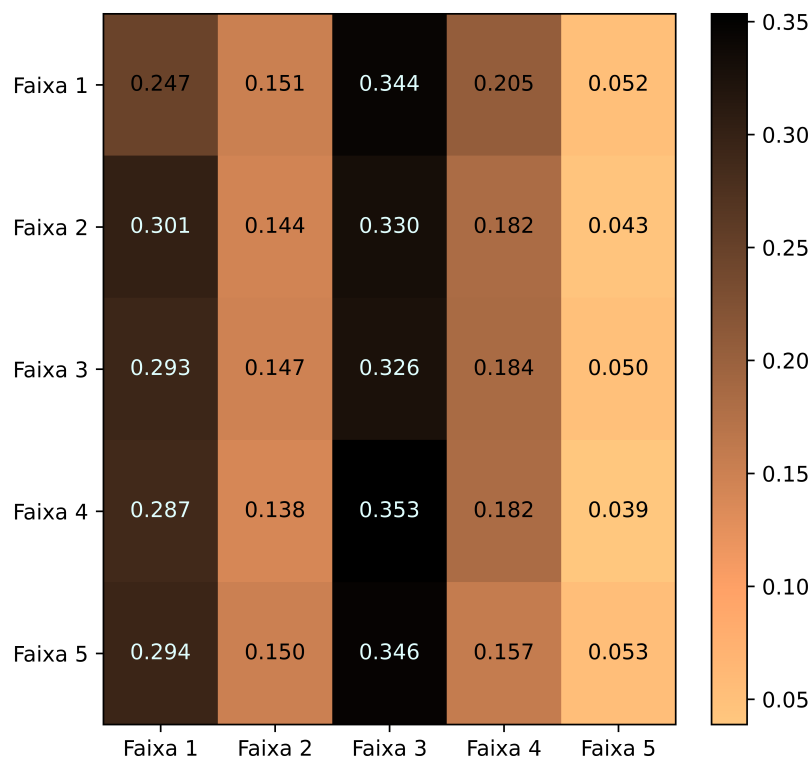


Fonte: Autor.

Outro elemento importante a ser investigado é como são as conexões entre as idades do estudo. Para isso cada indivíduo será categorizado em 5 faixas etárias: [0,20) (jovens), [20,30) (jovens adultos), [30,50) (adultos), [50,70) (adultos seniors) e maiores de 70 que representam os idosos. Na Figura nota-se como é a distribuição de idades da pesquisa. Ao analisar a distribuição de ligações por faixa etária é percebido que seguem uma distribuição geométrica na qual pode ser escrita como $P \propto (1-p)^x p$ ou $P \propto e^{-\lambda x} (1 - e^{-\lambda})$. Na Figura 11 temos o estudo das distribuições como se fosse da segunda forma, no eixo das abscissas temos o grau e no eixo das coordenadas temos o logaritmo da probabilidade, com isso é encontrado uma reta com Mínimos Quadrados na qual o coeficiente angular é igual ao λ e seu respectivo R^2 .

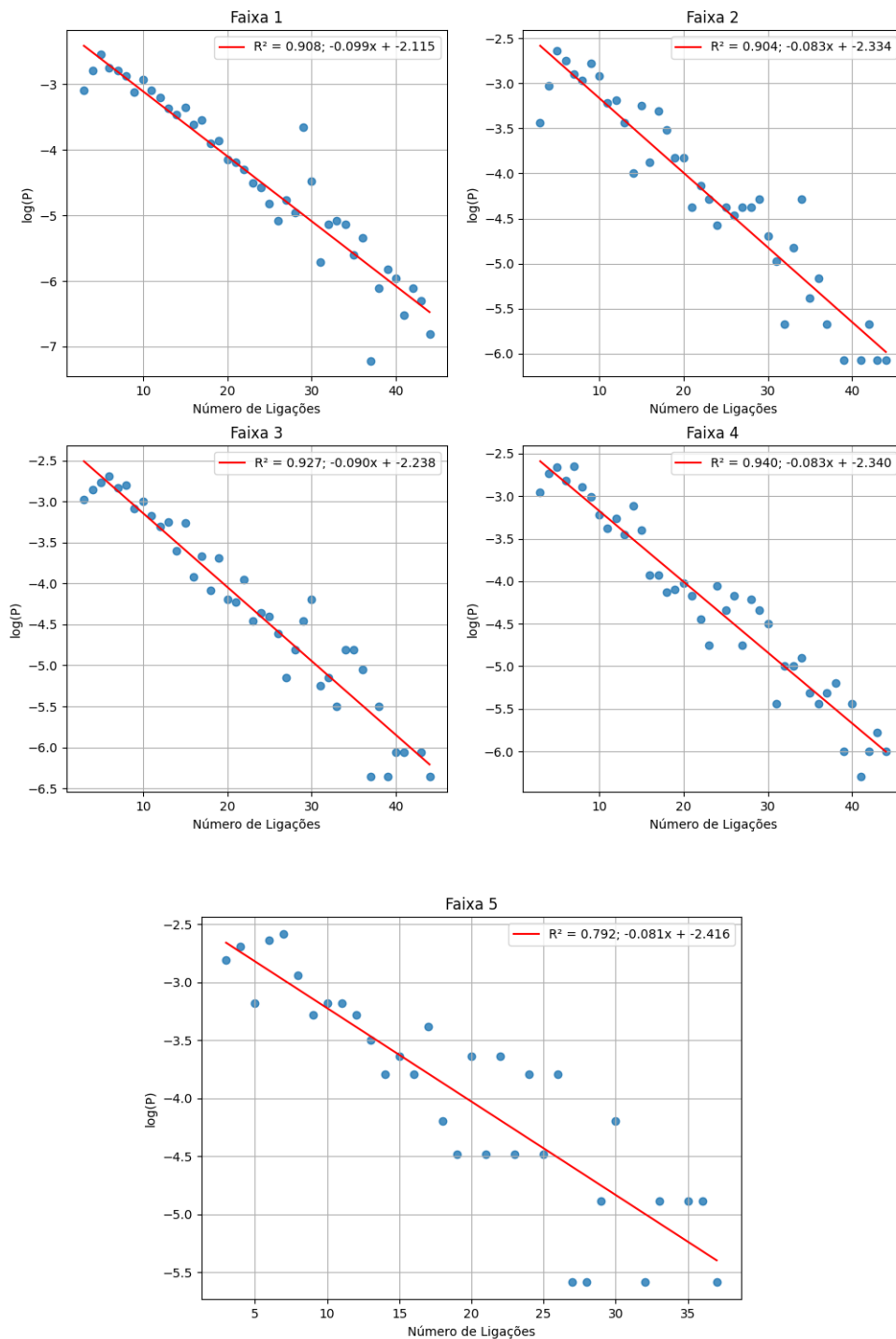
Além de examinarmos as ligações dentro de cada faixa etária, também é relevante analisarmos as interações entre as diferentes faixas etárias. Na Figura 10, podemos observar a fração média de ligações entre essas faixas etárias. Notavelmente, encontramos um número significativo de ligações entre pessoas com idades entre 30 e 50 anos, visto que essa faixa etária é bastante presente tanto no ambiente de trabalho como no âmbito familiar. Além disso, junto à faixa etária 1, essas idades representam as maiores interações. Por outro lado, a faixa etária mais avançada apresenta menos ligações, pois poucas pessoas chegam a essa fase da vida e, dentre elas, muitas mantêm contato com indivíduos mais jovens para cuidados médicos.

Figura 10 – Fração média de ligações entre faixas etárias



Fonte: Autor.

Figura 11 – Proporção de cada faixa etária no Banco de dados.



Fonte: Autor.

Apesar da grande quantidade de pessoas e contatos há um problema essencial: não há garantia de que os entrevistados se conectem em si. Por exemplo, no caso da França [20] fo-

ram feitas ligações aleatórias extraída da população francesa excluindo territórios ultramarinos, assim nossa análise em rede estaria bastante limitada.

3.2 Modelo de Redes

Para evitar esse problema [28] propõe um modelo para formamos uma rede baseado nos dados franceses. A partir dos dois dias de entrevista é calculado a média de contatos por dia de cada pessoa (arredondado para cima) e utiliza o Modelo de Configuração para a formação de uma Rede Sintética. No entanto, o modelo de formação de redes gera um agrupamento médio limitado e ele é muito importante na suavização de epidemias [29]. Nesse sentido Manzo propõe que antes de excluirmos um dado sítio do Modelo de Configuração, passamos por cada vizinho dele e os conectamos entre si com uma probabilidade p . Ou seja:

1. Quando um dado sítio i atinge o grau requerido pelo MC selecionamos todos os vizinhos $v(i)$;
2. Selecionamos dois sítios $v, n \in v(i)$ e conectamos com probabilidade p ;
3. Se ela existir os conectamos e salvamos para que ela não se repita no MC, caso contrário salvamos para que ela não se repita nesse algoritmo. Por fim também reduzimos em 1 o valor de κ_v e κ_n ;
4. Fazemos 2. até testarmos todas as ligações possíveis e paramos esse algoritmo para dar continuidade ao MC.

Figura 12 – Ilustração do Modelo de Manzo

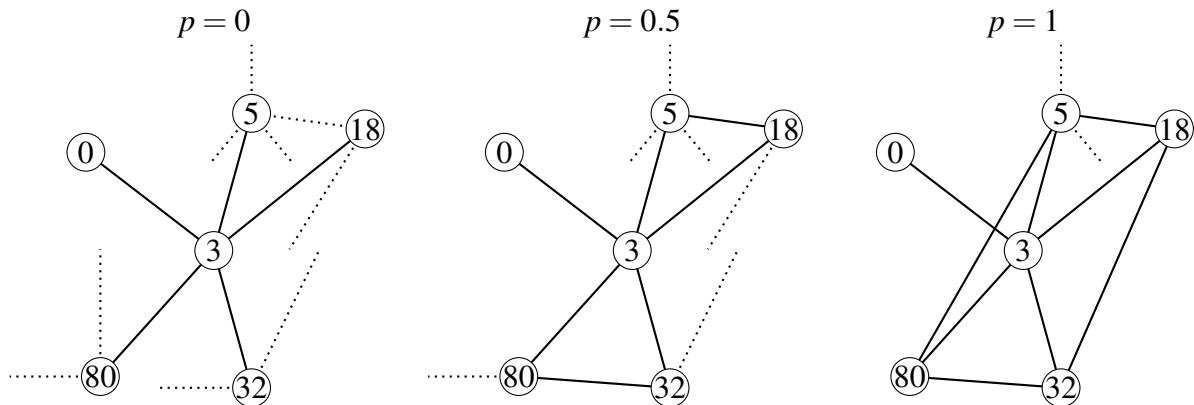


Figura 13 – Funcionamento do Modelo de Configuração, escolhemos dois sítios i e j aleatoriamente e os conectamos. O Algoritmo pode gerar várias topologias de redes, porém ainda limitadas pelas quantidades $\{k_i\}$ de graus impostas pelo MC.

Fonte: Elaborado pelo autor

O Modelo é mostrado na Figura 13, na qual mostramos o estágio quando encontramos um dado sítio que atinge o grau especificado. Assim como nos modelos de formação de rede existem várias redes que podem ser formadas dada uma probabilidade p . Outrossim, se o

Algorithm 2 Implementação de Manzo

Require: $edges$ **Require:** κ **Require:** $n_existir$

▷ É necessário salvar quais ligações não vão existir.

Require: p **Require:** $i \in G(\mathcal{N}, \mathcal{L})$

```

while  $v \in v(i)$  do
  while  $n \in v(i)$  do                                ▷ Nesse caso não há preferência na escolha de  $n$  e  $v$ .
    if  $(v \neq n)$  and  $(\kappa_n \neq 0)$  and  $(\kappa_v \neq 0)$  then
       $r \leftarrow \text{random\_number}([0,1])$ 
      if  $(r \leq p)$  then
        if  $(v, n \text{ not in } edges) \text{ and } (v, n \text{ not in } n\_existir)$  then
           $edges.insert((v, n))$                                 ▷ Essa matriz vem do MC
           $\kappa_v \leftarrow \kappa_v - 1$ 
           $\kappa_n \leftarrow \kappa_n - 1$ 
        end if
      else
        if  $n, v \text{ not in } n\_existir$  then
           $n\_existir.insert((v, n))$ 
        end if
      end if
    end while
  end while

```

sítio já tiver atingido o grau requerido pelo MC, então ele não precisa entrar nesse algoritmo. A partir desse modelo o autor consegue os seguintes resultados mostrados na Tabela 2 e os meus estão na Tabela 4.

Tabela 1 – Tabela do resultado de Manzo

| Probabi- -lidade | Grau Médio | Grau Mediano | Desvio Padrão Grau | Agrup- amento Médio | Corr Agrup- Grau | Menor Caminho Médio | Diâmetro |
|---------------------|---------------|-----------------|--------------------------|---------------------------|------------------------|---------------------------|--------------|
| $p = 0$ | 9.72 (0.00) | 8 (0.00) | 6.56 (0.00) | 0.01 (0.00) | -0.06 (0.01) | 3.47 (0.00) | 6 (0.00) |
| $p = 0.5$ | 9.72 (0.00) | 8 (0.00) | 6.56 (0.00) | 0.43 (0.00) | -0.62 (0.01) | 4.38 (0.03) | 7.45 (0.5) |
| $p = 1.0$ | 9.72 (0.00) | 8 (0.00) | 6.56 (0.00) | 0.57 (0.01) | -0.56 (0.01) | 5.52 (0.09) | 10.10 (0.59) |

Tabela 2 – Nessa tabela são apresentados resultados advindos do autor para a rede francesa feitas em 100 simulações de redes. Apresentando as características da rede e em parêntese o erro associado a cada valor.

Fonte: Elaborada pelo autor.

Os valores com probabilidade $p = 0$ batem perfeitamente com os resultados do autor, os outros valores se aproximam bastante do que aparece no artigo, porém ainda na mesma faixa de erro. Podemos ver que o Grau permanece o mesmo para que seja respeitado o MC já o agrupamento aumenta bastante com o incremento de p chegando a seu valor máximo em 0.56 que é limitado pelo MC.

Tabela 3 – Tabela do meu resultado.

| Probabi- -lidade | Grau Médio | Grau Mediano | Desvio Padrão Grau | Agrup- amento Médio | Corr Agrup- Grau | Menor Caminho Médio | Diâmetro |
|---------------------|---------------|-----------------|--------------------------|---------------------------|------------------------|---------------------------|-------------|
| p = 0 | 9.72(0.00) | 8.00(0.00) | 6.56(0.00) | 0.01(0.00) | -0.06(0.01) | 3.47(0.00) | 6.00(0.03) |
| p = 0.5 | 9.72(0.00) | 8.00(0.00) | 6.56(0.00) | 0.43(0.01) | -0.62(0.01) | 4.37(0.05) | 7.46(0.50) |
| p = 1.0 | 9.72(0.00) | 8.00(0.00) | 6.56(0.00) | 0.56(0.01) | -0.57(0.01) | 5.62(0.11) | 10.39(0.66) |

Tabela 4 – Nessa tabela são apresentados resultados advindos do meu código na qual foram feitas 1000 simulações de redes para a rede francesa que se aproximam bastante do apresentado no artigo junto com os desvios padrões entre parênteses. Os que não bateram certo estão na faixa de erro mostrada anteriormente.

Fonte: Elaborada pelo autor.

Assim a partir dos resultados mostrados na Figura 11, podemos construir um grafo $G(\mathcal{N}, \mathcal{L})$ na qual cada sítio i recebe k_i ligações que seguem a distribuição empírica dos dados e assim cada sítio terá $k_{i,j}$ ligações com a faixa etária j que seguirá uma distribuição multinomial a partir das probabilidades encontradas na Figura 10.

Nesse sentido jogaremos esse conjunto $k_{i,j}$ de ligações no modelo de configuração e será feita uma nova alteração considerando ligações por faixa etária. Contudo, na prática isso não funcionou satisfatoriamente, pois como não se há informação completa de como é a rede real tanto porque os entrevistados não se conectam ente si o algoritmo não converge, faltando de 9% a 20% das ligações.

Para solucionar esse problema é proposto uma última modificação, ao final do modelo de configuração será utilizado o seguinte algoritmo:

1. Será criado um nó sintético i' com faixa etária aleatória e com um conjunto $k_{i',j}$ de ligações com outras faixas etárias que segue a mesma distribuição da rede.
2. Esse novo sítio i' tentará ser conectado aleatoriamente com os nós que sobraram ligações até que todas as $k_{i',j}$ sejam colocadas ou seja tentado todos os sítios (em geral acontece a segunda opção).
3. Então se foi possível colocar ligações com o nó i' coloca-se ele na rede.
4. Repete 1-3 até o momento em que todas as ligações forem colocadas.

3.3 Modelo de Infecção

Como discutido anteriormente a pandemia do COVID-19 teve bastante impacto na sociedade e um deles foi no aumento nos estudos de modelos epidemiológicos, criando modelos mais robustos [30] considerando mais do que apenas infectados e suscetíveis. Além disso esses modelos estão em formato de equações diferenciais e isso tem implicações na modelagem na qual é considerado que todos os indivíduos conseguem propagar a doença para qualquer outro indivíduo. Além disso, nesse formato de modelagem estamos estudando o nosso sistema de forma macroscópica e quando se tenta modelar utilizando redes estamos tentando estudar de forma microscópica, portanto a escolha de qual modelo é essencial e como ele se comporta de forma microscópica é essencial. Então qual modelo iremos escolher?

Para responder essa pergunta temos que entender qual o propósito de vacinas. As vacinas têm como principal função reduzir a propagação da enfermidade e impedir que os indivíduos que por ventura se contaminem não apresentem os efeitos mais graves da doença. Por fim não só é considerado o fator humano mas também econômico como é o caso da superlotação de hospitais [31]. Nesse sentido não é apenas o espaço físico da UTI ou Hospital que está sendo utilizado, mas também de remédios, equipes de saúde e suprimentos, que poderiam estar sendo alocados em outras localidades e também prejudica pessoas que precisariam do espaço para outros problemas. À vista disso, usaremos como métricas para serem minimizadas são o número básico de reprodução R_0 , número de mortos M e número de Hospitalizações M_H .

Assim sendo, o modelo de contágio a ser utilizado é o chamado SEIHARD como base para nossa primeira aproximação. Nele temos os sítios que podem se apresentar como Suscetíveis, Expostos, Infectados Sintomáticos, Hospitalizados, Infectados Assintomáticos, Recuperados e Mortos. Esse modelo é uma adaptação de [32] e [33], podemos definir cada estágio da seguinte forma:

- **Suscetível (S):** é o sítio que tem conexões na rede e não foi infectado pelo vírus ainda;
- **Exposto (E):** um sítio é considerado exposto quando ele estava no estágio S e foi contaminado por alguém no estágio E ou A. Esse estágio é para representar os indivíduos que estão na fase de incubação do vírus, isso faz com que eles possam contaminar os outros indivíduos por não saberem que tem a doença;
- **Infectado Sintomático (I):** um sítio nesse estágio está infectado, porém ele não pode infectar outras pessoas. Fazendo isso estamos tentando emular uma pessoa que está sentindo os sintomas e por isso não teria mais contato com outras pessoas;
- **Hospitalizado (H):** nele o nó continua sem a capacidade de contaminar outros nós. Nessa primeira aproximação estamos considerando que podemos ter qualquer quantidade de sítios hospitalizados, o que não é verdade por isso a posteriori iremos atualizar essa parte para melhorar o modelo;

- **Infectado Assintomático (A):** a diferença desse estágio para o I é o fato de que ele pode infectar outras pessoas. Isso se deve ao fato de que se ele é assintomático então não tem conhecimento da doença por não sentir nenhum sintoma, portanto ele continua tendo as suas ligações e também não tem risco de morte ou necessidade de ser hospitalizado;
- **Recuperado (R):** caso o nó não morra ele será considerado Recuperado.
- **Morto (D):** Por fim, nesse o sítio perde todas as conexões e é nele que será contabilizado quantas pessoas morreram para utilizarmos como métrica de minimização.

Figura 14 – Ilustração do Modelo SEIHARD

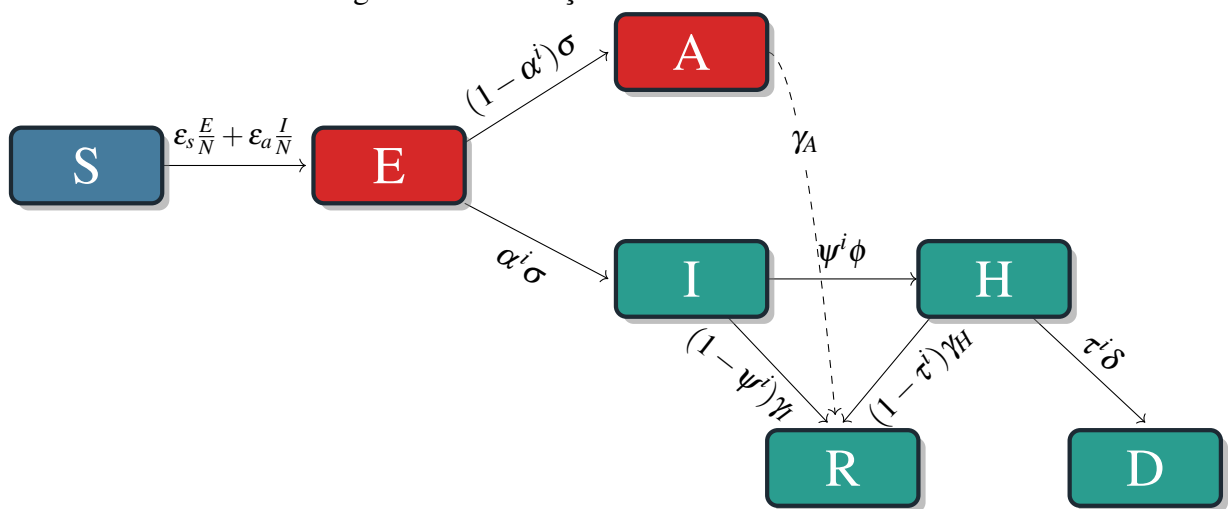


Figura 15 – Esquematização do Modelo SEIHARD de contágio. Os blocos em vermelho significam quem pode contaminar outros, os blocos em azul escuro quem não é infectado e em azul claro quem está isolado da rede.

Tabela 5: Tabela de parâmetros do modelo.

| Parâmetro | Definição | Valor [Referência] |
|--------------|---|--------------------------|
| ϵ_s | Proporção de Infecção de Sintomáticos | 0.5 day^{-1} |
| ϵ_a | Proporção de Infecção de Assintomáticos | 0.41 day^{-1} |
| α^i | Fração de Infectados que tornarão-se Sintomáticos | Tabela 6 [34] |
| σ | Transição exposta a infecções | $1/5.1 \text{ day}^{-1}$ |
| γ_A | Taxa de Recuperação de Assintomáticos | $1/7 \text{ day}^{-1}$ |
| γ_H | Taxa de Recuperação de Hospitalizados | $1/14 \text{ day}^{-1}$ |
| γ_I | Taxa de Recuperação de Sintomáticos | $1/7 \text{ day}^{-1}$ |
| ϕ | Taxa de Hospitalização | 0.025 day^{-1} |
| ψ^i | Probabilidade de Hospitalização | Tabela 7 |
| δ | Taxa de Mortalidade | 0.015 day^{-1} |
| τ^i | Probabilidade de Morte | Tabela 8 |

Fonte: Elaborada pelo autor.

| Faixa | 0 - 20 | 20 - 30 | 30 - 50 | 50 - 70 | ≥ 70 |
|-------|--------|---------|---------|---------|-----------|
| Valor | 29,1% | 37,4% | 41.68% | 39.4% | 31.3% |

Tabela 6: Nessa tabela são apresentados os valores de α^i , que é a probabilidade de se tornar sintomático pela idade.

Fonte: Elaborada pelo autor.

| Faixa | 0 - 20 | 20 - 30 | 30 - 50 | 50 - 70 | ≥ 70 |
|-------|--------|---------|---------|---------|-----------|
| Valor | 0.408% | 1.04% | 3.89% | 9.98 % | 17.5% |

Tabela 7: Nessa tabela são apresentados os valores de ψ^i , que é a probabilidade de um infectado ser hospitalizado.

Fonte: Elaborada pelo autor.

| Faixa | 0 - 20 | 20 - 30 | 30 - 50 | 50 - 70 | ≥ 70 |
|-------|--------|---------|---------|---------|-----------|
| Valor | 1.04% | 1.33% | 1.38% | 7.6% | 24% |

Tabela 8: Nessa tabela são apresentados os valores de τ^i , que é a probabilidade de um hospitalizado morrer. Esses dados foram coletados a partir do Open Data SUS do ano de 2020.

Fonte: Elaborada pelo autor.

4 CONCLUSÃO

REFERÊNCIAS

- [1] HUREMOVIĆ, D. Psychiatry of pandemics (a mental health response to infection outbreak) — brief history of pandemics (pandemics throughout history). Springer International Publishing, v. 10.1007/978-3-030-15346-5, may 2019.
- [2] DOBSON, A. P.; CARPER, E. R. Infectious diseases and human population history. *BioScience*, Oxford University Press (OUP), v. 46, n. 2, p. 115–126, fev. 1996. Disponível em: <<https://doi.org/10.2307/1312814>>.
- [3] GRENNAN, D. What is a pandemic? *JAMA*, American Medical Association (AMA), v. 321, n. 9, p. 910, mar. 2019. Disponível em: <<https://doi.org/10.1001/jama.2019.0700>>.
- [4] EULER, L. Solutio problematis ad geometriam situs pertinentis. *Commentarii Academiae Scientiarum Imperialis Petropolitanae*, v. 8, p. 128–140, 1736.
- [5] BOGUSLAWSKI, P. Modelling and analysing 3d building interiors with the dual half-edge data structure. 01 2011.
- [6] SZWARCFITER, J. L. *TEORIA COMPUTACIONAL DE GRAFOS*. 1. ed. [S.l.]: Elsevier, 2018. ISBN 9788535288841; 8535288848; 9788535288858; 8535288856.
- [7] GOLDBARG, M. *Grafos: Conceitos, algoritmos e aplicações*. [S.l.: s.n.], 2012. ISBN 978-85-352-5716-8.
- [8] BARABÁSI, A.-L. *Network Science*. [S.l.]: Cambridge University Press, 2016. ISBN 9781107076266; 1107076269.
- [9] NEWMAN, M. *Networks*. 2. ed. [S.l.]: Oxford University Press, 2018. ISBN 0198805098; 9780198805090.
- [10] COSTA, L. da F. et al. Characterization of complex networks: A survey of measurements. *Advances in Physics*, Informa UK Limited, v. 56, n. 1, p. 167–242, jan. 2007. Disponível em: <<https://doi.org/10.1080/00018730601170527>>.
- [11] KLEINBERG, J. The small-world phenomenon: An algorithmic perspective. *Conference Proceedings of the Annual ACM Symposium on Theory of Computing*, 04 2001.
- [12] FRANCO, B. M. . G. M. . S. Inside pagerank. *ACM Transactions on Internet Technology* 2005-feb 01 vol. 5 iss. 1, v. 5, feb 2005.
- [13] ERDÖS, P.; RÉNYI, A. On random graphs i. *Publicationes Mathematicae Debrecen*, v. 6, p. 290–297, 1959.

- [14] RUMORS, False Flags, and Digital Vigilantes: Misinformation on Twitter after the 2013 Boston Marathon Bombing. In: iConference 2014 Proceedings. iSchools, 2014. Disponível em: <<https://doi.org/10.9776/14308>>.
- [15] MOSSONG, J. et al. *POLYMOD social contact data*. Zenodo, 2017. Disponível em: <<https://zenodo.org/record/3746312>>.
- [16] MOSSONG, J. et al. Social contacts and mixing patterns relevant to the spread of infectious diseases. *PLoS Medicine*, Public Library of Science (PLOS), v. 5, n. 3, p. e74, mar. 2008. Disponível em: <<https://doi.org/10.1371/journal.pmed.0050074>>.
- [17] NIEL, H. et al. *Social contact data for Belgium (2006)*. Zenodo, 2020. Disponível em: <<https://zenodo.org/record/4059864>>.
- [18] LANDER, W. et al. *Social contact data for Belgium (2010-2011)*. Zenodo, 2020. Disponível em: <<https://zenodo.org/record/4302055>>.
- [19] ZHANG, J. et al. *Social contact data for China mainland*. Zenodo, 2020. Disponível em: <<https://zenodo.org/record/3878754>>.
- [20] BÉRAUD, G. et al. *Social contact data for France*. Zenodo, 2020. Disponível em: <<https://zenodo.org/record/3886590>>.
- [21] LEUNG, K. et al. *Social contact data for Hong Kong*. Zenodo, 2020. Disponível em: <<https://zenodo.org/record/3874808>>.
- [22] GRIJALVA, C. G. et al. *Peruvian Social Contact Data*. Zenodo, 2017. Disponível em: <<https://zenodo.org/record/1095665>>.
- [23] LITVINOVA, M. et al. *Social contact data for Russia*. Zenodo, 2020. Disponível em: <<https://zenodo.org/record/3874674>>.
- [24] Mahikul Wiriya et al. *Social contact data for Thailand*. Zenodo, 2020. Disponível em: <<https://zenodo.org/record/4086739>>.
- [25] Horby Peter et al. *Social Contact Data For Vietnam*. Zenodo, 2018. Disponível em: <<https://zenodo.org/record/1289474>>.
- [26] DODD, P. J. et al. *Social contact data for Zambia and South Africa (CODA dataset)*. Zenodo, 2019. Disponível em: <<https://zenodo.org/record/2548693>>.
- [27] MELEGARO, A. et al. *Social contact data for Zimbabwe*. Zenodo, 2020. Disponível em: <<https://zenodo.org/record/3886638>>.

- [28] MANZO, G.; RIJT, A. van de. Halting SARS-CoV-2 by targeting high-contact individuals. *Journal of Artificial Societies and Social Simulation*, Journal of Artificial Societies and Social Simulation, v. 23, n. 4, 2020. Disponível em: <<https://doi.org/10.18564/jasss.4435>>.
- [29] BLOCK, P. et al. Social network-based distancing strategies to flatten the COVID-19 curve in a post-lockdown world. *Nature Human Behaviour*, Springer Science and Business Media LLC, v. 4, n. 6, p. 588–596, jun. 2020. Disponível em: <<https://doi.org/10.1038/s41562-020-0898-6>>.
- [30] XIANG, Y. et al. COVID-19 epidemic prediction and the impact of public health interventions: A review of COVID-19 epidemic models. *Infectious Disease Modelling*, Elsevier BV, v. 6, p. 324–342, 2021. Disponível em: <<https://doi.org/10.1016/j.idm.2021.01.001>>.
- [31] FRENCH, G. et al. Impact of hospital strain on excess deaths during the COVID-19 pandemic — united states, july 2020–july 2021. *MMWR. Morbidity and Mortality Weekly Report*, Centers for Disease Control MMWR Office, v. 70, n. 46, p. 1613–1616, nov. 2021. Disponível em: <<https://doi.org/10.15585/mmwr.mm7046a5>>.
- [32] EIKENBERRY, S. E. et al. To mask or not to mask: Modeling the potential for face mask use by the general public to curtail the COVID-19 pandemic. *Infectious Disease Modelling*, Elsevier BV, v. 5, p. 293–308, 2020. Disponível em: <<https://doi.org/10.1016/j.idm.2020.04.001>>.
- [33] MALAGUTTI, R.; EISENCRAFT, M. Seir epidemic model applied to the covid-19 outbreak in são paulo. 11 2021.
- [34] JUNG, C.-Y. et al. Clinical characteristics of asymptomatic patients with COVID-19: A nationwide cohort study in south korea. *International Journal of Infectious Diseases*, Elsevier BV, v. 99, p. 266–268, out. 2020. Disponível em: <<https://doi.org/10.1016/j.ijid.2020.08.001>>.

APÊNDICE A – TESTANDO O APÊNDICE

A.1 Teste

A.1.1 teste

APÊNDICE B – TESTE2

B.1 teste

ANEXO A – TESTANDO O ANEXO