



**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

**ANÁLISE DE ESTRATÉGIAS DE VACINAÇÃO PARA COVID-19 BASEADA EM
REDES COMPLEXAS**

**FORTALEZA
2024**

CARLOS MIGUEL MOREIRA GONÇALVES

ANÁLISE DE ESTRATÉGIAS DE VACINAÇÃO PARA COVID-19 BASEADA EM REDES
COMPLEXAS

Dissertação apresentada ao Programa de Pós-Graduação em Modelagem e Métodos Quantitativos do Centro de Ciências da Universidade Federal do Ceará, como requisito parcial à 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.

FORTALEZA

2024

Dados Internacionais de Catalogação na Publicação
Universidade Federal do Ceará
Sistema de Bibliotecas
Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

Gonçalves, Carlos Miguel Moreira.

Análise de estratégias de vacinação para COVID-19 baseada em redes complexas / Carlos Miguel Moreira Gonçalves. – 2025.

112 f. : il. color.

Dissertação (mestrado) – Universidade Federal do Ceará, Centro de Ciências, Programa de Pós-Graduação em Modelagem e Métodos Quantitativos, Fortaleza, 2025.

Orientação: Prof. Dr. Leandro Chaves Rêgo.

Coorientação: Prof. Dr. Pablo Ignacio Fierens.

1. Redes (Matemática). 2. Epidemiologia. 3. COVID-19 (Doença) - Vacinação. 4. Medidas de centralidade.. I. Título.

CDD 510

CARLOS MIGUEL MOREIRA GONÇALVES

ANÁLISE DE ESTRATÉGIAS DE VACINAÇÃO PARA COVID-19 BASEADA EM REDES
COMPLEXAS

Dissertação apresentada ao Programa de Pós-Graduação em Modelagem e Métodos Quantitativos do Centro de Ciências da Universidade Federal do Ceará, como requisito parcial à obtenção do título de mestre em Modelagem e Métodos Quantitativos. Área de Concentração: Modelagem e Métodos Quantitativos.

Aprovada em: 05/08/2024.

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. Jesus Ossian da Cunha Silva
Universidade Federal do Ceará (UFC)

Prof. Dr. Ricardo Lopes de Andrade
Instituto Federal de Educação, Ciência e Tecnologia
de Pernambuco (IFPE)

AGRADECIMENTOS

À minha noiva Isa Crisna Lima Duarte pelo apoio em todas as áreas da minha vida: seja no pessoal ou no acadêmico ela estava lá para me apoiar, tornando a luta diária um peso menor e trazendo muita felicidade à minha vida. Aos meus pais e meu irmão que convivem comigo e me apoiam constantemente a melhorar e me tornar uma pessoa melhor.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior Brasil (CAPES) Código de Financiamento 001.

Ao Prof. Dr. Leandro Chaves Rêgo e ao Prof. Dr. Pablo Ignacio Fierens, pela excelente orientação, pelas conversas muito esclarecedoras e pelo apoio em toda a jornada acadêmica.

Aos professores participantes da banca examinadora Jesus Ossian da Cunha Silva e Ricardo Lopes de Andrade pelo tempo, pelas valiosas colaborações e sugestões.

Aos professores Ascânio Dias Araújo, Michael Ferreira de Souza e por terem cedido o espaço do Laboratório de Simulação Numérica (LSN) para aprimorar meus estudos e simulações além das ótimas discussões sobre computação, algoritmos de Física.

Aos colegas do Laboratório de Simulação Numérica, que me ensinaram muito e fizeram com que todo o processo do mestrado e doutorado fosse facilitado.

“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.” (Christakis, 2010).

RESUMO

A rápida disseminação, elevada taxa de mortalidade e sobrecarga nos sistemas hospitalares tornaram o vírus SARS-CoV-19 um desafio significativo para a humanidade. A emergência da situação exigia investimentos substanciais no desenvolvimento de vacinas como uma medida crucial para mitigar essa problemática. Entretanto, os custos financeiros e o tempo necessário para o desenvolvimento de vacinas representam obstáculos para alcançar uma cobertura vacinal abrangente. Portanto, compreender a estratégia mais eficaz para a imunização da população é de suma importância para conter a propagação do vírus e otimizar a alocação de recursos em políticas públicas. É relevante destacar que grande parte dessas políticas estava fundamentada no distanciamento social, o qual teve impactos significativos nas esferas econômica, educacional e de saúde mental. O propósito deste projeto consiste em propor uma abordagem baseada em modelagem, utilizando o modelo de contágio SEIHARDS em uma rede complexa de contatos derivada de uma pesquisa conduzida por meio de questionário. O objetivo é identificar estratégias ótimas de vacinação para reduzir a disseminação do vírus, minimizar as hospitalizações e mitigar o número de óbitos. Para isso foram coletados dados de uma pesquisa chamada POLYMOD, que foi feita em 8 países da Europa, que tinha como objetivo mapear as conexões entre pessoas. Essa pesquisa é a mais alinhada com o projeto pois ela detém dados sobre as faixas etárias dos entrevistados, algo bastante relevante na hospitalização e mortalidade do COVID-19. Todavia os entrevistados dessa pesquisa não haviam conexão entre si e foi necessário desenvolver o algoritmo chamado Modelo de Configuração Estratificado (MCE) para formar uma rede com essa limitação do banco de dados que leva em consideração as idades e adicione a duração de tempo de contato entre indivíduos. Somado a isso, o modelo proposto consegue aumentar o agrupamento da rede porque o Modelo de Configuração tradicional consegue agrupamentos baixos. Ademais, a partir de coleta de dados de artigos e do OpenDataSUS foi possível encontrar as constantes para o modelo epidemiológico proposto que alguns parâmetros variam com a faixa etária. Esse projeto analisou o comportamento do algoritmo proposto com a mudança do agrupamento, propôs métricas de centralidades para este estudo e analisou o comportamento e a sensibilidade do modelo de infecção em relação às variações no agrupamento e na construção da rede, mostrando que o modelo é pouco sensível aos incrementos no agrupamento e que com a inserção de ponderação nas arestas não há ganho nas métricas e há um aumento de até 5 vezes no tempo computacional. As medidas de centralidade que consideraram pesos nos nós e a estrutura da rede mostraram-se altamente eficazes para a vacinação contra a COVID-19, com estratégias

como o *PageRank* reduzindo a mortalidade em mais de 60%, o tempo total de hospitalização em 66% e necessitando vacinar de 31% da rede para inibir a proliferação do vírus em comparação à ausência de vacinação. Algumas centralidades que utilizam pesos em arestas apresentaram duas abordagens de cálculo: altruísta e individualista. A primeira o cálculo foi feito pressupondo que o indivíduo vai prevenir a morte dos outros na rede, enquanto o segundo o indivíduo quer prevenir a própria morte. Nesse sentido, abordagens altruístas revelaram-se estratégias melhores na redução da mortalidade e tempo hospitalizado, mas as individualistas foram melhores contra disseminação da doença. Por fim, foram identificadas as métricas de melhor desempenho, as métricas mais eficazes em média que foram o *Pagerank* e o *PageRank* com ponderação nas arestas, bem como a fronteira de Pareto para os dados analisados para as diferentes métricas, redes e valores de agrupamento.

Palavras-chave: redes (matemática); epidemiologia; COVID-19 (doença) - vacinação; medidas de centralidade.

ABSTRACT

The rapid spread, high mortality rate, and overloading of hospital systems turned the SARS-CoV-2 virus into a significant challenge for humanity. The urgency of the situation required substantial investments in vaccine development as a crucial measure to mitigate this issue. However, the financial costs and time needed for vaccine development represent obstacles to achieving comprehensive vaccination coverage. Therefore, understanding the most effective strategy for population immunization is of utmost importance to contain the virus's spread and optimize resource allocation in public policies. It is relevant to highlight that a large part of these policies was based on social distancing, which had significant impacts on economic, educational, and mental health spheres. The purpose of this project is to propose a modeling-based approach using the SEIHARDS contagion model on a complex contact network derived from a questionnaire-based survey. The objective is to identify optimal vaccination strategies to reduce the virus's spread, minimize hospitalizations, and mitigate the number of deaths. For this, data were collected from a survey called POLYMOD, conducted in 8 European countries, aiming to map connections between people. This survey is the most aligned with the project because it holds data on the age groups of the respondents, something quite relevant in COVID-19 hospitalization and mortality. However, the respondents of this survey had no connections among themselves, making it necessary to develop an algorithm called the Stratified Configuration Model (SCM) to form a network with this database limitation that takes into account ages and adds the duration of contact time between individuals. Additionally, the proposed model manages to increase the network's clustering because the traditional Configuration Model achieves low clustering. Moreover, by collecting data from articles and OpenDataSUS, it was possible to find the constants for the proposed epidemiological model, where some parameters vary with age group. This project analyzed the behavior of the proposed algorithm with changes in clustering, proposed centrality metrics for this study, and analyzed the behavior and sensitivity of the infection model concerning variations in clustering and network construction. It showed that the model is not very sensitive to increases in clustering and that with the insertion of weighting on the edges, there is no gain in the metrics and an increase of up to 5 times in computational time. The centrality measures that considered weights on nodes and the network structure proved to be highly effective for vaccination against COVID-19, with strategies like *PageRank* reducing mortality by more than 60%, total hospitalization time by 66%, and requiring vaccinating 31% of the network to inhibit the virus's proliferation compared to the absence of vaccination. Some

centralities that use weights on edges presented two calculation approaches: altruistic and individualistic. In the first, the calculation was done assuming that the individual will prevent the death of others in the network, while in the second, the individual wants to prevent their own death. In this sense, altruistic approaches proved to be better strategies in reducing mortality and hospitalization time, but the individualistic ones were better against the disease's spread. Finally, the best-performing metrics were identifiedthe most effective metrics on average were *PageRank* and *PageRank* with edge weightingas well as the Pareto frontier for the analyzed data for different metrics, networks, and clustering values.

Keywords: networks (mathematics); epidemiology; COVID-19 (disease) - vaccination.

LISTA DE FIGURAS

Figura 1 – Pontes de Königsberg	25
Figura 2 – Ilustração de uma rede não direcionada, das suas quantidades de ligações L , de nós N e da Matriz de Adjacência A	27
Figura 3 – Ilustração de Redes e suas Métricas Globais	30
Figura 4 – Ilustração do Modelo de Configuração, é escolhido dois nós i e j aleatoriamente e são conectados	38
Figura 5 – Esquematização do Modelo Suscetível-Infectado de contágio	40
Figura 6 – Modelo SI com $I(0) = 0.2$ e $\bar{\beta} = 0.4$. No início, há uma fração de infectados, e com o tempo, todos os indivíduos se tornarão infectados.	41
Figura 7 – Modelo SIS com $I(0) = 0.2$, $\bar{\gamma} = 0.3$ e $\bar{\beta} = 0.1$. Caso $\bar{\beta} < \bar{\gamma}$, com a taxa de recuperação maior que a taxa de infecção, para $t \rightarrow \infty$, o número de infectados se aproxima de zero e o de suscetíveis se aproxima de um, indicando a erradicação da doença.	42
Figura 8 – Modelo SIS com $I(0) = 0.2$, $\bar{\beta} = 0.3$ e $\bar{\gamma} = 0.1$. Há convergência para equilíbrio estável com $\bar{\beta} > \bar{\gamma}$, para $t \rightarrow \infty$, as frações de infectados e suscetíveis estabilizam em $1 - \frac{\bar{\gamma}}{\bar{\beta}}$ e $\frac{\bar{\gamma}}{\bar{\beta}}$ respectivamente.	43
Figura 9 – Modelo SIS com $I(0) = 0.2$, $\bar{\gamma} = 0.3$ e $\bar{\beta} = 0.1$. Caso $\bar{\beta} < \bar{\gamma}$, com a taxa de recuperação maior que a taxa de infecção, para $t \rightarrow \infty$, o número de infectados se aproxima de zero e o de suscetíveis se aproxima de um, indicando a erradicação da doença.	44
Figura 10 – Esquematização do Modelo Susceptível-Infectado-Removido de contágio .	44
Figura 11 – Modelo SIR com $I(0) = 0.2, S(0) = 0.8, R(0) = 0, \bar{\omega} = 0.4$ e $\bar{\beta} = 1.2$. É mostrado existe um ponto de máximo de $I(t)$, esse gráfico é bastante sensível aos parâmetros e seu comportamento pode mudar bastante com a alteração deles	45
Figura 12 – Ilustração do Modelo de Manzo e van Rijt que adiciona uma probabilidade de agrupamento p	54
Figura 13 – Ilustração do cálculo da taxa para se tornar infectado	58

Figura 14 – Esquematização do Modelo SEIHARDS de contágio. Os blocos em vermelho significam quem pode contaminar outros, os blocos em azul escuro mostra o compartimento ainda não infectado e em amarelo mostra o estágio na qual o indivíduo não interage mais com a rede. Os parâmetros em preto são as taxas e em roxo são as probabilidades de entrar no estágio	60
Figura 15 – Distribuição percentual de Faixas etárias do Brasil X POLYMOD	64
Figura 16 – Frequência e Duração de Interações (com/sem Contato Físico)	65
Figura 17 – Influência das idades nas conexões	66
Figura 18 – Frequência média de conexões por faixa etária	67
Figura 19 – Média e desvio-padrão dos tempos de contato entre faixas etárias	68
Figura 20 – Agrupamento da rede em função de p	69
Figura 21 – Variação da Correlação entre Agrupamento e Grau, Diâmetro em função de p	70
Figura 22 – Distribuição de Graus $p = 1.0$	71
Figura 23 – Resultados da diferença absoluta entre a matriz de proporções de conexões entre faixas etárias encontradas no modelo e no POLYMOD	72
Figura 24 – Fração de Suscetíveis, Infectados e Recuperados antes da aplicação de qualquer estratégia de vacinação	74
Figura 25 – Fração de Hospitalizados e Mortos antes da aplicação de qualquer estratégia de vacinação	75
Figura 26 – Fração de Expostos e Mortos antes da aplicação de qualquer estratégia de vacinação para diferentes valores de p	76
Figura 27 – Comparação entre a ponderação nas arestas e sem ponderação	77
Figura 28 – Gráfico para cada métrica por fração de vacinados para $p = 0.0$ e $p = 1.0$ sem ponderação nas arestas	79
Figura 29 – Gráfico para cada métrica por fração de vacinados para $p = 0.0$ e $p = 1.0$ com ponderação nas arestas	80
Figura 30 – Correlação entre o ranqueamento entre cada métrica	81
Figura 31 – Correlação entre o ranqueamento de cada métrica para cada agrupamento .	82
Figura 32 – Diferença relativa da centralidade com e sem peso nas arestas	83
Figura 33 – Diferença relativa da centralidade com e sem peso nas arestas	84
Figura 34 – Diferença relativa entre abordagem Individualista e Altruista	85

LISTA DE TABELAS

Tabela 1 – Valores de $\#Triângulos(v_i)$, $\#Tríades(v_i)$ e $C(v_i)$ do exemplo da Figura 3	30
Tabela 2 – Valores de diferentes métricas de centralidade calculados para a rede ilustrada na Figura 2, os valores foram calculados utilizando a biblioteca Networkx	34
Tabela 3 – Taxas de transição utilizadas no modelo SEIHARDS	60
Tabela 4 – Probabilidades de transição por faixa etária.	61
Tabela 5 – Tabela da variação dos parâmetros perante a vacinação	61
Tabela 6 – Melhores estratégias para cada valor de p na rede não ponderada. Há uma predominância de estratégias locais, ou seja, nós próximos do nó avaliado têm maior importância	86
Tabela 7 – Melhores estratégias para cada valor de p na rede ponderada. Há uma predominância da <i>WPagerank</i> (PR) e PR em relação as outras centralidades e ainda sim como no exemplo passado uma predominância de estratégias locais.	87
Tabela 8 – Tabela com todos os resultados para rede com arestas não ponderadas $p = 0.00$, em azul os menores valores e e em vermelho os maiores.	97
Tabela 9 – Tabela com todos os resultados para rede com arestas não ponderadas $p = 0.25$, em azul os menores valores e e em vermelho os maiores.	98
Tabela 10 – Tabela com todos os resultados para rede com arestas não ponderadas $p = 0.50$, em azul os menores valores e e em vermelho os maiores.	99
Tabela 11 – Tabela com todos os resultados para rede com arestas não ponderadas $p = 0.75$, em azul os menores valores e e em vermelho os maiores.	100
Tabela 12 – Tabela com todos os resultados para rede com arestas não ponderadas $p = 1.00$, em azul os menores valores e e em vermelho os maiores.	101
Tabela 13 – Tabela com todos os resultados para rede com arestas ponderadas $p = 0.00$, em azul os menores valores e e em vermelho os maiores.	102
Tabela 14 – Tabela com todos os resultados para rede com arestas ponderadas $p = 0.25$, em azul os menores valores e e em vermelho os maiores.	104
Tabela 15 – Tabela com todos os resultados para rede com arestas ponderadas $p = 0.50$, em azul os menores valores e e em vermelho os maiores.	106
Tabela 16 – Tabela com todos os resultados para rede com arestas ponderadas $p = 0.75$, em azul os menores valores e e em vermelho os maiores.	108

Tabela 17 – Tabela com todos os resultados para rede com arestas ponderadas $p = 1.00$,
em azul os menores valores e em vermelho os maiores. 110

LISTA DE ALGORITMOS

Algoritmo 1 – Modelo de Configuração	38
Algoritmo 2 – Algoritmo caso discreto do modelo SIS	47
Algoritmo 3 – Algoritmo caso contínuo do modelo SIS	48
Algoritmo 4 – Implementação de Manzo e van Rijt	54
Algoritmo 5 – Implementação do MCE	56
Algoritmo 6 – Implementação de Manzo e van de Rijt - Adaptado	57

LISTA DE ABREVIATURAS E SIGLAS

A-CGR	Centralidade de Gravidade com abordagem altruísta
CA	Centralidade de Autovetor
CB	Centralidade de Intermediação
CE	Centralidade de Excentricidade
CG	Centralidade de Grau
CGR	Centralidade de Gravidade com abordagem individualista
CH	Centralidade Harmônica
CI	Centralidade de Idade
CK	Centralidade <i>K-shell</i>
CL	Centralidade Laplaciana
CP	Centralidade de Proximidade
CR	Centralidade aleatória
GM	Centralidade de Grau Morte
PH	Probabilidade de pelo menos um dos vizinhos serem hospitalizados
PHA	Probabilidade de pelo menos um dos vizinhos serem hospitalizados dado que o sítio é assintomático
PM	Probabilidade de pelo menos um do sítio ser hospitalizado e morrer
PMA	Probabilidade de pelo menos um dos vizinhos morrerem dado que foram hospitalizados e que o sítio é assintomático
PR	<i>Pagerank</i>
UT	Utilidade

SUMÁRIO

1	INTRODUÇÃO	17
1.1	COVID-19	18
1.2	Estado da arte	19
1.3	Descrição do problema	21
1.3.1	<i>Objetivos</i>	22
1.4	Organização do trabalho	23
2	REFERENCIAL TEÓRICO	25
2.1	Conceitos fundamentais	26
2.2	Métricas estruturais de redes	27
2.2.1	<i>Métricas globais de redes</i>	28
2.2.2	<i>Métricas locais de centralidade</i>	30
2.2.3	<i>Métricas da literatura</i>	34
2.3	Modelos de formação de redes	35
2.3.1	<i>Rede de Erdős-Rényi</i>	36
2.3.2	<i>Modelo de Configuração</i>	37
2.4	Modelos de infecção ou difusão	38
2.4.1	<i>Modelo SI</i>	39
2.4.2	<i>Modelo SIS</i>	42
2.4.3	<i>Modelo SIR</i>	44
2.4.4	<i>Modelos em Redes</i>	46
2.4.5	<i>Modelo SEIHARD</i>	48
2.5	Fronteira de pareto	50
3	METODOLOGIA	52
3.1	Coleta de dados	52
3.2	Modelo de formação de redes proposto	53
3.3	Modelo de infecção proposto SEIHARDS	57
3.4	Estratégias de vacinação propostas	61
3.5	Ambiente computacional	63
4	RESULTADOS	64
4.1	Análise exploratória dos dados do POLYMOD	64

4.2	Resultados da rede gerada	68
4.3	Resultados do modelo de infecção	72
4.4	Resultado das estratégias de vacinação	77
5	CONCLUSÕES E TRABALHOS FUTUROS	88
	REFERÊNCIAS	90
	APÊNDICE A – RESULTADOS PARA TODAS AS ESTRATÉGIAS ..	97

1 INTRODUÇÃO

Desde as mais antigas civilizações humanas, elas têm enfrentado o problema de propagação de doenças em larga escala (Psychiatry..., 2019). Por exemplo, em 430 a.C. aconteceu a epidemia chamada Peste de Atenas que foi responsável pela morte de cerca de um terço da população dessa cidade, em 541 houve a primeira pandemia chamada Praga de Justiniano que ocorreu no mediterrâneo, e em 1347 aconteceu a mais devastadora pandemia na história da humanidade, a Peste Negra. Mais recentemente tivemos a Gripe Suína em 2009, a COVID-19 que aconteceu de 2019 a 2022 e a Sétima Pandemia da Cólica que teve início em 1961. Nesse sentido, o estudo de doenças se tornou cada vez mais necessário para cientistas seja para entender como uma doença afeta o nosso corpo, seja para modelar a propagação dela.

Com a expansão da humanidade nos últimos anos, impulsionada pelo comércio, desmatamento e turismo, houve uma maior interação entre humanos, e entre humanos e animais. Essa maior interação facilitou a propagação de doenças entre os países, o que pode ser classificado em quatro categorias distintas, de acordo com a taxa de contágio e a área de atuação. A primeira categoria é a endemia, que se refere a uma infecção com uma taxa de contágio controlada e previsível, atuando de forma contínua e restrita a uma determinada área geográfica ou população, podendo abranger desde uma cidade até um continente. A segunda categoria é o surto, que expressa um aumento repentino na ocorrência de casos da doença em pequenas áreas. Em seguida, temos a epidemia, que corresponde a um surto em grande escala. Finalmente, a pandemia é caracterizada como uma epidemia em escala mundial.

Dessa forma, a busca pelo entendimento das doenças emergiu como uma necessidade primordial desde os tempos mais remotos da história humana. Esse impulso de compreensão deu origem ao campo da Epidemiologia, uma disciplina voltada para a análise quantitativa e qualitativa dessas enfermidades. Essa área tem início com os estudos de Hipócrates (Beaglehole *et al.*, 2010) sobre como o ambiente favorece ou dificulta o surgimento de doenças. Apesar disso, foi apenas no século XIX que ganhou mais força com John Snow no estudo sobre a cólera em Londres. Snow descobriu que as mortes de cólera de 1848-49 e 1853-54 estavam relacionadas à água que os enfermos tomavam que era fornecida pela companhia Southpark. Outrossim, em 1760 Daniel Bernoulli foi o primeiro a tentar modelar matematicamente a propagação de doenças e as consequências da vacinação (Bernoulli e Blower, 2004; Dietz e Heesterbeek, 2002). Já no século XX, os modelos compartmentais de transmissão de doenças foram introduzidos (Kermack e McKendrick, 1927) e Richard Doll e Andrew Hill descobriram a

relação entre fumar e desenvolvimento de câncer (Doll e Hill, 1950).

Desde então a matemática mostrou-se cada vez mais essencial para o estudo de doenças não apenas pela análise de dados, mas também na modelagem, pois não é eticamente correto fazer experimentações utilizando doenças, principalmente em humanos, então modelar matematicamente se torna essencial para previsões de propagação. Ademais, o entendimento da biologia por trás também é necessário para ter um modelo mais verossímil e seus resultados tenham validação na realidade. Por fim, todo o conhecimento construído até aqui foi essencial para o entendimento e previsão da pandemia do novo Coronavírus que durou de 11 de março de 2020 até 5 de maio de 2023.

1.1 COVID-19

Em dezembro de 2019, alguns casos de uma pneumonia de causas desconhecidas surgiram em Wuhan, capital da província de Hubei, um grande centro de transporte da China (Singhal, 2020). Os pacientes que apresentaram esses sintomas faziam parte de um mercado de animais marinhos. Autoridades de saúde da China foram acionadas e foram coletadas amostras do patógeno para investigação e para caracterizar e controlar o desconhecido patógeno. No dia 31 de dezembro, o país notificou a OMS sobre o surto da doença e no dia seguinte o mercado da região foi fechado. No dia 7 de janeiro, cientistas conseguiram isolar o agente infecioso (Wang *et al.*, 2020) e foi identificado com mais de 95% de semelhança com o coronavírus apresentado em morcegos e mais de 70% com SARS-CoV. No dia 23 de janeiro, já havia casos espalhados em 32 províncias da China e no mesmo dia toda a população de Wuhan foi colocada em *lockdown*. Contudo, dois dias depois aconteceria a comemoração do Ano Novo Chinês e várias pessoas de outros países haviam viajado com intuito de participar do evento. No dia 11 de março, a OMS declara emergência de saúde pública do chamado Coronavírus ou COVID-19 e somente em 5 de maio de 2023 deixa de se tornar uma ameaça global.

Para evitar a proliferação do vírus, várias políticas públicas foram implementadas em todo o mundo. Estas políticas incluíam restrições de mobilidade, como *lockdowns* e quarentenas, a fim de reduzir a interação social e a transmissão do vírus. Além disso, a promoção do uso de máscaras faciais, a intensificação da testagem e rastreamento de contatos, bem como a implementação de medidas de distanciamento físico em ambientes públicos, foram amplamente adotadas (Cruz *et al.*, 2021).

No entanto, a centralização nas políticas de distanciamento social, que impuseram

a necessidade de confinamento domiciliar à população, deu origem a desafios de natureza econômica (Irawan e Alamsyah, 2021). Isso se deveu, em grande parte, à impossibilidade de uma parcela substancial da força de trabalho desempenhar suas funções, resultando em um aumento significativo nas taxas de desemprego e na deterioração das condições sociais, com notável impacto negativo na qualidade da educação (Gonzalez *et al.*, 2020). O fechamento de instituições de ensino, como escolas e universidades, que são locais propensos a grandes aglomerações, agravou ainda mais essas questões. Como resultado desse contexto complexo, emergiu uma preocupação crítica com a saúde mental (Pereira *et al.*, 2020).

Devido a isso a campanha de vacinação em massa também se tornou uma política crucial para o fim da pandemia e minimizar o impacto dela na saúde pública e na economia. Um dos principais desafios da vacinação é a complexidade do processo de fabricação, que requer instalações especializadas, reagentes e tecnologia de ponta. Além disso, a demanda mundial por vacinas tem sobre carregado a capacidade de produção existente, levando a atrasos na entrega de doses. A logística de distribuição também é complicada, com necessidade de armazenamento em temperaturas específicas para algumas vacinas, o que requer infraestrutura adequada. A obtenção de matérias-primas e ingredientes essenciais pode ser afetada por interrupções na cadeia de suprimentos global, causando a escassez de insumos.

1.2 Estado da arte

Paralelamente aos avanços na prevenção da pandemia, houve um considerável aprimoramento dos modelos epidemiológicos, resultando em mais sofisticados e especializados na compreensão da propagação de doenças infecciosas (Xiang *et al.*, 2021). Esses modelos têm sido essenciais para orientar estratégias de controle da COVID-19, incluindo a implementação de programas de vacinação (Scabini *et al.*, 2021; Bustamante-Castañeda *et al.*, 2021; Loyal e Chen, 2020). No entanto, é importante observar que os modelos epidemiológicos tradicionais, embora sejam valiosos para prever tendências em larga escala, podem não capturar detalhes específicos das interações locais e complexas que ocorrem em estruturas de rede (Pellis *et al.*, 2015).

Os modelos epidemiológicos baseados em redes representam uma abordagem que enfatiza a análise das interações individuais em sistemas complexos, buscando compreender a propagação de doenças em escalas locais e com detalhes de rede. Esses modelos consideram a estrutura da rede, que descreve as conexões interpessoais e os contatos sociais, e incorporam processos de transmissão de doenças em estruturas de rede realistas (Pei *et al.*, 2023).

(Pastor-Satorras e Vespignani; Eames e Keeling, 2002, 2003) utilizam o modelo de redes com propriedades de pequeno mundo para estudar a imunização e encontram expoentes críticos para a evolução da infecção. Enquanto que (Salathé e Jones; Kitsak *et al.*; Gong *et al.*; Miller e Hyman, 2010, 2010, 2013, 2007) estudam formas de encontrar os sítios centrais na propagação das doenças utilizando algumas métricas de redes.

No entanto, é relevante destacar que essa abordagem ainda não incorporou completamente os avanços em complexidade (Eikenberry *et al.*, 2020) alcançados pelos modelos epidemiológicos que não utilizam informações de contatos (Pellis *et al.*, 2015). Esses avanços estão relacionados à uma maior especialização em relação à doença com a introdução de novos componentes relevantes, como a distinção entre os sintomáticos e os assintomáticos, além de considerar aspectos relacionados à reação da sociedade, como a implementação de medidas de quarentena.

Entretanto os modelos de redes têm avançado por outras frentes como por exemplo a vacinação de uma população de uma mesma doença com cepas diferentes, (Li *et al.*, 2023) em estudo na qual a população se adapta à propagação da doença se utilizando de isolamento, (Silva *et al.*, 2023) estudam em redes com distribuição de grau em lei de potência e aplicando a teoria de campo médio para prever o comportamento da epidemia, em estudo da importância de viagens entre cidades apresentado por (Quiroga e Schimit; Rossa *et al.*, 2023, 2020), do transporte com (Scabini *et al.*, 2021), distanciamento social por (Maheshwari e Albert, 2020) e agrupamento que (Craig *et al.*, 2020) apresenta para o espalhamento da doença. Além disso, ainda no que tange COVID e vacinação existem estudos utilizando redes e modelos de infecção para estudar o espalhamento da aceitação ou não aceitação da vacinação como o de (Mallick *et al.*, 2023).

Esses estudos sobre a COVID-19 têm explorado diversas abordagens para entender a dinâmica da pandemia, desde a previsão de curvas de infecção (Xiang *et al.*, 2021) até a avaliação de estratégias de intervenção (Liu e Lou, 2022; Kitsak *et al.*, 2010), como distanciamento social, *lockdowns* e, principalmente, campanhas de vacinação. Através da análise de eficácia, cobertura vacinal e estratégias de implementação, é possível avaliar o impacto das vacinações em larga escala na redução da transmissão de patógenos e na proteção da saúde da população. Além disso, esses estudos contribuem para a tomada de decisões informadas sobre quais vacinas priorizar, quais grupos populacionais devem ser alvo e como otimizar recursos limitados. O estudo de estratégias de vacinação para COVID-19 não é novidade na literatura (Doostmohammadian *et*

al., 2020; Tetteh *et al.*, 2021; Petrizzelli *et al.*, 2022), entretanto os estudos estavam limitados a: redes tradicionais com pouca base no real, modelos epidemiológicos pouco especializados com a doença, sem levar consideração as idades dos indivíduos ou tempo de contato entre eles e o estudo de poucas métricas de centralidade.

Uma dificuldade para a utilização de redes para o estudo de dinâmica de infecção é a dificuldade de se obter dados sobre as redes reais de contato entre pessoas, que são complexas e exigem múltiplas interações para serem mapeadas de forma eficaz. Muitos bancos de dados disponíveis são limitados em tamanho e diversidade, frequentemente baseando-se em amostras pequenas que não capturam a totalidade das interações sociais. Além disso, esses dados geralmente não consideram atributos essenciais das conexões e dos indivíduos, como a frequência e a duração das interações, bem como características demográficas que podem influenciar a propagação de infecções.

Nesse sentido, utilizar Modelos de geração de redes, como os propostos por Erdős-Rényi, Watts-Strogatz, Barabási-Albert e o Modelo de Configuração (Newman, 2018), são uma alternativa a esse problema. Eles foram muito explorados devido à sua capacidade de representar diferentes características estruturais das redes. No entanto, esses modelos clássicos foram desenvolvidos principalmente para redes não ponderadas, onde as conexões entre os nós são tratadas de forma binária, desconsiderando a intensidade ou a frequência das interações. Essa limitação é particularmente significativa, pois fatores como o tempo de contato entre as pessoas e características intrínsecas, como a idade, desempenham um papel crucial na formação e dinâmica das redes.

Em contraste, o *Stochastic Block Model* (SBM) (Lee e Wilkinson, 2019) foi proposto como uma abordagem para redes ponderadas, oferecendo uma forma de reproduzir a estrutura comunitária ao modelar as probabilidades de conexões entre blocos de nós. No entanto, embora o SBM seja eficaz para identificar agrupamentos e padrões médios dentro da rede, ele apresenta limitações significativas quando se trata de preservar a variabilidade e a distribuição de graus e pesos das arestas.

1.3 Descrição do problema

A incorporação de modelos epidemiológicos complexos em conjunto com a modelagem de redes, considerando a estratificação etária, representa um desenvolvimento metodológico que, em grande parte, não havia sido abordado satisfatoriamente na literatura. Essa abordagem

combinada possibilita uma análise mais precisa e abrangente das interações sociais e da disseminação de doenças, levando em conta as diferenças demográficas relacionadas à idade. Ao fazer isso, abrem-se oportunidades para melhorar a modelagem de surtos epidêmicos e a eficácia das estratégias de intervenção, com foco na melhoria da vacinação em grupos específicos.

Vale ressaltar que, historicamente, a identificação de nós centrais em redes epidemiológicas dependia majoritariamente de métricas que têm a informação da estrutura da rede (Miller e Hyman, 2007; Kitsak *et al.*, 2010; Salathé e Jones, 2010). Alguns trabalhos recentes têm focado em abordagens que levam em consideração características dos indivíduos (Chen *et al.*, 2021; Klise *et al.*, 2022), mas ainda sim são bastante restritos.

Por fim, a implementação das campanhas de vacinação contra a COVID-19 em diferentes países, embora variem em detalhes, refletem uma combinação de considerações etárias, econômicas e de exposição ao risco, em vez de uma priorização exclusiva focada na otimização da prevenção da disseminação do vírus e na redução da mortalidade (Huh *et al.*, 2021; Rosen *et al.*, 2021; Cadeddu *et al.*, 2022; Jung, 2021). Enquanto alguns países, como Israel, adotaram critérios de vacinação simplificados, visando simultaneamente os indivíduos com maior risco de morte e de hospitalização, outros países europeus seguiram diretrizes da OMS para contextos de oferta limitada de vacinas, priorizando inicialmente profissionais de saúde e residentes de lares de idosos.

1.3.1 *Objetivos*

Este trabalho tem como principal objetivo estudar estratégias de vacinação contra COVID-19 baseadas em métricas de centralidade de vértices em redes complexas. Em particular, os objetivos específicos são os seguintes:

- Revisar a literatura sobre modelos epidemiológicos de COVID-19, sobre a rede de contatos entre pessoas e suas faixas etárias e sobre os parâmetros que são necessários para modelar a propagação da doença e como a vacina interfere nisso;
- Desenvolver um modelo de redes capaz de simular com precisão as nossas conexões sociais, considerando os diferentes padrões de interação entre indivíduos de distintas faixas etárias;
- Encontrar um modelo preciso para a propagação da doença, hospitalizações, doenças e efeitos da vacina;
- Encontrar valores das taxas do modelo epidemiológico e como ele se altera com a vacina-

ção;

- Fazer comparação de diversas estratégias de vacinação baseada em métricas de redes;
- Propor novas métricas que combinem estrutura e propriedades dos nós.

1.4 Organização do trabalho

Este trabalho está estruturado em cinco capítulos e uma versão preliminar do mesmo foi apresentada e publicada nos anais do Simpósio Brasileiro de Pesquisa Operacional (SBPO) (Gonçalves *et al.*, *In: ANAIS DO SIMPÓSIO BRASILEIRO DE PESQUISA OPERACIONAL*, 55.,2023, São José dos Campos, **Anais eletrônicos** ..., Galoá, 2023). O segundo capítulo deste trabalho é dedicado a uma revisão teórica sobre redes, englobando a exploração de parâmetros essenciais, métricas de centralidade, diferentes modelos de redes e modelos epidemiológicos. Serão abordados conceitos fundamentais que fornecem a base teórica necessária para a compreensão da dinâmica das redes, sua estrutura e sua aplicabilidade em contextos epidemiológicos. Além disso, serão apresentados modelos que permitem simular a propagação de doenças em redes.

No terceiro capítulo, é detalhada a metodologia adotada para este estudo. Isso inclui uma descrição do banco de dados selecionado, justificando sua escolha e fornecendo informações relevantes sobre sua estrutura e conteúdo. É apresentado também o modelo de redes utilizado, destacando suas características e como ele foi aplicado para representar a interconexão entre indivíduos em um contexto específico. Além disso, o modelo epidemiológico escolhido é discutido em detalhes, assim como a integração dos dados provenientes da literatura sobre a COVID-19, os registros de vacinação da Pfizer-BioNTech e os dados do OpenDataSUS para informar e validar o modelo.

No quarto capítulo, são discutidos os resultados proeminentes da análise dos dados da pesquisa utilizada para construir o modelo em redes, resultados dos modelos de redes, epidemiológicos e das estratégias de vacinação apontando aquelas que foram melhores e como variaram perante um aumento do agrupamento e a ponderação ou não de arestas com o tempo de contato entre indivíduos. Nele mostramos os resultados de cada métrica de centralidade perante três aspectos: fração do número de mortos, número total de dias em que os indivíduos ficaram hospitalizados e qual a fração de indivíduos é necessária vacinar para erradicar a doença. Um dos principais resultados é que o aumento do agrupamento tem uma mudança nos resultados da infecção, enquanto que a ponderação nas arestas é algo que não os altera significantemente.

O *Page Rank* ponderado com arestas foi a melhor centralidade apresentada e ela conseguiu erradicar a infecção com 45% de pessoas vacinadas enquanto que usando o critério idade, o principal parâmetro usado como estratégia de vacinação em vários países, foi preciso vacinar 85% das pessoas.

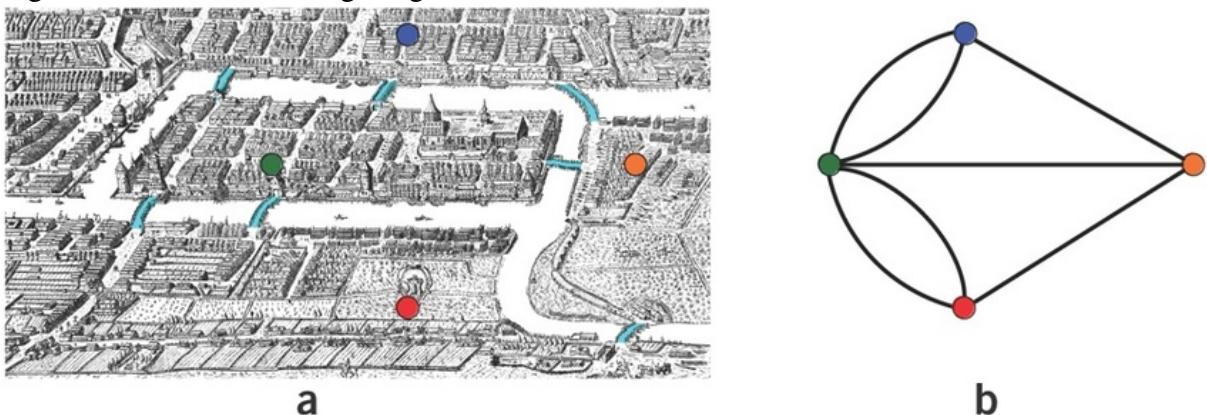
No último capítulo, serão apresentados as conclusões da dissertação descrevendo os resultados de forma resumida, as limitações desse trabalho e quais são os próximos passos para avançar nesta pesquisa. Ao final da dissertação é apresentado um apêndice com todos os resultados relevantes produzidos por esse trabalho.

2 REFERENCIAL TEÓRICO

A principal ferramenta para modelar o nosso problema surgiu em 1736. Nessa época existia a cidade de Königsberg (atual Kaliningrado), representada pela Figura 1, por ela 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 entre os moradores: seria possível sair de uma região e voltar para ela passando por todas as pontes apenas uma vez?

Leonhard Euler (Euler, 1736) 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 ligações entre esses pontos que representavam as pontes entre as regiões 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 não é possível sair de uma região e voltar para ela passando uma única vez em cada ponte de Königsberg.

Figura 1 – Pontes de Königsberg



Fonte: Chatterjee, Attanu (Chatterjee, 2015).

A prova de Euler mostra que quando se quer 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 Teoria dos 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 na computação Grafos estão relacionados à problemas de fluxo (Szwarcfiter, 2018; Goldbarg, 2012) e redes são utilizados para problemas

visando a estrutura do Grafo e suas interações (Barabási, 2016; Newman, 2018). Nesse trabalho usaremos os dois como sinônimos.

2.1 Conceitos fundamentais

Um Grafo $G(\mathcal{N}, \mathcal{L})$ é uma dupla na qual $\mathcal{N} = \{v_0, v_1, v_2, \dots, v_i, \dots\}$ é um conjunto não vazio de elementos chamados de vértices ou nós e \mathcal{L} é um conjunto não vazio de pares de elementos de \mathcal{N} chamados de arestas ou ligações. Dessa forma podemos definir $N = |\mathcal{N}|$ que mensura a quantidade de vértices que há em G e $L = |\mathcal{L}|$ que mensura o número de arestas. Essas quantidades também são chamadas de ordem e tamanho (Goldberg, 2012), respectivamente.

A partir dessa definição de redes é possível definir a Matriz de Adjacência que guarda a informação de conexões entre os nós. Seja uma matriz A quadrada de tamanho $N \times N$, cada elemento da matriz segue a seguinte regra:

$$A_{i,j} = \begin{cases} 1 & \text{se } (v_i, v_j) \in \mathcal{L} \\ 0 & \text{caso contrário.} \end{cases}$$

A partir dela surgem características importantes de redes. Se $A_{i,j} = A_{j,i} \forall v_i, v_j \in \mathcal{N}$ então a rede é chamada de não direcionada; se $\exists v_i \neq v_j$ tal que $A_{i,j} \neq A_{j,i}$ ela é chamada de direcionada. Isso pode ter diferentes interpretações a depender do contexto, 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 no segundo usamos redes direcionadas.

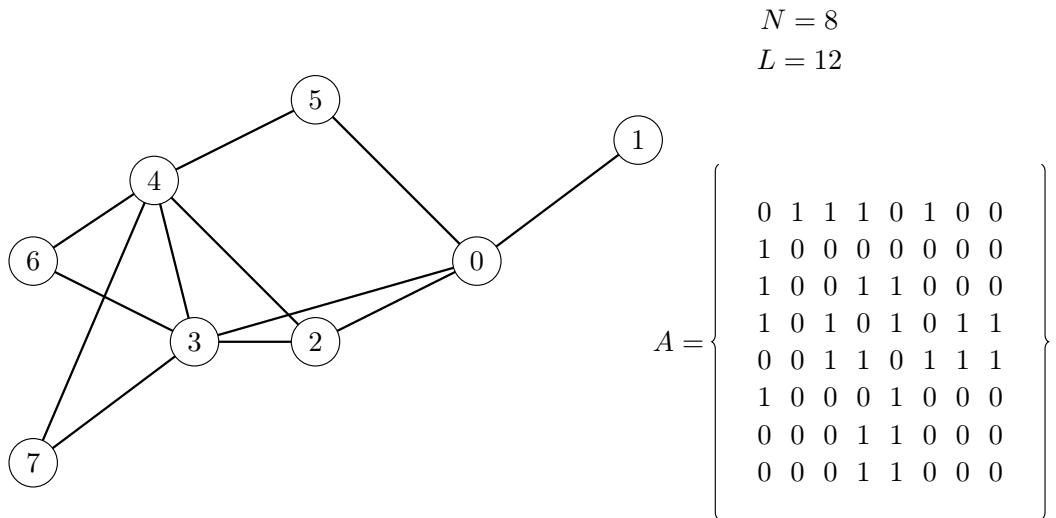
Uma alternativa para registrar a informação da Matriz de Adjacência é através da lista de vizinhos. Dois vértices v e μ são vizinhos se existe uma ligação entre v e μ , assim é denotado $\eta(v)$ o conjunto de vizinhos do vértice v para redes não direcionadas. Em redes direcionadas, existem dois tipos de vizinhos: os que se conectam a um nó e os que são conectados por um nó. Portanto, denotaremos $\eta_{in}(v)$ todos os vizinhos que têm ligação que chega em v e $\eta_{out}(v)$ todos os vizinhos que têm ligação que sai de v . A Figura 2 mostra um exemplo de uma rede não direcionada com 8 nós enumerados de 0 a 7, com 12 arestas e a matriz de adjacência $A_{i,j}$.

Por fim, é possível adicionar propriedades às nossas ligações. As propriedades 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. Além disso, pode-se também adicionar atributos aos vértices, representando alguma característica não-topológica dos mesmos, como, exemplo, idade de um indivíduo. Formalmente, um Grafo ponderado $G(\mathcal{N}, \mathcal{L}, \Theta, \Xi(v_i, v_j))$ é formado por um conjunto de nós \mathcal{N} , um conjunto de arestas \mathcal{L} e mapeamentos $\Theta : \mathcal{N} \mapsto \mathcal{S}$ e $\Xi(v, \mu) : \mathcal{L} \mapsto \mathcal{R}$ nos quais \mathcal{R} e \mathcal{S} são propriedades dos nós e das arestas, respectivamente, que podem ser quantitativas ou qualitativas, vetoriais ou não. Nesse caso é definida uma generalização da Matriz de Adjacência, seja $w_{v,\mu}$ o peso da ligação (v, μ) então definimos a Matriz de Adjacência Ponderada como:

$$W_{i,j} = \begin{cases} w_{v_i, v_j} & \text{se } (v_i, v_j) \in \mathcal{L} \\ 0 & \text{caso contrário.} \end{cases}$$

Figura 2 – Ilustração de uma rede não direcionada, das suas quantidades de ligações L , de nós N e da Matriz de Adjacência A



Fonte: Elaborado pelo autor.

2.2 Métricas estruturais de redes

Como discutido anteriormente, é muito importante estudarmos a estrutura da rede. Para um entendimento melhor dessa estrutura existem várias métricas (Costa *et al.*, 2007) relacionadas à topologia da rede podendo ser classificadas como globais ou locais. Métricas globais de redes são medidas que fornecem informações sobre as propriedades e características de uma rede como um todo, em oposição às métricas locais que se concentram em elementos individuais dentro da rede, como nós ou arestas. Ao final desta seção, a Figura 3 e as Tabelas 1-2 apresentam um exemplo com aplicações dessas métricas.

2.2.1 Métricas globais de redes

Sejam dois vértices $v, \mu \in \mathcal{N}$ é dito que existe um caminho entre v e μ se existe uma sequência de vértices $\zeta_1, \zeta_2, \dots, \zeta_m | (\zeta_s, \zeta_{s+1}) \in \mathcal{L}, \zeta_1 = v, \zeta_m = \mu$, na qual não há repetição de vértices ou arestas. O comprimento de um caminho é dado por sua quantidade de arestas. Podem existir vários caminhos entre v e μ , portanto a distância geodésica entre dois vértices $d_{v,\mu}$ é definida como o menor comprimento dentre todos os caminhos entre v e μ . Mediante essa definição, é possível construir a métrica global $\langle d \rangle$ chamada distância geodésica média da rede:

$$\langle d \rangle = \sum_{v \neq \mu} \frac{d_{v,\mu}}{N \cdot (N - 1)}. \quad (2.1)$$

Esse valor carrega a informação sobre o quanto em média é necessário caminhar na rede de um nó para outro. No caso de redes ponderadas para esse caminho é atribuído um peso $\sum_s w_{\zeta_s, \zeta_{s+1}}$, portanto o menor caminho é aquele que minimiza $\sum_s w_{\zeta_s, \zeta_{s+1}}$ e seu valor será a distância geodésica $d_{v,\mu}^w$. Contudo a concepção de menor caminho entre dois pontos pode depender do problema, por exemplo para se achar a resistência equivalente em uma rede de resistores pode se utilizar da mínima resistência ou de máxima corrente. Portanto, em situações em que um maior peso indica uma maior proximidade entre os nós, a distância geodésica entre dois nós é dada pela soma dos recíprocos dos pesos das arestas do caminho geodésico.

Seja G' um grafo $G'(\mathcal{N}', \mathcal{L}')$ tal que $\mathcal{N}' \subseteq \mathcal{N}$ e $\mathcal{L}' \subseteq \mathcal{L}$, G' então é definido como subgrafo de G , pois contém vértices e arestas que também estão em G . Assim uma componente de um grafo G é um subgrafo G' que satisfaz duas condições: conectividade interna e maximalidade. Conectividade interna significa que, para qualquer par de vértices $\mu, v \in \mathcal{N}'$, existe um caminho em G' que conecta μ a v , ou seja, o subgrafo G' é conexo. Maximalidade implica que não é possível adicionar nenhum outro vértice ou aresta ao subgrafo G' sem perder a propriedade de conectividade interna, isto é, G' é maximal com respeito à conectividade.

Outro valor importante é o chamado diâmetro $d = \max_{v,\mu} d_{v,\mu}$ que é a maior distância geodésica da rede, ele reflete o maior número de etapas necessárias para viajar entre os dois vértices mais distantes na rede, fornecendo uma medida da “extensão” da rede em termos da sua conectividade. Esse valor está relacionado a um tipo de rede chamada redes de pequeno mundo, (Watts e Strogatz, 1998) definem que uma rede é dita de pequeno mundo quando $\langle d \rangle \propto \log(N)$, ou seja redes com essa propriedades apresentam uma pequena distância entre qualquer par de nós e facilmente a informação pode propagar na rede.

Uma outra métrica a ser avaliada é o número de conexões que cada nó contém. É

definido o grau k_v como o número de conexões (ou vizinhos) que o nó v possui para redes não direcionadas. Para redes direcionadas existe uma divisão entre k_v^{in} e k_v^{out} para os nós que pertencem a $\eta_{in}(v)$ e $\eta_{out}(v)$, respectivamente. Pode-se achar os valores k_v^{in} e k_v^{out} a partir da Matriz de Adjacência da seguinte forma:

$$k_v^{in} = \sum_j A_{j,i}, \quad k_v^{out} = \sum_j A_{i,j}. \quad (2.2)$$

Nessa formulação é possível definir o grau em redes ponderadas k_v^w substituindo $A_{i,j}$ por $W_{i,j}$. Além disso é possível também definir o grau médio $\langle k \rangle$ da rede não direcionada e não ponderada

$$\langle k \rangle = \sum_{v \in \mathcal{N}} \frac{k_v}{N} = 2 \frac{L}{N}. \quad (2.3)$$

Esse valor tem bastante importância quantitativa quando o objetivo é trabalhar com distribuições de graus. Analogamente, pode-se definir o grau de entrada ou de saída médio, bem como o grau ponderado médio da rede.

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.

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

$$rc(G) = \frac{\sum_{i,j} A_{i,j} A_{j,i}}{\sum_{i,j} A_{i,j}}. \quad (2.4)$$

Por fim existe uma métrica para analisar o Agrupamento de uma Rede, ou seja, quanto que os vizinhos de um vértice estão unidos. Essa métrica será definida a partir da probabilidade de um vizinho do nó v_i estar conectado com outro vizinho de v_i

$$C(G) = \frac{\sum_{(i,j,k):i \neq j \neq k} A_{i,j} A_{j,k} A_{k,i}}{\sum_{(i,j,k):i \neq j \neq k} A_{i,j} A_{j,k}} = 3 \frac{\#\text{Triângulos}}{\#\text{Tríades}}, \quad (2.5)$$

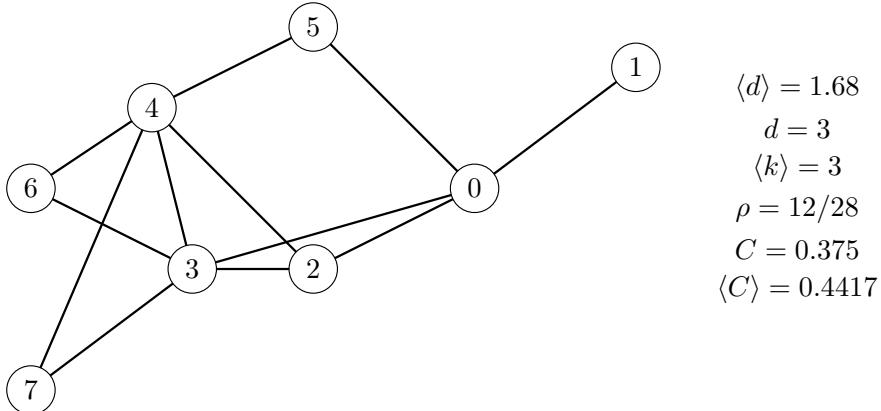
em que $\#$ significa número de alguma coisa, no caso estamos contando o número de triângulos e o número de tríades que aparecem na rede. Essa é a definição do que chamamos de Agrupamento

Global. Podemos também definir o Agrupamento Local $C(v)$ e por ele o Agrupamento Médio $\langle C \rangle$:

$$C(v) = 2 \frac{\#\text{Triângulos que contém o nó } v}{\#\text{Tríades que contém o nó } v}, \quad \langle C \rangle = \sum_{v \in \mathcal{N}} \frac{C(v)}{N}. \quad (2.6)$$

A Figura 3 ilustra uma rede e suas Métricas Globais: menor caminho médio $\langle d \rangle$, diâmetro d , grau médio $\langle k \rangle$, densidade ρ , Agrupamento C , Agrupamento Médio $\langle C \rangle$. Os valores do $\#\text{Triângulos}$, $\#\text{Tríades}$ e o Agrupamento se encontram na Tabela 1.

Figura 3 – Ilustração de Redes e suas Métricas Globais



Fonte: Elaborado pelo autor.

Tabela 1 – Valores de $\#\text{Triângulos}(v_i)$, $\#\text{Tríades}(v_i)$ e $C(v_i)$ do exemplo da Figura 3

nó	0	1	2	3	4	5	6	7
$\#\text{Triângulos}(v_i)$	1	0	2	4	3	0	1	1
$\#\text{Tríades}(v_i)$	6	0	3	10	10	1	1	1
$\#C(v_i)$	0.167	0.000	0.667	0.400	0.300	0.000	1.000	1.000

Fonte: Elaborado pelo autor.

2.2.2 Métricas locais de centralidade

Ao estudar redes o interesse está voltado em entender suas estruturas e propriedades, isso já é possível dadas as métricas que apresentamos anteriormente. É importante destacar que essas métricas globais oferecem uma análise abrangente da topologia da rede, revelando suas propriedades e comportamentos coletivos que podem estar diluídas na rede ou concentradas em nós específicos. Portanto estudar os nós em si e a sua importância perante a rede é necessário; para isso são usadas as chamadas Métricas Locais de Centralidade.

Quando se aborda o tema das Métricas de Centralidade, o foco está em analisar a importância topológica de cada nó dentro da rede. Contudo, em que sentido essa importância deve

ser considerada? A resposta a essa questão é dependente do contexto do problema em análise. Por exemplo, em uma rede de citações acadêmicas, pode ser de interesse identificar qual autor possui maior relevância (Leydesdorff e Rafols, 2011), enquanto em uma rede criminosa, o objetivo pode ser determinar qual indivíduo deve ser removido para que a rede seja desarticulada (Andrade *et al.*, 2021). Para cada um desses casos, assim como para muitos outros, existem métricas específicas que são adequadas para alcançar os objetivos propostos.

A métrica inicial a ser analisada consiste na contagem das conexões individuais de cada nó, representada pelo grau k_v . Ao conduzir essa avaliação, identificamos o nó central como aquele que apresenta o grau k_v mais elevado, objetivando, assim, encontrar o nó com a maior quantidade de conexões que é a Centralidade de Grau (CG).

A segunda métrica importante é a Centralidade de Excentricidade (CE). Essencialmente, a CE de um nó é determinada calculando-se o recíproco da maior distância que separa esse nó de qualquer outro nó na rede. Ela exibe a distância máxima que um nó precisa percorrer para alcançar o ponto mais remoto da rede. O nó cuja maior distância é a menor entre todos os nós da rede é então identificado como o mais central. Formalmente $CE(v) \forall v \in \mathcal{N}$ é definida como:

$$CE(v) = \frac{1}{\max_{\mu} d_{v,\mu}}. \quad (2.7)$$

Nessa mesma ideia é propício avaliar outras duas métricas parecidas: Centralidade de Proximidade (CP) e Centralidade Harmônica (CH). Em ambas o interesse é 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 calculada pelo inverso da média aritmética e a segunda com a média harmônica:

$$CP(v) = \frac{N-1}{\sum_{\mu \neq v} d_{v,\mu}}, \quad (2.8)$$

$$CH(v) = \frac{1}{N-1} \sum_{\mu \neq v} \frac{1}{d_{v,\mu}}. \quad (2.9)$$

Uma generalização da CP e CE é a Centralidade de Média- p (C_p) (Andrade e Rêgo, 2019). A ideia desta medida de centralidade é usar a noção de média generalizada das distâncias.

Ela é definida como:

$$C_p(v) = \begin{cases} \left(\frac{\sum_{\mu \neq v} d_{v,\mu}^p}{N-1} \right)^{-\frac{1}{p}} & \text{se } p \neq 0, \\ \left(\prod_{\mu \neq v} d_{v,\mu} \right)^{-\frac{1}{N-1}} & \text{se } p = 0. \end{cases} \quad (2.10)$$

Entretanto a importância de um nó não está associada somente a sua distância com os outros nós da rede, existem aqueles que são importantes por intermediarem conexões entre nós. Por exemplo, considere a rede mundial de 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 a única conexão entre o exterior e o país.

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

$$CB(v) = \frac{1}{(N-1)(N-2)} \sum_{\mu \neq \zeta} \frac{\mathbb{L}_{\mu,\zeta}^v}{\mathbb{L}_{\mu,\zeta}}, \quad (2.11)$$

em que μ e ζ são diferentes do nó v .

Ao analisar uma rede empresarial, é essencial compreender a relevância do CEO dentro da organização. Contudo, é importante observar que o CEO mantém comunicação direta apenas com os líderes de cada região da empresa, enquanto esses líderes nem sempre se comunicam diretamente entre si. Diante disso, surge a questão: onde se manifesta a importância do CEO? A resposta a essa indagação reside no fato de que o CEO estabelece conexões com indivíduos de grande influência dentro da rede, um aspecto que não é capturado pelas métricas de análise de rede discutidas até então. Nesse contexto, introduzimos a métrica de Centralidade de Autovetor (CA) como:

$$CA(v_i) = \frac{1}{\lambda} \sum_{v_j} A_{i,j} CA(v_j), \quad (2.12)$$

na qual λ é o maior autovalor associado à matriz de adjacência de rede. A importância dessa conexão é que os autovetores e autovalores capturam informações sobre como a informação flui na rede. Se um nó estiver conectado a outros nós importantes (ou seja, nós com alta centralidade), ele terá uma alta centralidade de autovetor, mesmo que tenha poucas conexões diretas. Em contrapartida, se um nó estiver conectado principalmente a nós periféricos, terá uma baixa centralidade de autovetor.

Com o advento das páginas da *internet* é preciso metrificar quais são os *sites* mais importantes. Para isso seria importante não só levar em consideração quem está em posição privilegiada, mas também como essa posição é mitigada pelo número de ligações, portanto (Brin

e Page, 1998) publicaram uma nova métrica que consegue quantificar isso chamada de PR. Essa métrica é utilizada pelo *Google* para analisar a importância de *sites* na *internet*. Considere um usuário que navega pela *internet* e está em um *site* v_j e deseja ir para outro. Existe uma probabilidade α de o usuário seguir um *link* dentro do *site* atual e ir para outro que ele aponta, e uma probabilidade $1 - \alpha$ de ele não seguir essa estrutura e ir para um *site* aleatório. Eles definiram o *PageRank* do nó v_j como:

$$PR(v_j) = 1 - \alpha + \alpha \times \sum_{j \neq i} A_{i,j} \frac{PR(v_i)}{k_{v_j}^{out}}. \quad (2.13)$$

(Carmi *et al.*, 2007) formularam o algoritmo *K-Shell* que se concentra em decompor a rede não direcionada em camadas, onde cada camada representa um subconjunto de nós com uma centralidade mínima, conhecida como o “grau mínimo” K. Inicialmente, os nós são ordenados por grau em ordem decrescente, e aqueles com menor grau são iterativamente removidos até que nenhum nó com grau menor ou igual a K permaneça na rede. Esse processo gera uma sequência de camadas (*K-Shells*), onde as camadas mais internas contêm nós de alta centralidade, enquanto as camadas mais externas contêm nós de menor centralidade. O seguinte algoritmo realiza a descomposição em *K-Shells*:

1. É inicializado $k = 1$. Este é o nível inicial de k-shell.
2. Encontre e remova todos os vértices de grau k . Atribua a esses vértices o k-shell k .
3. Após a remoção, reduza em 1 o grau de todos os vértices adjacentes aos removidos.
4. Se ainda existirem vértices de grau menor ou igual a k no grafo residual, repita 2 e 3 para esses vértices.
5. Incremente k para $k + 1$ e verifique se existem vértices de grau k no grafo residual. Se sim, repita os passos 2 a 4.
6. Continue o processo até que não reste nenhum vértice no grafo.
7. O algoritmo termina quando todos os vértices são atribuídos a um k-shell e o grafo está completamente decomposto.

A Tabela 2 resume os valores de diferentes métricas de centralidade para uma rede específica, conforme ilustrado na Figura 2. Cada coluna representa um nó individual da rede, numerados de 0 a 7, e cada linha descreve uma métrica específica de centralidade: Centralidade de Excentricidade (CE), Centralidade de Grau (CG), Centralidade *K-shell* (CK), Centralidade de Proximidade (CP), Centralidade de Harmonia (CH), Centralidade de Intermediação (CB) e Centralidade Autovetor (CA). Observa-se uma variabilidade entre os nós nas medidas de Centralidade de Excentricidade (CE), com nós como o 1, 5, 6 e 7 exibindo valores inferiores,

indicando uma posição mais periférica na rede. O nó 3 destaca-se com valores altos tanto na Centralidade de Grau (CG) quanto na Centralidade de Intermediação (CB), sublinhando sua influência central como ponto de conexão e intermediário crucial nas interações da rede.

Tabela 2 – Valores de diferentes métricas de centralidade calculados para a rede ilustrada na Figura 2, os valores foram calculados utilizando a biblioteca Networkx

	nó 0	nó 1	nó 2	nó 3	nó 4	nó 5	nó 6	nó 7
CE	0.5	0.333	0.5	0.5	0.5	0.333	0.333	0.333
CG	0.571	0.143	0.429	0.714	0.286	0.714	0.286	0.286
CK	2	1	2	2	2	2	2	2
CP	0.7	0.438	0.636	0.778	0.583	0.7	0.538	0.538
CH	5.5	3.5	5.0	6.0	5.833	4.5	4.333	4.333
CB	0.333	0.0	0.032	0.294	0.032	0.214	0.0	0.0
CA	0.353	0.101	0.385	0.511	0.239	0.486	0.285	0.285

Fonte: (Hagberg *et al.*, 2008).

2.2.3 Métricas da literatura

Na seção anterior, foram discutidas diversas métricas tradicionais empregadas na análise de redes e grafos. Embora essas métricas sejam fundamentais e amplamente reconhecidas no campo da análise de redes, é importante reconhecer que elas apresentam limitações. As suas principais dificuldades estão em ao considerar ponderação nas arestas metrificar somente com viés de estrutura da rede e pouco se leva em conta uma real mistura entre esses dois valores.

(Rêgo e Santos, 2019) criaram uma métrica para estudar co-autoria em artigos acadêmicos levando em conta a força entre as ligações dos vizinhos de cada autor e também como essa ligação é importante perante outras ligações entre vizinhos. A métrica de Utilidade (UT) é definida como:

$$UT(v) = \sum_{\mu \in \eta(v)} \left(\frac{w_{v,\mu}}{k_v^w} + \frac{w_{v,\mu}}{k_\mu^w} + \frac{w_{v,\mu}^2}{k_v^w k_\mu^w} \right) \quad (2.14)$$

O primeiro termo quantifica o quão importante é a ligação entre o sítio v com μ em relação às ligações de v , já o segundo é em relação às ligações de μ e o terceiro termo é um termo misto, capturando a sinergia da conexão entre os sítios. Para redes sem ponderação $w_{v,\mu} = 1$, $k_v^w = k_v$ (Jackson e Wolinsky, 1996). (Qi *et al.*, 2012) apresentam a chamada Centralidade Laplaciana (CL) na qual tentam mensurar a importância de um nó pela falta que ele faz se o retirarmos da rede (essa premissa vai ser muito presente na literatura atual). Porém como os

autores mostram essa métrica não é totalmente local, carregando informação de vizinhos de vizinhos, contudo ela não é definida bem em redes direcionadas. Seja W a Matriz de Adjacência ponderada e seja a matriz X uma matriz diagonal com cada elemento $X_{i,i} = \sum_{j \in \eta(v_i)} w_{v_i, v_j}$, pode-se definir uma matriz chamada de Laplaciana (Newman, 2018) L^a como $L_{i,j}^a = X_{i,j} - W_{i,j}$. Sejam λ_i os autovalores de L^a . É possível definir a Energia do grafo G como:

$$E(G) = \sum_i \lambda_i^2 = \sum_i X_{i,i}^2 + 2 \times \sum_{i < j} w_{v_i, v_j}^2, \quad (2.15)$$

Seja G_v o grafo G com a ausência do nó v , assim é possível definir a CL como:

$$CL(v) = \frac{E(G_v) - E(G)}{E(G)}. \quad (2.16)$$

Por fim (Andrade *et al.*, 2021) estabelecem uma centralidade baseada na lei da gravitação universal de Newton, chamada de Centralidade de Gravidade com abordagem individualista (CGR), ela leva em consideração as distâncias entre os nós e a ponderação nos nós da rede. A intuição dessa métrica está atrelada ao fato de que pessoas com características relevantes como conhecimento, idade, etnias, posição social, dentre outras os tornam mais ou menos influentes e atrativos na sua rede de contatos. Assim CGR é calculada como:

$$CGR(v, s) = \sum_{\mu \neq v} \frac{\Theta_v \Theta_\mu}{(d_{v,\mu}^w)^s}. \quad (2.17)$$

Na qual Θ_v é uma propriedade do nó v e s um parâmetro de controle que para a lei da gravitação $s = 2$.

2.3 Modelos de formação de redes

Na seção anterior, foram abordadas as métricas utilizadas para quantificar a estrutura de redes. Então, dado um grafo G é possível extrair suas propriedades. Contudo, surge a questão de saber se é possível criar uma rede que contenha determinadas propriedades quando essas são solicitadas. Portanto, nesta seção, serão discutidos os principais modelos empregados na formação de redes artificiais. Dois modelos de formação de redes amplamente utilizados na teoria das redes complexas são a Rede de Erdős-Rényi (ER) e o Modelo de Configuração. A Rede de Erdős-Rényi é uma abordagem clássica que gera redes aleatórias, onde a probabilidade de existência de uma aresta entre dois nós é independente e pré-definida. Por outro lado, o Modelo de Configuração permite criar redes com nós possuindo graus específicos, sendo útil para estudar como diferentes distribuições de grau afetam as propriedades da rede, como redes

de escala-livre, e para a geração de redes sintéticas com características semelhantes às redes reais, o que é valioso em estudos empíricos e modelagem de sistemas complexos. Ambos os modelos desempenham papéis fundamentais na análise e modelagem de redes complexas em diversos domínios.

2.3.1 Rede de Erdős-Rényi

Um dos primeiros modelos que conseguiu ter sucesso na geração sintética de redes foi o de Erdős-Rényi que iniciou os estudos de grafos (ou redes) aleatórias (Erdős e Rényi, 1959). O modelo é considerado estocástico com tempo discreto partindo do pressuposto que as conexões entre os nós é algo aleatório e conseguimos retirar propriedades importantes da rede. Apesar de ser um modelo simples e do fato de que poucas de suas propriedades aparecem em redes reais, é interessante estudar esse modelo por motivos históricos e também podemos aprender muito com ele.

Seja $G(\mathcal{N}, p)$ um grafo não direcionado que segue o modelo de Erdős-Rényi. Nesse modelo temos fixados o conjunto de vértices \mathcal{N} e temos uma probabilidade p para que cada ligação possível seja criada. Podemos construir a rede da seguinte forma:

1. Crie N vértices;
2. Selecione um vértice v e outro vértice μ 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 possíveis ligações.

Dada a descrição do modelo podemos estudar como suas métricas se comportam. O primeiro parâmetro a ser estudado é o grau. Dado um nó qualquer da rede, vamos calcular a probabilidade de ele possuir grau k . Observe que a partir de um determinado nó, existem $N - 1$ possíveis ligações e como cada ligação é formada de forma independente com probabilidade p , a probabilidade $P(k)$ de que um nó qualquer tenha grau k é:

$$P(k) = \binom{N-1}{k} p^k (1-p)^{N-1-k}. \quad (2.18)$$

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

$$\langle k \rangle = \sum_{k=1}^{N-1} P(k)k = \sum_{k=1}^{N-1} k \binom{N-1}{k} p^k (1-p)^{N-1-k} = (N-1)p. \quad (2.19)$$

No caso em que N é grande e p é pequeno, é obtido o resultado que a distribuição tende a uma de

Poisson com parâmetro $\langle k \rangle$. 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.

Outra métrica a ser avaliada é o agrupamento C da rede. Podemos ver pela definição que demos na seção anterior que o agrupamento é a probabilidade de dois nós estarem ligados dado que estão ligados ao mesmo nó. No caso da rede Binomial, como a probabilidade das arestas são independentes, isso é igual a probabilidade p :

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

2.3.2 *Modelo de Configuração*

Diferentemente do Modelo de Erdös-Rényi, no Modelo de Configuração (MC) (Newman, 2018), seja um grafo com uma quantidade conhecida de N nós e também informações sobre o grau de cada nó, representado por k_v . O MC possui uma restrição de que a soma de todos os graus deve ser um número par, condição que é decorrência da relação observada na Equação 2.3. O modelo, Algoritmo 1, é aplicado da seguinte maneira:

1. Criam-se os N vértices da rede;
2. Cada vértice v recebe $\kappa_v = k_v$ ligações a serem criadas;
3. Os vértices são ordenados de forma decrescente de grau da sequência $\{\kappa_i\}$ para facilitar a convergência do modelo;
4. Para cada nó v_i , é realizado um processo aleatório de seleção de outro nó v_j na rede, desde que v_j ainda possua capacidade para formar novas ligações, denotada por κ_j . Nesse cenário, estabelecemos uma conexão entre os nós v_i e v_j , reduzindo em 1 unidade os valores de κ_i e κ_j .
5. É repetido 3 e 4 até que todas as ligações na rede sejam colocadas.

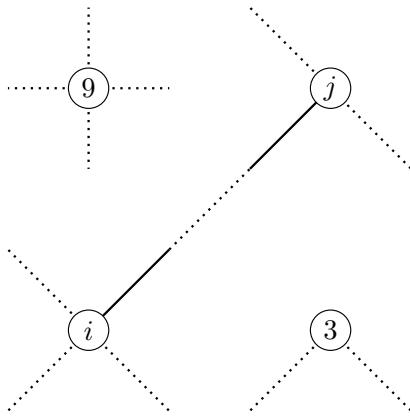
Algoritmo 1: Modelo de Configuração

```

1 for  $v \in \mathcal{N}$  do
2   |  $\kappa_v \leftarrow k_v;$ 
3 end
4  $\mathcal{L} \leftarrow \emptyset$ 
5  $\text{ordem} \leftarrow \text{ordenar}(\{v\}, \{k_v\})$  // Ordena os nós por número de meias arestas,
  em ordem decrescente.
6 for  $v \in \text{ordem}$  do
7   | while  $\kappa_v > 0$  do
8     |   |  $\mu \leftarrow \mathcal{N} \setminus \{v\}$  // Escolhido ao acaso
9     |   | if  $[(v, \mu) \notin \mathcal{L}] \text{ and } [\kappa_\mu \neq 0]$  then
10    |   |   |  $\mathcal{L} \leftarrow \mathcal{L} \cup \{(v, \mu)\};$ 
11    |   |   |  $\kappa_v \leftarrow \kappa_v - 1;$ 
12    |   |   |  $\kappa_\mu \leftarrow \kappa_\mu - 1;$ 
13    |   | end
14  | end
15 end

```

Figura 4 – Ilustração do Modelo de Configuração, é escolhido dois nós i e j aleatoriamente e são conectados



Fonte: Elaborado pelo autor.

2.4 Modelos de infecção ou difusão

Propagação de doenças (Arruda *et al.*, 2018; Wang *et al.*, 2019; Boguñá e Pastor-Satorras, 2002; Nowzari *et al.*, 2016), vírus de internet (Zhang *et al.*, 2016; Gan *et al.*, 2014; Yang *et al.*, 2013) e rumores (Chierichetti *et al.*, 2011; Moreno *et al.*, 2004; Yang *et al.*, 2010) 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 são as pessoas que difundem ou pela forma que a rede é construída é essencial para entender a difusão. Além disso, usar de modelagem matemática é essencial para boa parte desses problemas, pois apesar de termos dados sobre propagação

de rumores 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 no contexto de saúde pública. Em vista disso, nesta seção serão discutidos sobre os modelos de infecção clássicos da literatura, que apesar de simples são uma boa aproximação para a epidemiologia matemática.

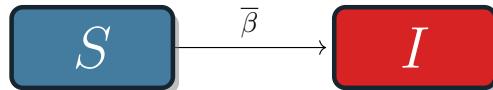
Na literatura existe a abordagem utilizando as equações diferenciais, no entanto em modelos tradicionais é assumida uma mistura homogênea, o que significa que um indivíduo infectado pode infectar qualquer outro indivíduo; outrossim esse tipo de modelagem, mesmo não considerando homogeneidade, é determinística. Para prever com precisão a dinâmica de uma epidemia, é necessário de considerar o papel preciso de cada indivíduo e como as suas relações contribuem para a proliferação como um todo. É nesse sentido que é proposto um modelo com uma abordagem utilizando redes complexas, pois assim haverá o entendimento de como acontece o espalhamento de uma doença e quais os principais agentes.

Na modelagem, cada indivíduo pode se encontrar em um estado ou comportamento da infecção, esses estados podem refletir diferentes níveis de exposição ao agente patogênico, status de vacinação, ou qualquer outro fator que influencie a transmissão da doença. A interação entre os compartimentos é modelada através de equações diferenciais ou em redes. No que tange redes, para passar de um estado para outro existem taxas que podem ser de tempo contínuo ou discreto. Em caso de tempo discreto, após o incremento de um período (por exemplo, um dia) cada nó tem uma chance aleatória de alterar seu estágio com uma probabilidade que considera a taxa e, se necessário, os vizinhos dele. Enquanto no tempo contínuo, cada intervalo de tempo subsequente segue uma distribuição exponencial da soma de todas as taxas da rede, resultando na atualização de um único nó escolhido de forma aleatória proporcionalmente a sua respectiva taxa para o comportamento subsequente (Gillespie, 1976).

2.4.1 *Modelo SI*

O modelo SI (Newman, 2018) descreve a propagação de uma doença em uma população, onde os indivíduos podem estar suscetíveis ou infectados, mas não se recuperam ou adquirem imunidade. A Figura 5 ilustra um esquema em diagrama de blocos, mostrando a única transição possível neste modelo.

Figura 5 – Esquematização do Modelo Suscetível-Infectado de contágio



Fonte: Elaborado pelo autor.

O modelo SI é um dos modelos mais simples de propagação de doenças, útil para entender a dinâmica básica de infecção. Ele assume uma população homogênea onde todos os indivíduos têm a mesma probabilidade de entrar em contato uns com os outros. Em cenários mais complexos, modelos mais sofisticados (como SIR ou SEIR) podem ser necessários para capturar características adicionais da propagação da doença, como recuperação ou períodos de incubação. Essa abordagem oferece uma base sólida para a compreensão inicial da dinâmica das doenças infecciosas e serve como um ponto de partida para modelos mais elaborados que incorporam mais detalhes sobre a transmissão e a progressão das doenças.

No modelo SI, a população é dividida em dois compartimentos: $S(t)$, que representa a fração de indivíduos suscetíveis (aqueles que podem ser infectados) no tempo t , e $I(t)$, que representa o número de indivíduos infectados no mesmo instante. A transmissão da doença ocorre quando indivíduos suscetíveis entram em contato com indivíduos infectados e a taxa de transmissão por indivíduo é β . Isto significa que, em média, cada indivíduo tem β contatos com outros indivíduos por unidade de tempo. A taxa de novas infecções depende diretamente da probabilidade de um encontro entre um suscetível e um infectado.

Dessa forma, a taxa média de novos casos de infecção por unidade de tempo é dada por $\bar{\beta}SI$. Isto porque cada infectado tem $\bar{\beta}$ contatos por unidade de tempo e a fração desses contatos que são com indivíduos suscetíveis é $S(t)$. Para modelar a dinâmica da infecção, precisamos de uma equação diferencial que descreva a taxa de variação de $I(t)$. A equação diferencial resultante é:

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

$$\frac{dI}{dt} = +\bar{\beta} \cdot S(t)I(t). \quad (2.22)$$

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:

$$\frac{dS}{dt} + \frac{dI}{dt} = 0 \rightarrow S + I = 1. \quad (2.23)$$

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. Substituindo (2.23) em (2.22) temos:

$$\frac{dI}{dt} = \bar{\beta}(1 - I(t))I(t) \quad (2.24)$$

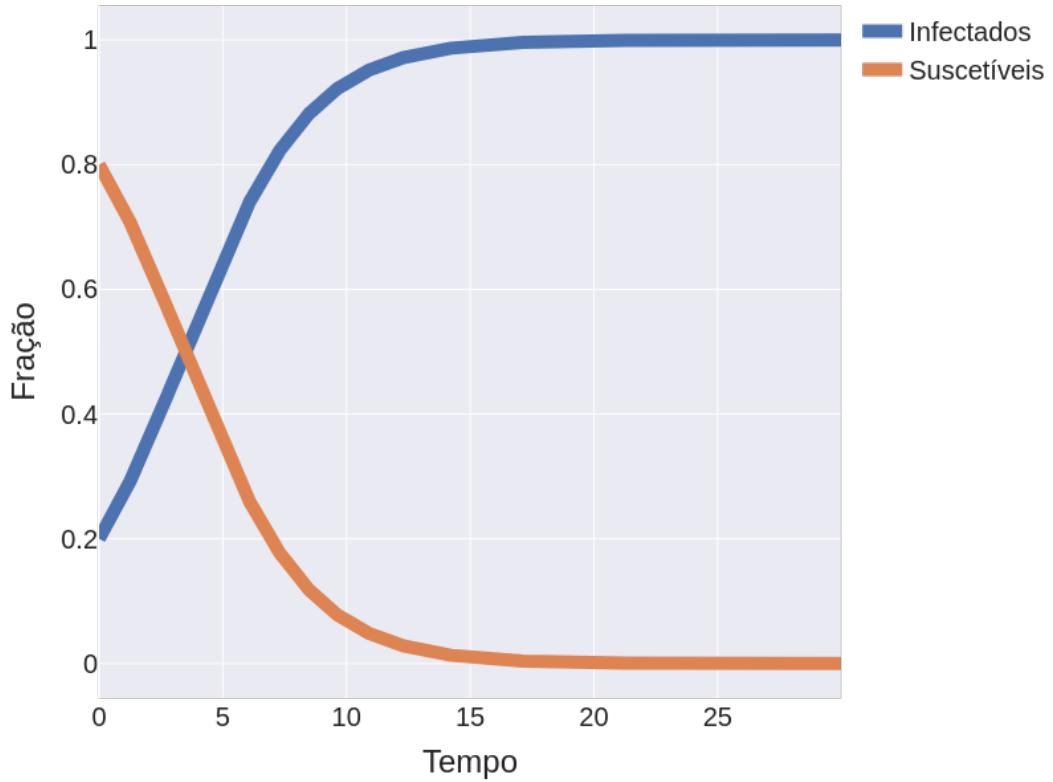
$$\frac{1}{I(t)(1 - I(t))} \frac{dI}{dt} = \bar{\beta} \quad (2.25)$$

Essa equação é resolvida a partir de integração dos dois lados obtendo:

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

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 $I(0)$ e, quando $t \rightarrow \infty$, $I(t)$ converge para 1 pois com o passar do tempo todos os nós se tornarão infectados como é mostrado na Figura 6.

Figura 6 – Modelo SI com $I(0) = 0.2$ e $\bar{\beta} = 0.4$. No início, há uma fração de infectados, e com o tempo, todos os indivíduos se tornarão infectados.



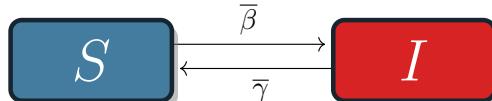
Fonte: Elaborado pelo autor.

Este modelo simplificado proporciona uma visão introdutória das dinâmicas de crescimento exponencial das infecções, ele se comporta bem para a modelagem de HIV (Júnior *et al.*, 2021) pois ou o indivíduo ao contrair a doença se torna imune ou porque após infectado

jamais retorna para o compartimento de suscetível. Apesar disso ainda existem várias outras doenças que não são modeladas dessa maneira.

2.4.2 *Modelo SIS*

Figura 7 – Modelo SIS com $I(0) = 0.2$, $\bar{\gamma} = 0.3$ e $\bar{\beta} = 0.1$. Caso $\bar{\beta} < \bar{\gamma}$, com a taxa de recuperação maior que a taxa de infecção, para $t \rightarrow \infty$, o número de infectados se aproxima de zero e o de suscetíveis se aproxima de um, indicando a erradicação da doença.



Fonte: Elaborado pelo autor.

O modelo SIS, esquematizado na Figura 7, é o modelo matemático que descreve a dinâmica de uma população dividida em dois compartimentos suscetíveis (S) e infectados (I) na qual é possível um indivíduo passar de infectado e voltar a ser suscetível, criando um ciclo contínuo de infecção e recuperação na população. São exemplos de doenças que podem ser modeladas a meningite, doenças venéreas e tuberculose (Júnior *et al.*, 2021). O modelo segue com uma lógica idêntica ao SI adicionando um termo $\bar{\gamma}$ que representa a taxa de recuperação do indivíduo, obtendo o seguinte conjunto de equações:

$$\frac{dS}{dt} = -\bar{\beta} \cdot S \cdot I + \bar{\gamma} \cdot I, \quad (2.27)$$

$$\frac{dI}{dt} = +\bar{\beta} \cdot S \cdot I - \bar{\gamma} \cdot I. \quad (2.28)$$

A partir desse conjunto de equações ainda é possível encontrar uma solução analítica, obtendo as seguintes expressões para S e I :

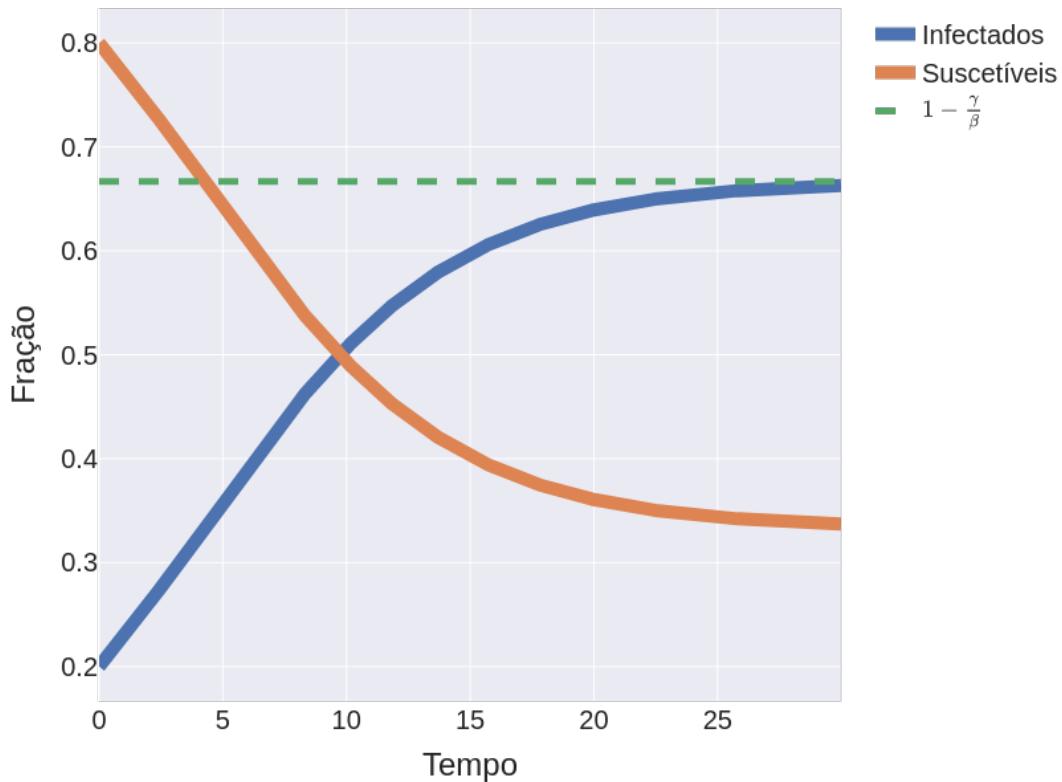
$$I(t) = \left(1 - \frac{\bar{\gamma}}{\bar{\beta}}\right) \frac{Ce^{(\bar{\beta}-\bar{\gamma})t}}{1 + Ce^{(\bar{\beta}-\bar{\gamma})t}}, \quad (2.29)$$

$$S(t) = 1 - I(t). \quad (2.30)$$

com $C = I(0)/(1 - I(0) - \bar{\gamma}/\bar{\beta})$. Assim a Equação 2.30 têm três tipos de comportamento dependendo de $\bar{\beta}$ e de $\bar{\gamma}$. Se $\bar{\beta} = \bar{\gamma}$ significa que a taxa de indivíduos que se tornam infectados é equivalente à taxa de recuperação que faz com que os infectados voltem a ser suscetíveis. Transforma as duas equações em $I = \frac{I(0)}{I(0)+\bar{\beta}t}$ e $S(t) = \frac{\bar{\beta}t}{1-S(0)+\bar{\beta}t}$ que mostra que a fração de infectados decai com $1/t$ e quando tempo tende a infinito o número de infectados tende para zero

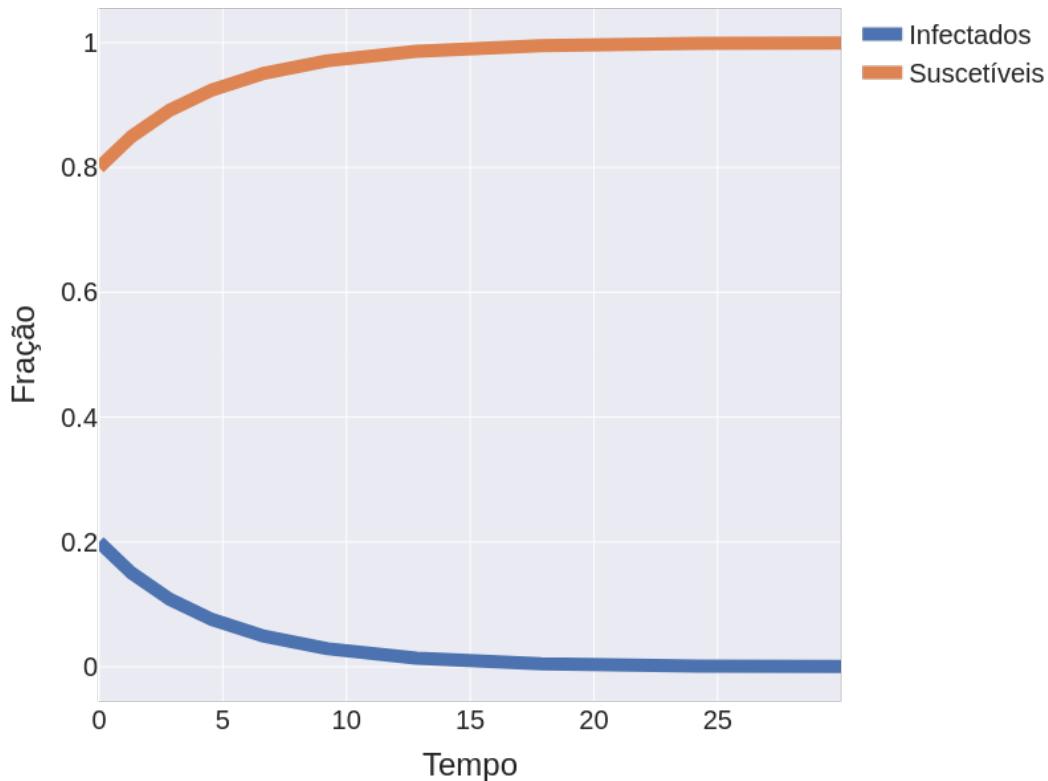
enquanto os suscetíveis converge para o tamanho da amostra. Se $\bar{\beta} > \bar{\gamma}$, significa que a taxa de infecção é maior que a de recuperação, e para $t \rightarrow \infty$ converge para um equilíbrio estável na qual $I(\infty) = 1 - \frac{\bar{\gamma}}{\bar{\beta}}$ e $S(\infty) = \frac{\bar{\gamma}}{\bar{\beta}}$, ou seja mesmo no infinito ainda existirá a doença no modelo (veja Fig. 8). Se $\bar{\beta} < \bar{\gamma}$, ou seja, a taxa de recuperação é maior que a taxa de infecção, e para $t \rightarrow \infty$ o número de infectados tende a 0 e o de suscetíveis tende a 1, ou seja, livre da doença (veja Fig. 9).

Figura 8 – Modelo SIS com $I(0) = 0.2$, $\bar{\beta} = 0.3$ e $\bar{\gamma} = 0.1$. Há convergência para equilíbrio estável com $\bar{\beta} > \bar{\gamma}$, para $t \rightarrow \infty$, as frações de infectados e suscetíveis estabilizam em $1 - \frac{\bar{\gamma}}{\bar{\beta}}$ e $\frac{\bar{\gamma}}{\bar{\beta}}$ respectivamente.



Fonte: Elaborado pelo autor.

Figura 9 – Modelo SIS com $I(0) = 0.2$, $\bar{\gamma} = 0.3$ e $\bar{\beta} = 0.1$. Caso $\bar{\beta} < \bar{\gamma}$, com a taxa de recuperação maior que a taxa de infecção, para $t \rightarrow \infty$, o número de infectados se aproxima de zero e o de suscetíveis se aproxima de um, indicando a erradicação da doença.



Fonte: Elaborado pelo autor.

2.4.3 *Modelo SIR*

Por fim o último dos modelos básicos para o estudo de infecções é inspirado por doenças que após a infecção o indivíduo deixa de ser infectado, contudo, não pode ser infectado novamente, isso pode ter a interpretação de que o indivíduo ganhou imunidade ou morreu. Esse modelo é o SIR (suscetível, infectado e removido) na qual adiciona um novo compartimento R, conforme esquematização da Figura 10. Um exemplo de doença que pode ser modelada é a catapora no qual um indivíduo após infectado fica imune à enfermidade.

Figura 10 – Esquematização do Modelo Susceptível-Infectado-Removido de contágio



Fonte: Elaborado pelo autor.

O modelo SIR pode ser descrito pelas equações

$$\frac{dS}{dt} = -\bar{\beta} \cdot S \cdot I \quad (2.31)$$

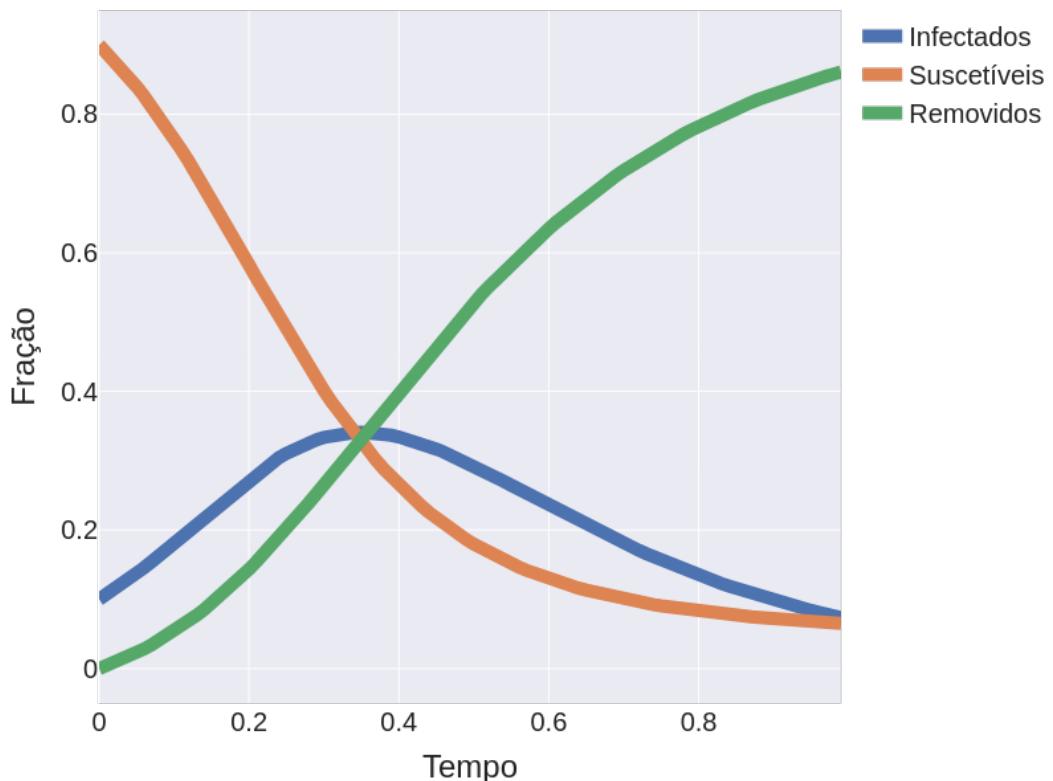
$$\frac{dI}{dt} = \bar{\beta} \cdot S \cdot I - \bar{\omega} \cdot I \quad (2.32)$$

$$\frac{dR}{dt} = \bar{\omega} \cdot I \quad (2.33)$$

$$S + I + R = 1, \quad (2.34)$$

com $I(0) = I_0$, $R(0) = 0$ e $S(0) = S_0$ e $\bar{\omega}$ a taxa de remoção. Diferente dos modelos anteriores, esse não tem solução analítica para a equação diferencial acoplada precisando de uma solução usando métodos computacionais, contudo ainda é possível extrair análises a partir dessas equações. A partir da equação $\frac{dI}{dt}$ e igualando ela a zero existe um ponto crítico de máximo em $S = \frac{\bar{\omega}}{\bar{\beta}}$ que não aparecia nos modelos anteriores, enquanto $S > \frac{\bar{\omega}}{\bar{\beta}}$ o modelo apresenta numa fase endêmica, na qual o número de infectados cresce até chegar no ponto crítico, quando $S < \frac{\bar{\omega}}{\bar{\beta}}$ o número de infectados decresce até chegar no final da endemia. Esse valor crítico é chamado de número ou taxa de reprodutibilidade basal R_0 que representa o número médio de pessoas que são infectadas por um único indivíduo.

Figura 11 – Modelo SIR com $I(0) = 0.2$, $S(0) = 0.8$, $R(0) = 0$, $\bar{\omega} = 0.4$ e $\bar{\beta} = 1.2$. É mostrado existe um ponto de máximo de $I(t)$, esse gráfico é bastante sensível aos parâmetros e seu comportamento pode mudar bastante com a alteração deles



Fonte: Elaborado pelo autor.

Reescalonando o tempo no modelo SIR como $\tau = \bar{\omega}t$, temos $\frac{d}{dt} = \bar{\omega}\frac{d}{d\tau}$. Substituindo essa relação nas equações diferenciais originais, obtendo $\frac{dS}{d\tau} = -\frac{\bar{\beta}}{\bar{\omega}}SI$, $\frac{dI}{d\tau} = \frac{\bar{\beta}}{\bar{\omega}}SI - I$ e $\frac{dR}{d\tau} = I$. É definido o parâmetro reescalonado $\bar{\beta}^* = \frac{\bar{\beta}}{\bar{\omega}}$, resultando nas equações $\frac{dS}{d\tau} = -\bar{\beta}^*SI$, $\frac{dI}{d\tau} = \bar{\beta}^*SI - I$ e $\frac{dR}{d\tau} = I$. Assim, o parâmetro $\bar{\omega}$ é normalizado a 1, e a dinâmica do sistema é descrita exclusivamente por $\bar{\beta}^*$ no tempo reescalonado. Isso simplifica a análise ao reduzir o número de parâmetros independentes e permitindo uma interpretação mais clara da relação entre transmissão e recuperação.

2.4.4 Modelos em Redes

Toda a modelagem descrita anteriormente foi apresentada no contexto de equações diferenciais, porém ela tem suas limitações como discutido anteriormente, contudo elucidam bem sobre como é feita esse tipo de modelagem. No contexto de redes complexas existem duas formas de construir a modelagem (Newman, 2018): de tempo contínuo ou de tempo discreto. Por exemplo, num modelo SI cada nó vai receber um rótulo descrevendo em qual compartimento ele está. No caso de tempo discreto um nó v no compartimento S tem uma probabilidade $p = (1 - \beta)^{k_v^I}$ de não ser contaminado, na qual β é diferente do que aparece nas equações diferenciais, enquanto no modelo de redes ele é uma probabilidade associada ao indivíduo, na modelagem por equações diferenciais é algo médio para todo o sistema, isso se aplica também ao γ e ω e k_v^I é o número de vizinhos de v que estão no estágio I . Portanto, a probabilidade de ser contaminado por pelo menos um nó é de

$$p_v = 1 - (1 - \beta)^{k_v^I}. \quad (2.35)$$

Uma vez calculada a probabilidade de contágio, pode-se decidir se o vértice passará ao estado I , em cujo caso muda-se o próximo estado do vértice para I . Ou seja, depois de uma interação no modelo na qual passa uma unidade de tempo (depende da unidade de β) os nós suscetíveis têm probabilidade p_i de serem contaminados, por isso é chamado de tempo discreto. No caso do modelo SIR ou SIS, existe uma probabilidade γ ou ω dele sair do estado infectado. Contudo, essa probabilidade não depende dos seus vizinhos, então cada nó que esteja infectado tem uma probabilidade γ ou ω de depois de uma unidade de tempo de mudar de estágio. O Algoritmo 2 mostra um pseudo-código de como funciona para o caso discreto o modelo SIS.

Para tempo contínuo no modelo SI, tradicionalmente, cada nó tem um tempo $T_v^{s \rightarrow i} \sim \text{Exp}(k_v^I \beta)$ (distribuição exponencial com parâmetro $k_v^I \beta$) para ser contaminado, com

Algoritmo 2: Algoritmo caso discreto do modelo SIS

Input: $G(\mathcal{N}, \mathcal{L})$: rede.

Input: $random()$: função que retorna um número ao acaso em $(0, 1)$.

```

1 for  $v_i \in \mathcal{N}$  do
2   | estado $[i] \leftarrow$  Inicialização de estados; // S ou I
3 end
4  $t \leftarrow 0$ 
5 while  $t \leq tempo$  do
6   | for  $v_i \in \mathcal{N}$  do
7     |   | if estado $[i] = S$  then
8       |     | num_inf  $\leftarrow 0$ 
9       |     | for  $v_j \in \eta(v_i)$  do
10      |       |       | if estado $[j] = I$  then
11        |         |         | num_inf  $\leftarrow num\_inf + 1$ 
12       |       |       | end
13      |     |   | end
14      |     |   |  $p \leftarrow 1 - (1 - \beta)^{num\_inf}$ 
15      |     |   | if random()  $< p$  then
16        |       |   | estado $[i] \leftarrow I$ 
17       |     |   | end
18     |   | else
19       |     |   | if random()  $< \gamma$  then
20         |       |   | estado $[i] \leftarrow S$ 
21       |     |   | end
22     |   | end
23   | end
24   |  $t \leftarrow t + 1$ 
25 end
```

$k_v^I = |\{v_j \in \eta(v) : \text{estado}[j] = I\}|$ (contudo há estudos com tempos não exponenciais (Feng *et al.*, 2007)). Portanto, o tempo Δt para a próxima atualização na rede segue uma distribuição exponencial com parâmetro igual a soma dos parâmetros de cada variável $\Delta t \sim \text{Exp}(\sum_{v \in S} k_v^I \beta)$, pois o mínimo entre várias variáveis exponenciais independentes segue uma distribuição exponencial com a soma dos parâmetros.

A escolha do nó a ser atualizado é feita a partir do cálculo de $M = \sum_{v \in S} k_v \beta$. Seja $S = \{\mu_1, \mu_2, \dots, \mu_{|S|}\}$ uma ordem do conjunto de nós suscetíveis. Assim é gerado um valor r aleatório entre 0 e 1, o nó μ_j a ser atualizado é primeiro vértice tal que $\sum_{i=1}^j k_{\mu_i}^I \beta \geq r \cdot M$. Para modelos como SIS e SIR segue a mesma ideia, sem considerar os vizinhos, o tempo para mudar de estágio é $T_v^{i \rightarrow s} \sim \text{Exp}(\gamma)$ ou $T_v^{i \rightarrow r} \sim \text{Exp}(\omega)$, agora M é calculado com a soma de $\sum_{v \in S} k_v^I \beta + \sum_{v \in I} \gamma$ ou de $\sum_{v \in S} k_v^I \beta + \sum_{v \in I} \omega$ e assim o tempo que passa é $\Delta t \sim \text{Exp}(M)$ e o nó escolhido é tal que a soma de todas as taxas até ele seja maior que $r \cdot M$ (Gillespie, 1976). O

Algoritmo 3 mostra um pseudo-código de como funciona para o caso contínuo o modelo SIS.

Algoritmo 3: Algoritmo caso contínuo do modelo SIS

Input : $G(\mathcal{N}, \mathcal{L})$: rede.
Input : $random()$: função que retorna um número ao acaso em $(0, 1)$.
Input : $exp(\lambda)$: função que retorna um número aleatório com distribuição exponencial de parâmetro λ .

```

1 for  $v \in \mathcal{N}$  do
2   | // Inicialização de estados
3   |    $S \leftarrow S \cup \{v\}$  ou  $I \leftarrow I \cup \{v\}$ 
4 end
5  $t \leftarrow 0$ 
6 while  $t \leq tempo$  do
7   |  $M \leftarrow 0$ 
8   | // Os nós são visitados numa ordem fixa
9   | for  $v_i \in \mathcal{N}$  do
10  |   | if  $v_i \in S$  then
11  |   |   |  $k_{v_i}^I \leftarrow |I \cap \eta(v_i)|$ 
12  |   |   |  $rate[i] \leftarrow k_{v_i}^I \cdot \beta$ 
13  |   | else
14  |   |   |  $rate[i] \leftarrow \gamma$ 
15  |   | end
16  |   |  $M \leftarrow M + rate[i]$ 
17  | end
18  |  $t \leftarrow t + exp(M)$ 
19  |  $\Delta \leftarrow random() \cdot M$ 
20  |  $m \leftarrow 0$ 
21  |  $i \leftarrow 0$ 
22  | while  $m \leq \Delta$  do
23  |   |  $i \leftarrow i + 1$ 
24  |   |  $m \leftarrow m + rate[i]$ 
25  | end
26  | if  $v_i \in S$  then
27  |   |  $I \leftarrow I \cup \{v_i\}$ 
28  |   |  $S \leftarrow S \setminus \{v_i\}$ 
29  | else
30  |   |  $S \leftarrow S \cup \{v_i\}$ 
31  |   |  $I \leftarrow I \setminus \{v_i\}$ 
32 end
33 end
```

2.4.5 Modelo SEIHARD

Os modelos de infecção mais tradicionais apresentados anteriormente têm se mostrado eficazes na simulação de epidemias e no entendimento geral da propagação de doenças em

uma população. No entanto, esses modelos tendem a tratar a população de forma homogênea, sem considerar as particularidades que podem influenciar a dinâmica da infecção em diferentes subgrupos. A pandemia de COVID-19 evidenciou as limitações desses modelos tradicionais ao destacar a importância de fatores como idade e comorbidades na propagação e nos desfechos da doença. Indivíduos mais velhos e aqueles com condições de saúde preexistentes, como diabetes ou doenças cardíacas, apresentaram riscos significativamente maiores de complicações graves e mortalidade. Esses fatores são determinantes para a dinâmica da infecção e, portanto, requerem uma modelagem que vá além dos pressupostos uniformes dos modelos tradicionais.

Além disso, a COVID-19 possui características peculiares na sua dinâmica de infecção que não são capturadas adequadamente pelos modelos tradicionais. Um exemplo é a fase assintomática da doença, durante a qual indivíduos infectados podem transmitir o vírus sem apresentarem sintomas. Esta fase é particularmente importante para a modelagem, pois a diferenciação entre indivíduos assintomáticos e sintomáticos pode influenciar drasticamente a previsão de quem está em maior risco de mortalidade. Além disso, as hospitalizações se tornaram um fator crítico na gestão da pandemia, não apenas pela sua alta frequência, mas também pela sobrecarga que impõem ao sistema de saúde. Contudo a literatura (Liu e Lou, 2022; Xiang *et al.*, 2021) apresenta vários tipos de modelos que foram utilizados para modelar a epidemia do Coronavírus e é importante entender qual é o melhor balanceamento de variáveis para não ficar muito complexo ou muito simplificado.

Neste contexto, o modelo apresentado por (Eikenberry *et al.*, 2020) oferece uma abordagem mais específica para a COVID-19 e será a base deste trabalho. O modelo SEIHARD é uma adaptação de *frameworks* SEIR para explorar o impacto do uso de máscaras pela população geral na dinâmica de transmissão da COVID-19, contudo isso não será abordado neste trabalho e o modelo proposto admite uma evolução temporal da taxa β que também não será abordada. Apesar de não considerar explicitamente a idade ou comorbidade como um fator, o modelo é projetado para capturar a complexidade adicional da pandemia, incluindo a presença de indivíduos assintomáticos e a influência da hospitalização.

O modelo é um sistema de equações diferenciais determinísticas não lineares que estratifica a população em diferentes compartimentos: suscetíveis (S), expostos (E), infecciosos sintomáticos (I), infecciosos assintomáticos (A), hospitalizados (H), recuperados (R) e mortos (D). A dinâmica da infecção no modelo é governada por taxas de transição entre os compartimentos, onde indivíduos suscetíveis podem se tornar expostos após contato com infecciosos. A fase

exposta representa o período de incubação do vírus, após o qual os indivíduos podem se tornar sintomáticos ou assintomáticos. Os sintomáticos têm uma probabilidade de serem hospitalizados, enquanto os assintomáticos se recuperam ou continuam a transmitir o vírus. A dinâmica é modelada pelo sistema de equações

$$\frac{dS}{dt} = -\beta(t)(I + \eta A) \frac{S}{N}, \quad (2.36)$$

$$\frac{dE}{dt} = \beta(t)(I + \eta A) \frac{S}{N} - \sigma E, \quad (2.37)$$

$$\frac{dI}{dt} = \alpha \sigma E - \phi I - \gamma_I I, \quad (2.38)$$

$$\frac{dA}{dt} = (1 - \alpha) \sigma E - \gamma_A A, \quad (2.39)$$

$$\frac{dH}{dt} = \phi I - \delta H - \gamma_H H, \quad (2.40)$$

$$\frac{dR}{dt} = \gamma_I I + \gamma_A A + \gamma_H H, \quad (2.41)$$

$$\frac{dD}{dt} = \delta H. \quad (2.42)$$

Na qual $\beta(t)$ é a taxa de contato infeccioso, que pode variar ao longo do tempo; η representa a infecciosidade relativa dos indivíduos assintomáticos em comparação aos sintomáticos; σ é a taxa de transição dos expostos para os infecciosos; α é a fração dos expostos que se tornam sintomáticos; ϕ é a taxa com que os sintomáticos são hospitalizados; γ_I , γ_A e γ_H são as taxas de recuperação para os indivíduos sintomáticos, assintomáticos e hospitalizados, respectivamente; e δ é a taxa de mortalidade entre os hospitalizados.

2.5 Fronteira de pareto

No âmbito epidemiológico, a avaliação de estratégias baseia-se em diversas métricas, como a redução da taxa de infecção, a rapidez na obtenção da imunidade coletiva, os custos financeiros envolvidos e os impactos socioeconômicos resultantes. Devido ao potencial conflito entre esses critérios, identificar a estratégia de vacinação ideal torna-se um desafio complexo e multidimensional no planejamento de saúde pública.

Para lidar com essa complexidade, aplicamos o conceito de Fronteira de Pareto (Branke *et al.*, 2008) proveniente da otimização multiobjetivo. Formalmente, em um problema com múltiplos objetivos a serem otimizados simultaneamente, uma solução é considerada ótima de Pareto se não existe outra solução viável que melhore pelo menos um dos objetivos sem piorar algum outro. A Fronteira de Pareto é o conjunto de todas as soluções ótimas de Pareto,

representando os melhores compromissos possíveis entre os diferentes critérios em conflito.

Matematicamente, seja S o conjunto de soluções viáveis em um problema de otimização multiobjetivo com funções objetivo f_1, f_2, \dots, f_k , que devem ser minimizadas. Uma solução $x^* \in S$ é dita ótima de Pareto se não existir outra solução $x \in S$ tal que $f_i(x) \leq f_i(x^*)$ para todo $i = 1, 2, \dots, k$, e $f_j(x) < f_j(x^*)$ para pelo menos um índice j . A Fronteira de Pareto é então definida como o conjunto de todas as soluções ótimas de Pareto em S , permitindo identificar soluções onde a melhoria em um dos objetivos só é possível às custas do comprometimento de pelo menos outro. Ao analisar a Fronteira de Pareto, é possível selecionar estratégias que se alinham melhor às prioridades e restrições específicas do contexto em questão, seja dando preferência à máxima eficácia, ao menor custo ou a um equilíbrio entre diversos fatores.

3 METODOLOGIA

3.1 Coleta de dados

A utilização de Redes Complexas como ferramenta de modelagem para estudos sobre a COVID-19 é intuitiva, pois o espalhamento das epidemias ocorre através de redes de interações humanas. Para a implementação dessa abordagem, é essencial dispor de dados sobre redes de contágio que incluam informações relevantes para a análise, como idade, tempo de contato, pertencimento a grupos de risco, entre outros aspectos.

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 que tinha o intuito da coleta de dados para entender quais seriam os padrões de contatos da Europa (Mossong *et al.*, 2008). Apoiado nesse estudo, pesquisas posteriores se inspiraram no formato do POLYMOD (que podem ser encontrados em <https://socialcontactdata.org/>). Essa pesquisa foi feita baseada em um questionário de maio de 2005 até setembro de 2006 em 8 diferentes países da Europa a partir de ligações aleatórias ou entrevistas presenciais. Os participantes da pesquisa foram recrutados de forma a serem amplamente representativos de toda a população em termos de distribuição geográfica, idade e sexo, mas houve uma superamostragem para crianças e adolescentes devido ao seu papel significativo na disseminação de agentes infecciosos. Esta superamostragem foi justificada pelo alto número de contatos que eles mantêm e pela natureza desses contatos, que tendem a ocorrer dentro do próprio grupo etário e serem de longa duração.

Ao final do dia a pessoa teria que preencher o diário anotando sobre informações pessoais, ambiente (escola, trabalho, casa, ...) e as pessoas com as quais esteve fisicamente próximo no dia, a idade aproximada (ou exata) dessas pessoas, duração da conversa, frequência na qual têm contato, dentre outras informações. 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.

Apesar dos dados obtidos a partir do estudo POLYMOD, é crucial notar que este não constrói uma rede de conexões entre os participantes. Tal limitação é inerente à metodologia adotada pelo estudo, que se baseia em discagem e entrevistas aleatórias para coleta de informações. A escolha por esse método implica que as conexões entre os participantes não são mapeadas diretamente na forma de uma rede interconectada. A rede gerada pelos dados

seria aproximadamente formada por vários grafos estrelas desconectados em que os nós centrais da estrela são os participantes das entrevistas. Portanto, para ser possível formar uma rede e utilizar os dados anteriores é necessário um modelo que leve em consideração a distribuição de graus, idade e as conexões entre idades. Um modelo muito utilizado na literatura é o *Stochastic Block Model* (SBM) (Lee e Wilkinson, 2019) no qual cada nó é separado em um bloco e os nós pertencentes a cada bloco são conectados com probabilidade p_{B_v, B_μ} , na qual B_v é o bloco de v . A principal desvantagem dessa abordagem é que a rede gerada apresenta uma característica média não resgatando, por exemplo, a distribuição de graus dos dados. (Klise *et al.*, 2022) apresenta uma outra forma de construir a rede na qual separa a rede em camadas relacionadas aos locais onde os indivíduos visitam (casa, escola ou trabalho) entretanto para a base de dados do POLYMOD isso não é possível pois não se tem noção da totalidade das redes.

3.2 Modelo de formação de redes proposto

Para evitar esse problema (Manzo e Rijt, 2020) propõe um modelo para formação de uma rede baseado no modelo de configuração para um problema similar. Com base nas entrevistas realizadas em dois dias, calcula-se a média diária de contatos por pessoa (arredondada para cima), utilizando 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 esta é uma propriedade muito importante na suavização de epidemias (Block *et al.*, 2020). Nesse sentido, Manzo e van de Rijt propõem que antes de ser excluído um dado nó do Modelo de Configuração (Algoritmo 4), passamos por cada par de vizinhos dele e os conectamos entre si com uma probabilidade p . Ou seja:

1. Quando um dado nó v atinge o grau requerido pelo MC são selecionados todos os vizinhos $\eta(v)$;
2. São escolhidos dois nós aleatoriamente $\mu, \zeta \in \eta(v)$ e são conectados com probabilidade p ;
3. Se a ligação existir 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_μ e κ_ζ , se a ligação entre μ e ζ for adicionada na rede;
4. Repetimos 2 até que todas as ligações possíveis sejam testadas e o algoritmo termina para dar continuidade ao MC.

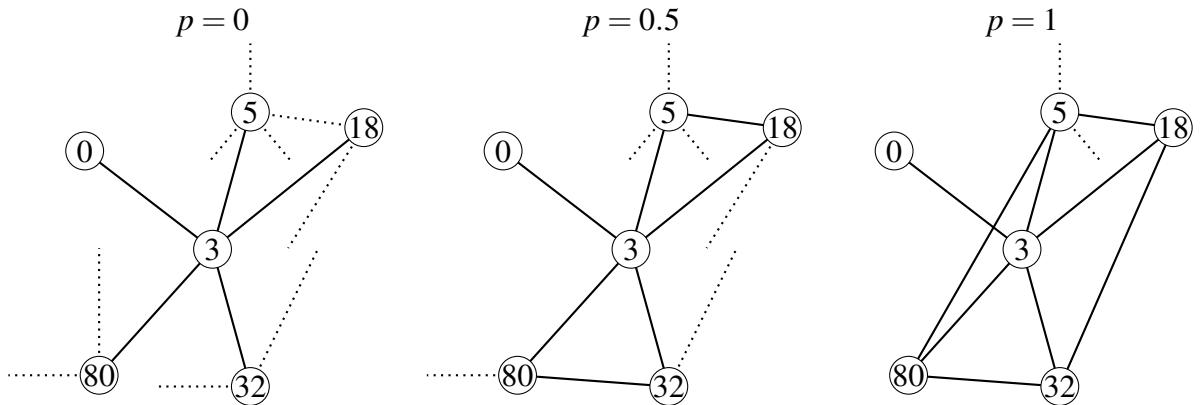
Algoritmo 4: Implementação de Manzo e van Rijt

Input: $G(\mathcal{N}, \mathcal{L})$: rede já construída
Input: $\mathcal{L}_{\text{test}}$: possíveis ligações já testadas // Não são testadas duas vezes.
Input: $\text{random}()$: função que retorna um número ao acaso em $(0, 1)$.

```

1 for  $v \in \mathcal{N}$  do
2   for  $\mu \in \eta(v)$  do
3     for  $\zeta \in \eta(v) \setminus \{\mu\}$  do
4       if ( $\kappa_\zeta \neq 0$ ) and ( $\kappa_\mu \neq 0$ ) then
5         if ( $\text{random}() \leq p$ ) then
6           if ( $\mu, \zeta \notin \mathcal{L} \cup \mathcal{L}_{\text{test}}$ ) then
7              $\mathcal{L} \leftarrow \mathcal{L} \cup \{(\mu, \zeta)\}$  // Essa matriz vem do MC
8              $\kappa_\mu \leftarrow \kappa_\mu - 1$ 
9              $\kappa_\zeta \leftarrow \kappa_\zeta - 1$ 
10            end
11          end
12         $\mathcal{L}_{\text{test}} \leftarrow \mathcal{L}_{\text{test}} \cup \{(\mu, \zeta)\}$ 
13      end
14    end
15  end
16 end
```

Figura 12 – Ilustração do Modelo de Manzo e van Rijt que adiciona uma probabilidade de agrupamento p



Fonte: Elaborado pelo autor

O modelo é mostrado na Figura 12, quando um dado nó v atinge o grau necessário para cada $\mu \in \eta(v)$ os nós são conectados entre si com uma probabilidade p se o grau de v não tiver atingido seu valor. Para que esse algoritmo se adapte a problemática utilizaremos os dados do POLYMOD. A partir deles é possível construir a distribuição $\{k_v\}_{v \in \mathcal{N}}$ de graus e $\{f_v\}_{v \in \mathcal{N}}$ de faixas etárias.

Nesse contexto será proposto uma nova atualização no MC, consideraremos o conjunto $\kappa_{v,f}$ de conexões de um dado nó v de faixa etária f_v com um nó de faixa etária

f , as quais serão incorporadas ao modelo de configuração, resultando em uma atualização considerando as conexões por faixa etária na qual será chamado de Modelo de Configuração Estratificado (MCE). Para a implementação do MCE, é fundamental dispor tanto da distribuição empírica de graus quanto de uma matriz M tal que $M_{f,a}$ registra a proporção média de conexões de um nó na faixa etária f que vão pra faixa etária a . Adicionalmente, a distribuição F_w das faixas etárias é calculada a partir de um conjunto de dados. O procedimento para a implementação do MCE é delineado nos passos seguintes:

1. São criados N nós na rede formando-se o conjunto \mathcal{N} ;
2. Atribui-se a cada nó v uma faixa etária f_v , seguindo uma distribuição empírica F_w ;
3. Com base nos dados coletados, são geradas k_v meias arestas para cada nó $v \in \mathcal{N}$, seguindo a distribuição de graus empírica da faixa etária f_v ;
4. Para cada nó v , o número de meias arestas que vão se conectar com cada faixa etária a , $\kappa_{v,f}$, é escolhido conforme a distribuição multinomial $\text{Multi}(k_v, M_{f_v,:})$;
5. Os nós são ordenados em ordem decrescente de grau, de acordo com a sequência $\{\kappa_v\}$, para facilitar a convergência do modelo;
6. Para cada nó v , uma faixa etária a que ainda tenha meias arestas $\kappa_{v,a}$ a serem conectadas. Um nó μ dessa faixa etária é escolhido aleatoriamente entre todos os nós, excluindo v , que não possuam conexões com v , e que ainda tenham meias arestas disponíveis κ_{μ,f_v} para estabelecer;
7. Para cada nó μ selecionado dessa maneira, estabelece-se uma conexão com v , e os contadores κ_{μ,f_v} e $\kappa_{v,a}$ são decrementados em 1;
8. O processo termina para o nó v quando $\kappa_{v,a} = 0$ para todas as faixas a ou não exista um nó v_j que possa ser selecionado como descrito no passo 6;
9. Logo após deve-se utilizar o Algoritmo de Manzo e van de Rijt adaptado (explicado no seguinte parágrafo) com parâmetro p para incrementar o agrupamento;
10. Os passos de 6 a 9 são repetidos para todos os nós da rede.

O Algoritmo 5 é o pseudocódigo do MCE. Assim como o modelo tradicional o MCE tem um agrupamento limitado e baixo, portanto será implementado uma nova versão do Modelo de Manzo e van de Rijt para o MCE. Contudo ele precisa fazer algumas alterações para poder se adaptar ao MCE que utiliza as faixas etárias. O Algoritmo 6 é um pseudo-código com o algoritmo adaptado utilizando as faixas etárias, ele é chamado logo após todas as ligações $\{\kappa_{v,a}\}$ serem colocadas ou quando todos os nós forem checados.

Algoritmo 5: Implementação do MCE

```

1  $\mathcal{L} \leftarrow \emptyset$ 
2 for  $v \in \mathcal{N}$  do
3    $f_v \leftarrow$  segundo a distribuição  $F_w$ 
4    $k_v \leftarrow$  segundo a distribuição empírica de graus da faixa  $f_v$ 
5    $\{\kappa_{v,:}\} \leftarrow$  segundo a distribuição multinomial Multi( $k_v, M_{f_v,:}$ )
6   lista[ $f_v$ ]  $\leftarrow$  lista[ $f_v$ ]  $\cup \{v\}$ 
7 end
8 ordem  $\leftarrow$  ordenar( $\{v\}, \{k_v\}$ ) // Ordena os nós por número de meias arestas,
  em ordem decrescente.
9 for  $v \in$  ordem do
10  for  $a < \text{tamanho}(F_w)$  do
11    embaralha(lista[f])
12    for  $\mu \in \text{lista}[a] \setminus \{v\}$  do
13      if  $[\kappa_{\mu,f_v} \neq 0] \text{ and } [(v, \mu) \notin \mathcal{L}]$  then
14         $\mathcal{L} \leftarrow \mathcal{L} \cup \{(v, \mu)\}$ 
15         $\kappa_{v,a} \leftarrow \kappa_{v,a} - 1$ 
16         $\kappa_{\mu,f_v} \leftarrow \kappa_{\mu,f_v} - 1$ 
17      end
18      if  $\kappa_{v,a} == 0$  then break ;
19    end
20    Manzo_Rijt( $v, p$ )
21  end
22 end

```

Algoritmo 6: Implementação de Manzo e van de Rijt - Adaptado

Input : v : nó sob consideração

Input : p : probabilidade de conexão

Input : $G(\mathcal{N}, \mathcal{L})$: rede já construída

Input : $\mathcal{L}_{\text{test}}$: possíveis ligações já testadas // Não são testadas duas vezes.

Input : $\text{random}()$: função que retorna um número ao acaso em $(0, 1)$.

```

1 for  $\mu \in \eta(v)$  do
2   if  $soma(\kappa_{\mu,:}) \neq 0$  then
3     lista_vizinhos[ $f_\mu$ ]  $\leftarrow$  lista_vizinhos[ $f_\mu$ ]  $\cup$  { $\mu$ }
4   end
5 end
6 for  $\mu \in \eta(v)$  do
7   for  $a < tamanho(F_w)$  do
8     embaralha(lista_vizinhos[ $f_\mu$ ])
9     for  $\zeta \in lista\_vizinhos[a] \setminus \{v\}$  do
10       if  $\kappa_{\mu,a} == 0$  then break ;
11       if  $\kappa_{\zeta,f_\mu} \neq 0$  then
12         if ( $\text{random}() \leq p$ ) then
13           if  $(\mu, \zeta) \notin \mathcal{L} \cup \mathcal{L}_{\text{test}}$  then
14              $\mathcal{L} \leftarrow \mathcal{L} \cup \{(\mu, \zeta)\}$  // Essa matriz vem do MC
15              $\kappa_{\mu,a} \leftarrow \kappa_{\mu,a} - 1$ 
16              $\kappa_{\zeta,f_\mu} \leftarrow \kappa_{\zeta,f_\mu} - 1$ 
17           end
18         end
19          $\mathcal{L}_{\text{test}} \leftarrow \mathcal{L}_{\text{test}} \cup \{(\mu, \zeta)\}$ 
20       end
21     end
22   end
23 end
```

3.3 Modelo de infecção proposto SEIHARDS

Baseando-se no modelo proposto por (Eikenberry *et al.*, 2020), serão feitas melhorias

que incluem a incorporação de faixas etárias e a possibilidade de reinfecção da doença. Além disso, considerando que o modelo será implementado em redes, serão introduzidos compartimentos nos quais não ocorrerá contaminação. Finalmente, o modelo usado não é determinista como o SEIHARD original, mas estocástico. O modelo SEIHARDS é composto pelos estágios: Suscetíveis (**S**), Expostos (**E**), Infectados Sintomáticos (**I**), Hospitalizados (**H**), Infectados Assintomáticos (**A**), Recuperados (**R**) e Mortos (**D**) que são modelados da seguinte forma:

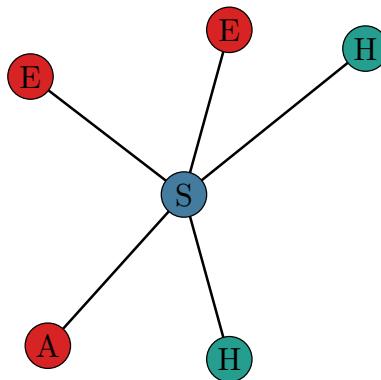
- **Suscetível (S)**: o nó neste estágio pode ser infectado pelos seus vizinhos que estão no estágio **A** ou no estágio **E**, essa taxa será ε_S vezes o número de contatos que estão no estágio **E** somado ε_A vezes o número de contatos no estágio **A** resultando na Equação 3.1, para o caso sem ponderação nas arestas. Para o caso com ponderação em arestas deve-se utilizar o peso multiplicando as taxas, resultando na Equação 3.2. Um exemplo do cálculo para o caso sem ponderação é ilustrado na Figura 13.

$$\Lambda_v = \frac{|\mu \in \eta(v) \cap E| \cdot \varepsilon_S + |\mu \in \eta(v) \cap A| \cdot \varepsilon_A}{\langle k \rangle} \quad (3.1)$$

$$\Lambda_v^w = \frac{\varepsilon_S \left(\sum_{\mu \in \eta(v) \cap E} w_{v,\mu} \right) + \varepsilon_A \left(\sum_{\mu \in \eta(v) \cap A} w_{v,\mu} \right)}{\langle k^w \rangle} \quad (3.2)$$

Figura 13 – Ilustração do cálculo da taxa para se tornar infectado

$$\Lambda = 2 \cdot \varepsilon_S + 1 \cdot \varepsilon_A = 2 \cdot 0.5 + 1 \cdot 0.41 = 1.41 \text{ dia}^{-1}$$



Fonte: Elaborado pelo autor.

- **Exposto (E)**: um nó é considerado exposto quando ele estava no estágio **S** e foi contaminado por alguém no estágio **E** ou **A**. Este estágio é para representar os indivíduos que estão na fase de incubação do vírus, isso faz com que eles possam contaminar os outros

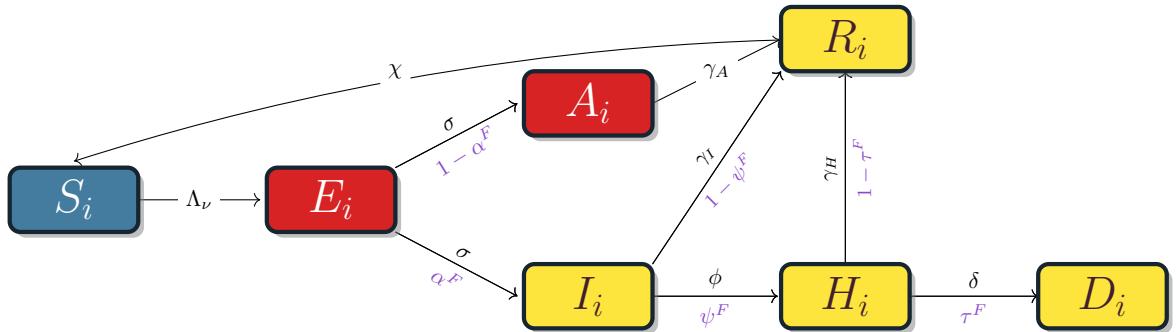
indivíduos por não saberem que têm a doença. Deste estágio, ele pode passar para o estágio **I** com uma probabilidade α^F , dependente da faixa etária f , ou para o estágio **A** com probabilidade $(1 - \alpha^F)$ mas com a mesma taxa σ .

- **Infectado Sintomático (I):** um nó neste estágio está infectado, porém ele não pode infectar outras pessoas. Isso é feito para modelar um indivíduo que está sentindo os sintomas e, por isso, foi isolado e não teria mais contato com outras pessoas. Uma pessoa neste estágio tem probabilidade ψ^F de se hospitalizar, neste caso o tempo até hospitalização segue uma taxa ϕ^F . Caso não seja necessária a hospitalização, a taxa do tempo de recuperação de sintomáticos é dada por γ_I .
- **Hospitalizado (H):** nele o nó continua sem a capacidade de contaminar outros nós. Nesse estágio ele tem uma probabilidade τ^F de ir para o comportamento Morto com taxa δ e uma $(1 - \tau^F)$ de se recuperar com taxa γ_H ;
- **Infectado Assintomático (A):** a diferença deste 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 seus contatos e também não tem risco de morte ou necessidade de ser hospitalizado. Um indivíduo neste estágio se recupera com uma taxa γ_A .
- **Recuperado (R):** caso o nó não morra ele será considerado recuperado na qual consegue voltar a ser suscetível novamente com uma taxa χ .
- **Morto (D):** Por fim, neste estado o nó perde todas as conexões e é nele que serão contabilizadas quantas pessoas morreram para utilizarmos como uma das métricas de avaliação das estratégias de vacinação.

Na Figura 14 está esquematizado como funciona a dinâmica de infecção, sem considerar a vacinação das pessoas. No modelo, é assumido que o tempo de transição entre os estados segue uma distribuição exponencial com o parâmetro descrito como um rótulo das arestas na Figura 14. Os valores de cada parâmetro estão indicados nas Tabelas 3 e 4. Estes valores foram obtidos de 2 formas: ou pela busca em artigos ou utilizando a base de dados do openDataSUS para Síndrome Respiratória Aguda Grave de 2020 (BRASIL. Ministério da Saúde,), um portal brasileiro no qual existem dados de internações COVID-19 e outras doenças gripais. O valor $1/\gamma_H$ ($1/\phi$), foi calculado a partir da média da diferença do tempo de uma pessoa ter sido internada para o tempo que ela saiu do hospital (ou morreu), já a probabilidade τ^F foi obtida calculando quantas pessoas na faixa etária F morreram dado que tinham sido

internadas. Ademais, pelo fato de que o openDataSUS é limitado aos dados de internações, tornou-se necessária a busca de mais dados em artigos científicos.

Figura 14 – Esquematização do Modelo SEIHARDS de contágio. Os blocos em vermelho significam quem pode contaminar outros, os blocos em azul escuro mostra o comportamento ainda não infectado e em amarelo mostra o estágio na qual o indivíduo não interage mais com a rede. Os parâmetros em preto são as taxas e em roxo são as probabilidades de entrar no estágio



Fonte: Elaborado pelo autor.

Tabela 3 – Taxas de transição utilizadas no modelo SEIHARDS

Parâmetro	Definição	Valor [dia ⁻¹]	Notas
ε_E	Taxa de contágio Expostos	0.500	(1)
ε_A	Taxa de contágio Assintomática	0.410	(2)
σ	Taxa E→A, E→I (incubação)	0.196	(1)
γ_A	Taxa de recuperação de Assintomáticos	0.143	(1)
γ_I	Taxa de recuperação de Sintomáticos	0.143	(1)
γ_H	Taxa de recuperação de Hospitalizados	0.083	(3)
ϕ	Taxa I→H	0.144	(3)
δ	Taxa de Mortalidade de Hospitalizados	0.073	(3)
χ	Taxa R→S (imunidade post-COVID)	0.025	(4)

Fontes: (1) (Eikenberry *et al.*, 2020); (2) (Niu *et al.*, 2020); (3) calculados usando o openDataSUS; (4) (Kirkcaldy *et al.*, 2020).

Considerando agora a vacinação, serão utilizados os dados da vacina da Pfizer-BioNTech, que apresenta uma grande eficácia (Hassine, 2021) e foi a primeira vacina a ser utilizada. Como as vacinas só se tornaram disponíveis depois que a COVID-19 já tinha se espalhado por todas as regiões, adotaremos a convenção de que os indivíduos serão vacinados quando a difusão da doença já tenha entrado em um regime estacionário e o número de infectados tenha se estabilizado. Ao chegar nesse ponto são vacinados uma fração F dos nós e devido à característica da vacina apenas os indivíduos nos estágios Suscetível, Recuperados ou Assinto-

máticos de acordo com um dado critério de prioridade. Com a vacinação, os parâmetros ε_S , ε_A , Λ_v , α^F , ψ^F , e τ^F que influenciam a velocidade de propagação da infecção e a probabilidade de assintomatismo são alterados. No modelo, as mortes estão associadas aos indivíduos que passam pelo estágio sintomático, e esses parâmetros são críticos para a minimização do impacto. As modificações ocorrem por meio da multiplicação da eficácia da vacina, e os novos valores estão detalhados na Tabela 5.

Tabela 4 – Probabilidades de transição por faixa etária.

Parâmetro	Definição	0 - 20	20 - 30	30 - 50	50 - 70	≥ 70	Notas
α^F	Prob. E→I	29.10%	37.40%	41.68%	39.40%	31.30%	(1)
ψ^F	Prob. I→H	0.408%	1.040%	3.890%	9.980%	17.500%	(2)
τ^F	Prob. H→D	1.040%	1.330%	1.380%	7.600%	24.000%	(3)

Fontes: ⁽¹⁾ (Jung *et al.*, 2020); ⁽²⁾ (Filho *et al.*, 2022); ⁽³⁾ calculada usando o openDataSUS.

Tabela 5 – Tabela da variação dos parâmetros perante a vacinação

Definição	Não Vacinado	Vacinado
Taxa de contágio Sintomática	ε_S	$0.058 \times \varepsilon_S$
Taxa de contágio Assintomática	ε_A	$0.058 \times \varepsilon_A$
Prob. E→I	α^F	$0.346 \times \alpha^F$
Prob. I→H	ψ^F	$0.034 \times \psi^F$
Prob. H→D	τ^F	$0.034 \times \tau^F$

Fonte: (Haas *et al.*, 2021).

3.4 Estratégias de vacinação propostas

As estratégias de vacinação são baseadas nas centralidades definidas nas Seções 2.2.2-2.2.3. Dessas centralidades, a CGR utiliza atributos dos nós por definição. Portanto, precisamos definir quais serão os seus respectivos valores no cálculo, para isso foram definidas duas abordagens: individualista e altruísta. Na abordagem altruísta, será calculada a centralidade do nó v considerando o risco do nó v infectar outros nós e eles morrerem, ou seja $\Theta_v = 1 - \alpha^F(v)$, $\Theta_\mu = \tau^F(\mu)$. Por outro lado, na abordagem individualista, o cálculo da centralidade do nó v considera o risco do nó v ser infectado por um vizinho e morrer, ou seja $\Theta_v = \tau^F(v)$, $\Theta_\mu = 1 - \alpha^F(\mu)$. Ambos os parâmetros são encontrados na Tabela 4. Portanto, a CGR para o caso individualista é

$$CGR(v, s) = \sum_{\mu \neq v} \frac{(1 - \alpha^F(v)) \cdot \tau^F(\mu)}{(d_{v,\mu}^w)^s}; \quad (3.3)$$

já a Centralidade de Gravidade com abordagem altruísta (A-CGR) é definida como

$$A-CGR(v, s) = \sum_{\mu \neq v} \frac{(1 - \alpha^F(\mu)) \cdot \tau^F(v)}{(d_{v,\mu}^w)^s}. \quad (3.4)$$

Uma outra mudança é na UT, a qual é definida originalmente apenas para atributos nas arestas e não para atributos dos nós. Propomos calcular a UT com a seguinte modificação:

$$UT(v) = \sum_{\mu \in \eta(v)} \left(\frac{w_{v,\mu} \Theta_v}{k_v^w} + \frac{w_{v,\mu} \Theta_\mu}{k_\mu^w} + \frac{w_{v,\mu}^2 \Theta_v \cdot \Theta_\mu}{k_v^w k_\mu^w} \right), \quad (3.5)$$

em que Θ_v e Θ_μ , são os pesos de cada vértice que podem ser a abordagem altruísta ou individualista.

Além das centralidades previamente discutidas, propomos cinco novas medidas para complementar as abordagens tradicionais presentes na literatura: Probabilidade de pelo menos um dos vizinhos serem hospitalizados (PH), Probabilidade de pelo menos um dos vizinhos serem hospitalizados dado que o sítio é assintomático (PHA), Probabilidade de pelo menos um do sítio ser hospitalizado e morrer (PM), Probabilidade de pelo menos um dos vizinhos morrerem dado que foram hospitalizados e que o sítio é assintomático (PMA) e Centralidade de Grau Morte (GM). As fórmulas correspondentes a essas centralidades são definidas como segue:

– PH é dada por:

$$PH(v) = 1 - \prod_{\mu \in \eta(v)} \left(1 - \psi^{F(\mu)} \right);$$

– PHA por:

$$PHA(v) = \left(1 - \alpha^{F(v)} \right) \left(1 - \prod_{\mu \in \eta(v)} \left(1 - \psi^{F(\mu)} \right) \right);$$

– PM por:

$$PM(v) = 1 - \prod_{\mu \in \eta(v)} \left(1 - \alpha^{F(\mu)} \cdot \tau^{F(\mu)} \cdot \psi^{F(\mu)} \right);$$

– PMA por:

$$PMA(v) = \left(1 - \alpha^{F(v)} \right) \left(1 - \prod_{\mu \in \eta(v)} \left(1 - \alpha^{F(\mu)} \cdot \tau^{F(\mu)} \cdot \psi^{F(\mu)} \right) \right);$$

– e GM por:

$$GM(v) = k \alpha^{F(v)} \cdot \psi^{F(v)} \cdot \tau^{F(v)}.$$

Essas centralidades foram elaboradas tendo em mente a aplicabilidade em modelos epidemiológicos. Elas são formuladas de maneira simples e possuem uma lógica intuitiva: focar na vacinação de nós (ou sítios) cujos vizinhos possuem maior probabilidade de morrer, ou que eles próprios estejam em maior risco. Essa abordagem visa otimizar a alocação de recursos em cenários de controle de surtos, identificando nós críticos cuja imunização teria o maior impacto na contenção da propagação da doença.

3.5 Ambiente computacional

Para a realização dos experimentos e simulações presentes neste estudo, foram utilizadas duas máquinas com diferentes configurações de hardware. A seguir, descrevemos as especificações técnicas de ambas as máquinas utilizadas:

- Máquina 1:
 - Modelo de Hardware: ASUS System Product Name
 - Memória RAM: 16,0 GiB
 - Processador: AMD Ryzen 5 4600g com gráficos Radeon, 12 núcleos de processamento
- Máquina 2:
 - Modelo de Hardware: INTEL H61
 - Memória RAM: 16,0 GiB
 - Processador: Intel Core i7-3770 CPU @ 3.40GHz, 8 núcleos de processamento

Ambas as máquinas utilizam o sistema operacional Ubuntu 22.04.4 LTS, com arquitetura de 64 bits. O ambiente gráfico é gerido pelo GNOME versão 42.9 e o sistema de janelas Wayland. Para os cálculos de tempo computacional das centralidades foram utilizadas a Máquina 1.

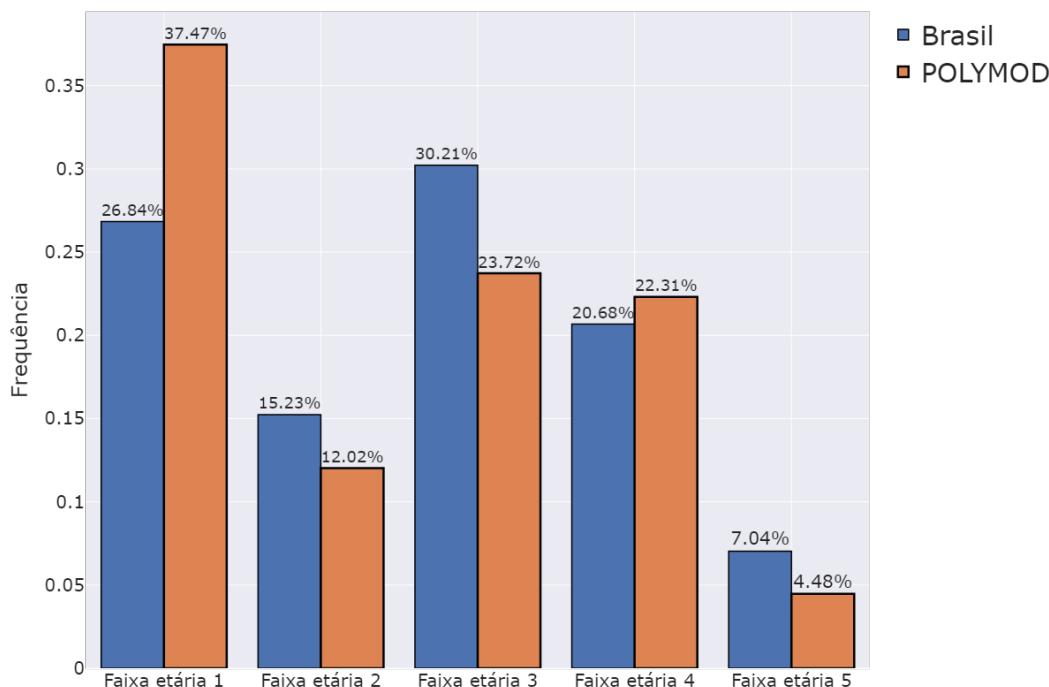
Para a execução das principais simulações deste estudo, foi empregada a linguagem de programação C em conjunto com as bibliotecas mtwister (Matsumoto *et al.*, 1998) e igraph (Csárdi *et al.*, 2024). Para a análise e visualização dos dados obtidos, recorreu-se à linguagem Python, utilizando bibliotecas como matplotlib (Hunter, 2007), pyplot e numpy (Harris *et al.*, 2020), que oferecem recursos avançados para processamento de dados e geração de gráficos informativos.

4 RESULTADOS

4.1 Análise exploratória dos dados do POLYMOD

A partir da coleta de dados realizada pelo POLYMOD, o propósito é elucidar a estruturação das interconexões interpessoais e analisar as implicações decorrentes dessas relações visando uma modelagem em redes. Para isso cada indivíduo foi 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; a distribuição de cada faixa etária é mostrada na Figura 15 que foi obtida a partir do Censo de 2022 do IBGE (IBGE,) em comparação às faixas do POLYMOD. Esta categorização reveste-se de relevância considerável, haja vista que a estimação dos parâmetros utilizados é só viável com a categorização das idades.

Figura 15 – Distribuição percentual de Faixas etárias do Brasil X POLYMOD



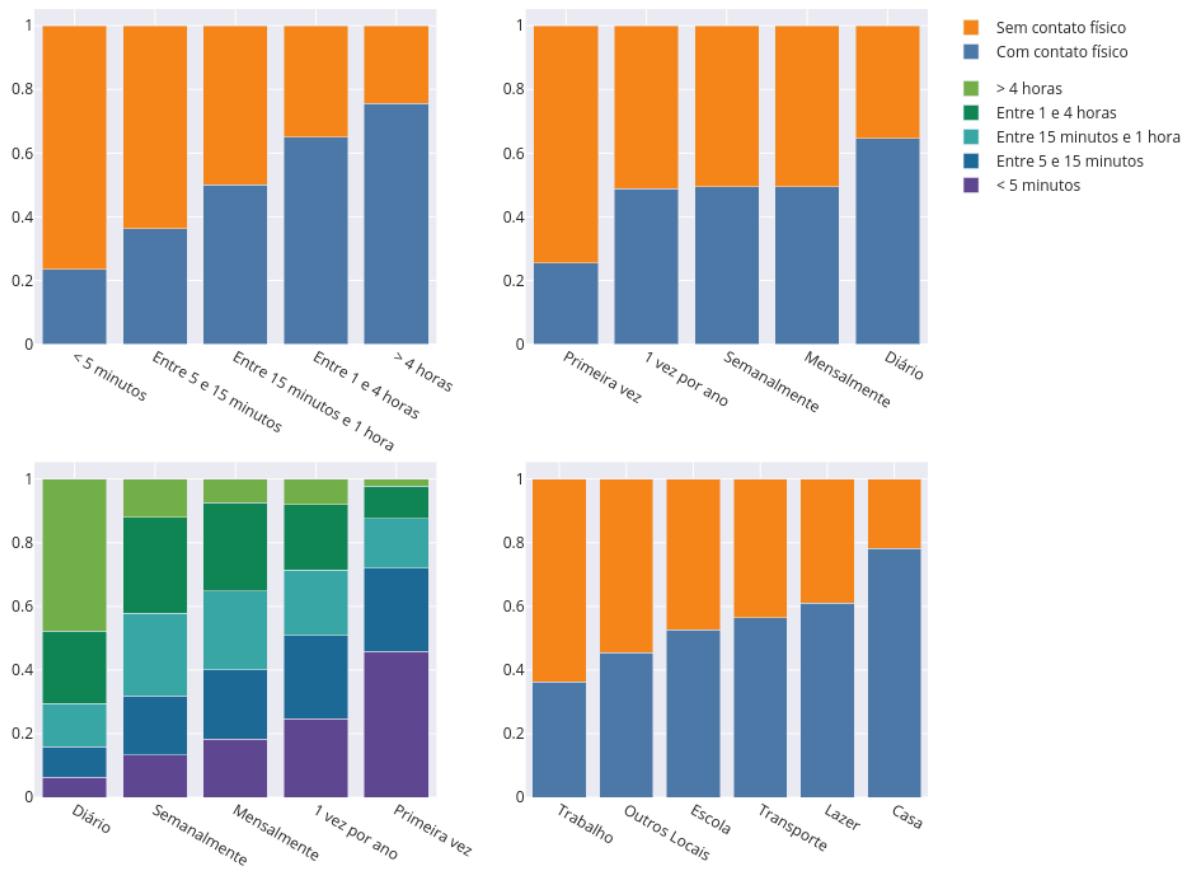
Fonte: Elaborado pelo autor.

Nota: Este gráfico de barras ilustra a distribuição percentual de diferentes faixas etárias no Brasil do Censo 2022 IBGE em comparação com os dados do estudo POLYMOD. As faixas etárias são categorizadas em cinco grupos distintos. Observa-se que, na Faixa Etária 1, o estudo POLYMOD apresenta uma maior frequência em comparação com o Brasil. Na Faixa Etária 3, a situação se inverte, com o Brasil mostrando uma maior frequência em comparação ao estudo POLYMOD. A Faixa Etária 5 exibe a menor frequência em ambos os casos, com o Brasil ligeiramente superior ao POLYMOD. Esse será um dos *inputs* do modelo posteriormente.

Com base nos dados obtidos, foi realizada uma investigação para avaliar a duração, frequência e a presença de contato físico nos encontros em análise. A Figura 16 mostra os padrões dos contatos do POLYMOD; observa-se que, à medida que a frequência ou a duração do

contato aumenta, a probabilidade de ocorrer contato físico também cresce. Além disso, quanto mais frequente é o contato maiores as chances da conversa durar mais, por exemplo em contatos diários existe mais de 50% de chance de que o contato dure mais que 1 hora. Por fim, o último gráfico da Figura 16 evidencia que, no ambiente domiciliar, a probabilidade de ocorrência de contato físico é significativamente maior. Por contraste, no ambiente profissional, as chances são substancialmente menores, sendo considerado o cenário menos propenso a esse tipo de interação.

Figura 16 – Frequência e Duração de Interações (com/sem Contato Físico)



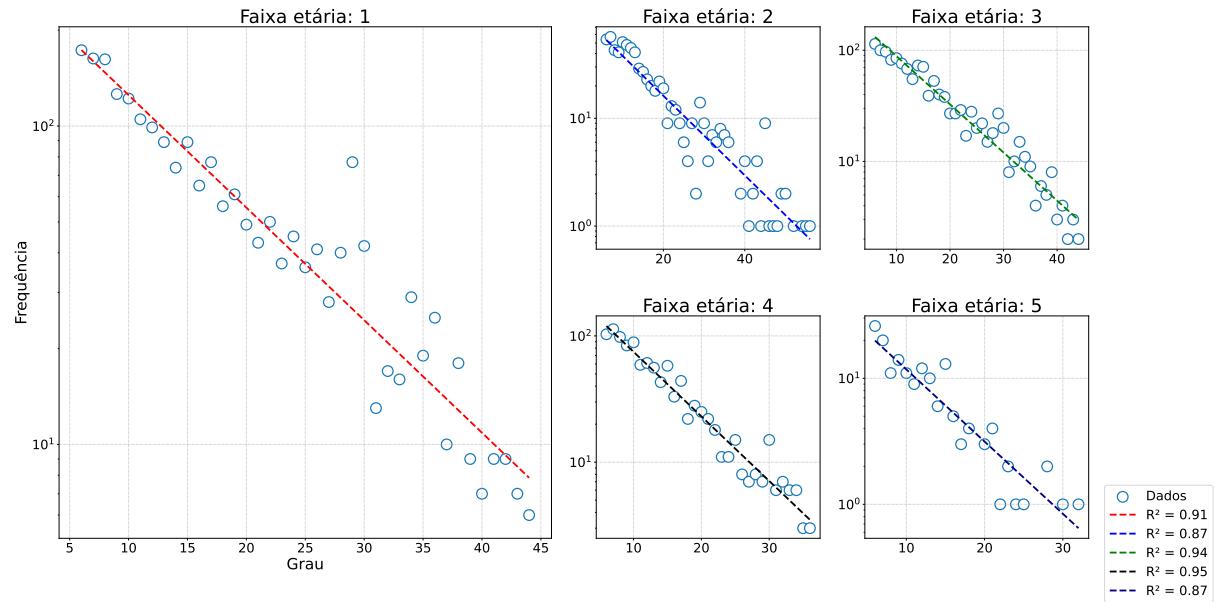
Fonte: Elaborado pelo autor.

Nota: Os dois gráficos da primeira linha mostram que, quanto mais frequente ou mais duradoura é uma interação, maior é a chance de acontecer contato físico. No terceiro gráfico (inferior esquerdo), quanto mais frequente o contato mais prolongado ele se torna. O quarto gráfico (inferior direito) revela que em locais como trabalho e escola, a maioria das interações ocorre sem contato físico, ao passo que, em casa, o contato físico é mais comum.

Um elemento importante que foi investigado é como as idades influenciam na quantidade de conexões que o próprio indivíduo possui. A Figura 17 ilustra a distribuição de graus por faixa etária dos respondentes da pesquisa do POLYMOD. Ao analisar a distribuição de ligações por faixa etária é percebido que a cauda segue uma distribuição geométrica, que pode ser escrita como $P \propto (1 - p)^x p$ ou $P \propto e^{-\lambda x} (1 - e^{\lambda})$. Considerando a forma $P \propto e^{-\lambda x} (1 - e^{\lambda})$,

na Figura 17, determina-se o coeficiente angular $-\lambda$ e seu respectivo R^2 , por faixa etária, através do método dos mínimos quadrados.

Figura 17 – Influência das idades nas conexões

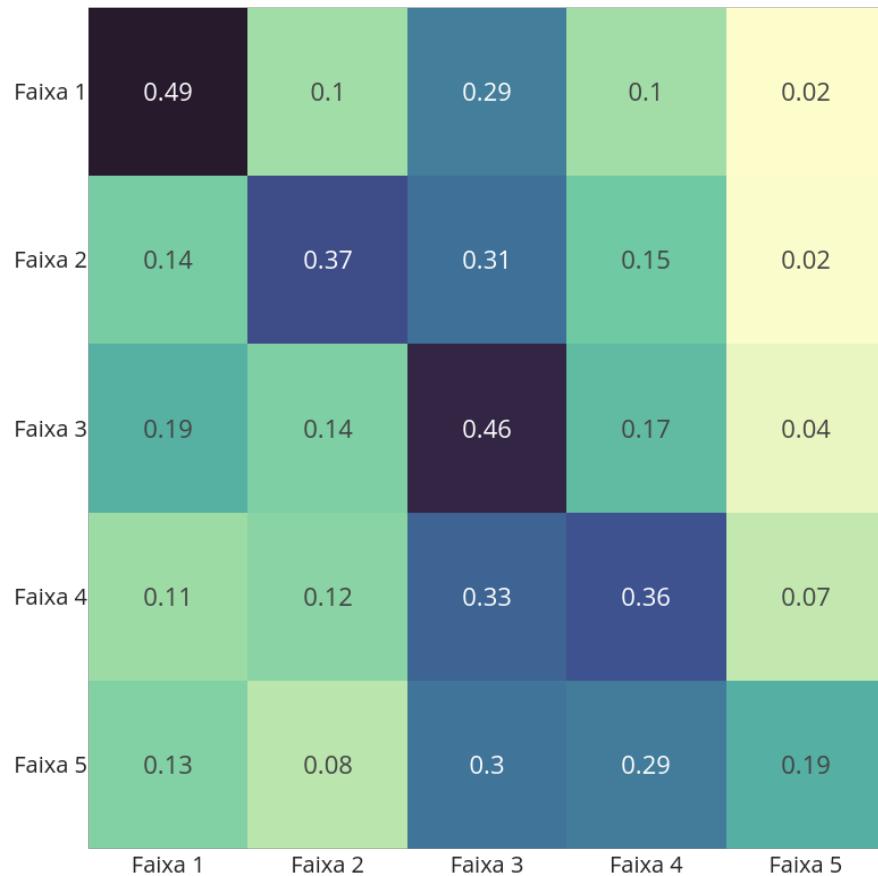


Fonte: Elaborado pelo autor.

Nota: Ilustra a influência das idades nas conexões da rede social, observando uma distribuição geométrica nas ligações por faixa etária.

Também foi analisada a relação entre a faixa etária de um indivíduo que preencheu o questionário do POLYMOD e a distribuição de suas ligações por faixas etárias. A Figura 18 mostra a média da frequência de ligações entre diferentes faixas etárias. Cada linha representa uma faixa etária e a fração das conexões que ela estabelece com outras faixas. Nota-se que as ligações mais frequentes ocorrem entre faixas etárias com elas mesmas, conforme evidenciado pelos valores mais altos na diagonal principal com exceção da última linha, e com a Faixa etária 3 pois, provavelmente, essa é a faixa dos adultos com filhos e também que tem que cuidar dos pais mais idosos. Já as faixas etárias mais distantes têm menor interação, como indicado pelos valores mais baixos nas extremidades da matriz.

Figura 18 – Frequência média de conexões por faixa etária



Fonte: Elaborado pelo autor.

Nota: As ligações mais frequentes ocorrem entre faixas etárias iguais (exceto a última linha) e com a Faixa Etária 3, indicando interação predominante entre grupos semelhantes e adultos cuidando de filhos e pais idosos; faixas etárias distantes têm menor interação.

Por último, a Figura 19 mostra a esquerda a média do tempo de contato (em minutos) entre diferentes faixas etárias, enquanto a figura à direita exibe o desvio padrão desse tempo. Observa-se que os maiores tempos ocorrem na interação com a faixa etária 1, evidenciando o fato de que indivíduos nesta faixa requerem maiores cuidados dos seus responsáveis mais velhos. A exceção se dá na faixa etária 5, que possui maior tempo de contato consigo mesma. Na matriz de variância, os contatos que exibem menor variabilidade são aqueles com a faixa etária 5, exceto os contatos da faixa 5 consigo mesma, indicando maior consistência na duração desses contatos. Por outro lado, a maior variabilidade dos contatos ocorrem nas interações com a faixa 1.

Figura 19 – Média e desvio-padrão dos tempos de contato entre faixas etárias

	Faixa 1	Faixa 2	Faixa 3	Faixa 4	Faixa 5		Faixa 1	Faixa 2	Faixa 3	Faixa 4	Faixa 5
Faixa Etária	173.36	132.58	151.53	122.38	119.5		118.05	117.75	123.27	110.97	109.13
	132.58	123.83	93.86	89.14	78.31		117.75	112.14	111.13	107.44	96.37
	151.53	93.86	99.52	81.92	89.53		123.27	111.13	112.82	101.88	102.43
	122.38	89.14	81.92	110.09	87.53		110.97	107.44	101.88	116.49	99.57
	119.5	78.31	89.53	87.53	135.62		109.13	96.37	102.43	99.57	114.35
	Faixa 1	Faixa 2	Faixa 3	Faixa 4	Faixa 5		Faixa 1	Faixa 2	Faixa 3	Faixa 4	Faixa 5

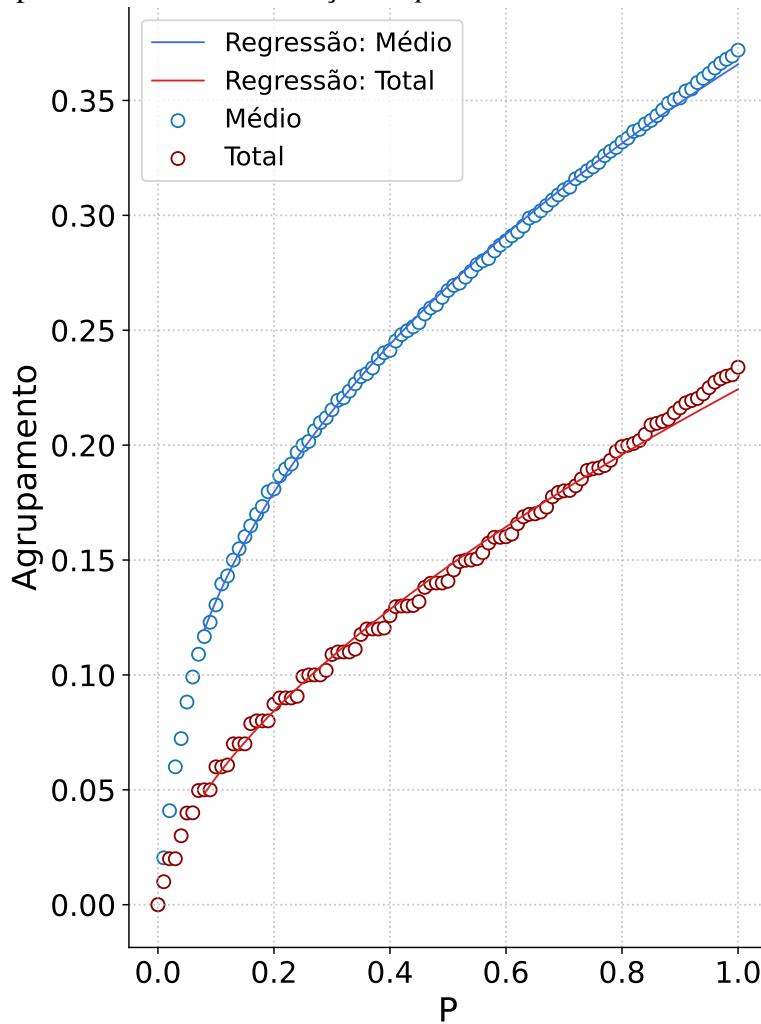
Fonte: Elaborado pelo autor.

Nota: Matriz da média e desvio padrão do tempo de contato (em minutos) entre diferentes faixas etárias. As interações com a faixa etária 1 têm maior média e variância, enquanto as interações com a faixa 5 tem menor média e menos variância, exceto as interações da faixa 5 consigo mesma.

4.2 Resultados da rede gerada

Os dados das Figuras 17 e 18 foram utilizados como *inputs* do Modelo de Configuração Estratificado (Algoritmo 5). Especificamente, com os dados da Figura 17 foi obtida a distribuição empírica do grau de cada nó de cada faixa etária e com os dados da Figura 18 foram estimados os parâmetros de uma distribuição multinomial para estabelecer quantas ligações vão para cada faixa etária. A partir disso e dado um valor de p (parâmetro que aumenta o coeficiente de agrupamento) foi construída a rede para acontecer a simulação da epidemia. A Figura 20 mostra a variação do agrupamento com o incremento de p . Observa-se que o agrupamento, C , cresce de acordo com $C \propto p^\alpha$, tanto o agrupamento médio quanto o total no qual $\alpha_{medio} = 0.44$ e $\alpha_{total} = 0.60$, respectivamente.

Figura 20 – Agrupamento da rede em função de p

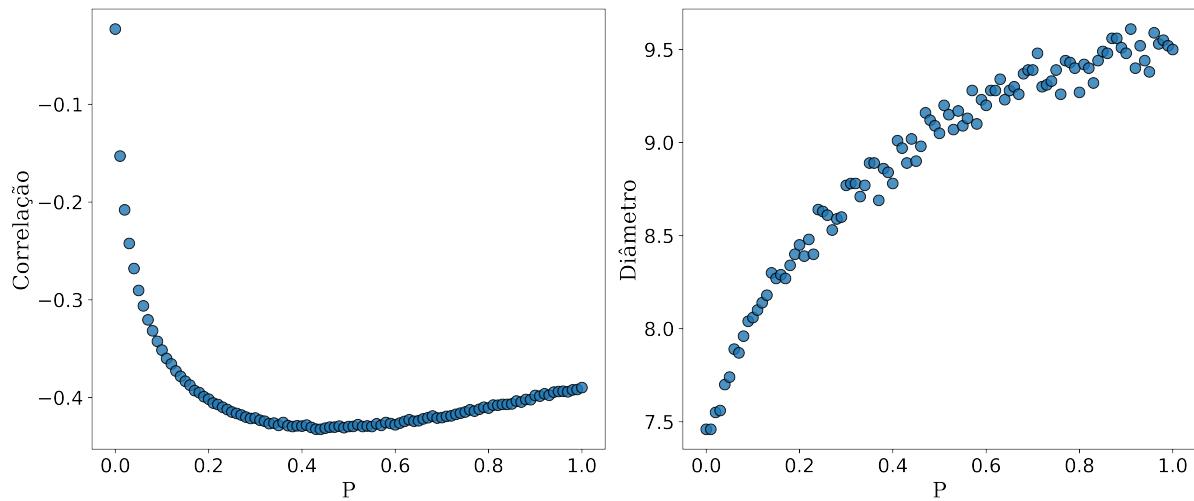


Fonte: Elaborado pelo autor.

Nota: Evolução do agrupamento da rede em função do incremento do valor de p , foi utilizada uma regressão linear para saber qual a função que rege encontrando $C \propto p^\alpha$ com $R^2 = 99\%$.

A Figura 21 apresenta a esquerda a correlação linear entre o Agrupamento e o Grau com o incremento de p , nele é possível notar que as correlações são sempre negativas, ou seja nós com maior grau tendem a ter menor agrupamento, algo observado em redes sociais (Bokányi *et al.*, 2023), pois para atingir o mesmo agrupamento local precisa-se de mais ligações em nós com maior grau. Não se sabe ao certo se isso é uma limitação do algoritmo, da distribuição de graus obtida nos dados ou algo inevitável para esse tipo de problema. O gráfico a direita mostra que o diâmetro aumenta com o incremento de p isso acontece pois como o número de ligações é, em média, fixo formam-se ilhas na rede fazendo com que localmente seja mais conectado, mas para chegar em pontos mais distantes seja mais difícil.

Figura 21 – Variação da Correlação entre Agrupamento e Grau, Diâmetro em função de p

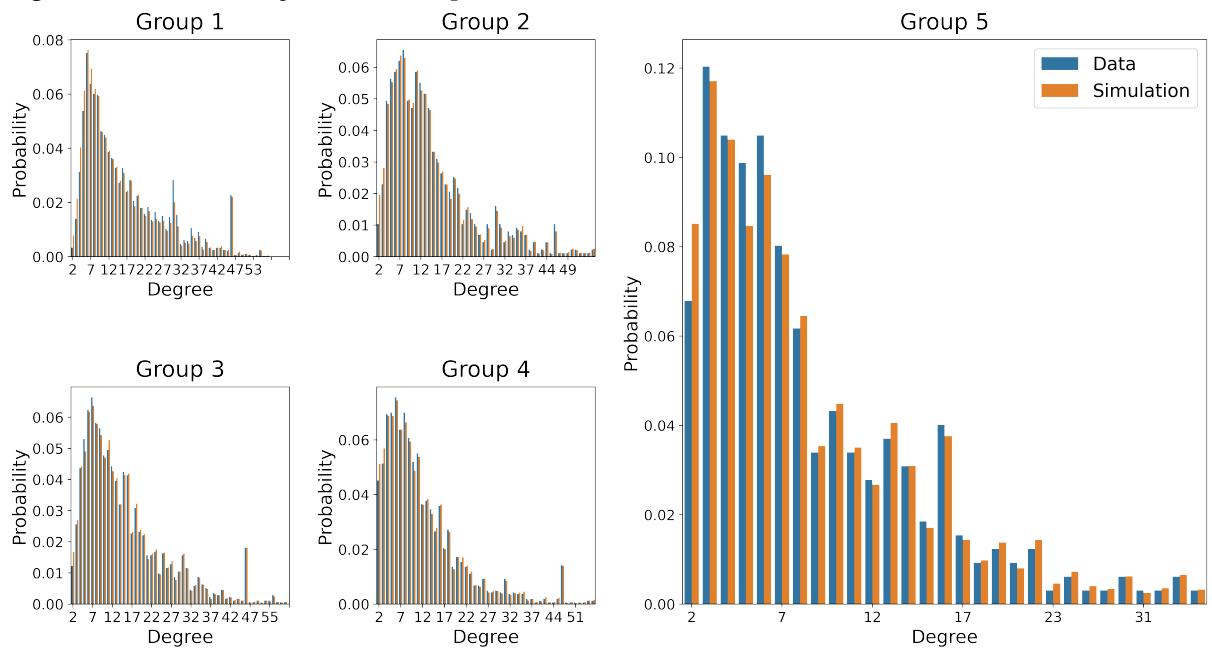


Fonte: Elaborado pelo autor.

Nota: (Esquerda) Correlação entre agrupamento médio e grau; com o aumento de p , a correlação diminui, indicando que nós com maior grau têm menor agrupamento. (Direita) Diâmetro da rede em função de p ; o diâmetro aumenta com p devido à formação de ilhas na rede, tornando mais difícil alcançar pontos distantes.

Após a geração da rede com o algoritmo proposto, foram calculados alguns os dados da rede e observou-se que para qualquer valor de p a distribuição de graus tinha a mesma distribuição que nos dados como mostra a Figura 22, isso poderia não acontecer dado que a união de ligações do modelo são aleatórias o que poderia causar sobra de ligações. Contudo o algoritmo apresentou uma perca de apenas 3% das ligações. Na Figura 23, mostra-se a diferença absoluta ao final do algoritmo da proporção de conexões de cada faixa etária em comparação com a Figura 18, tendo uma diferença maior entre as conexões com faixa etária 1.

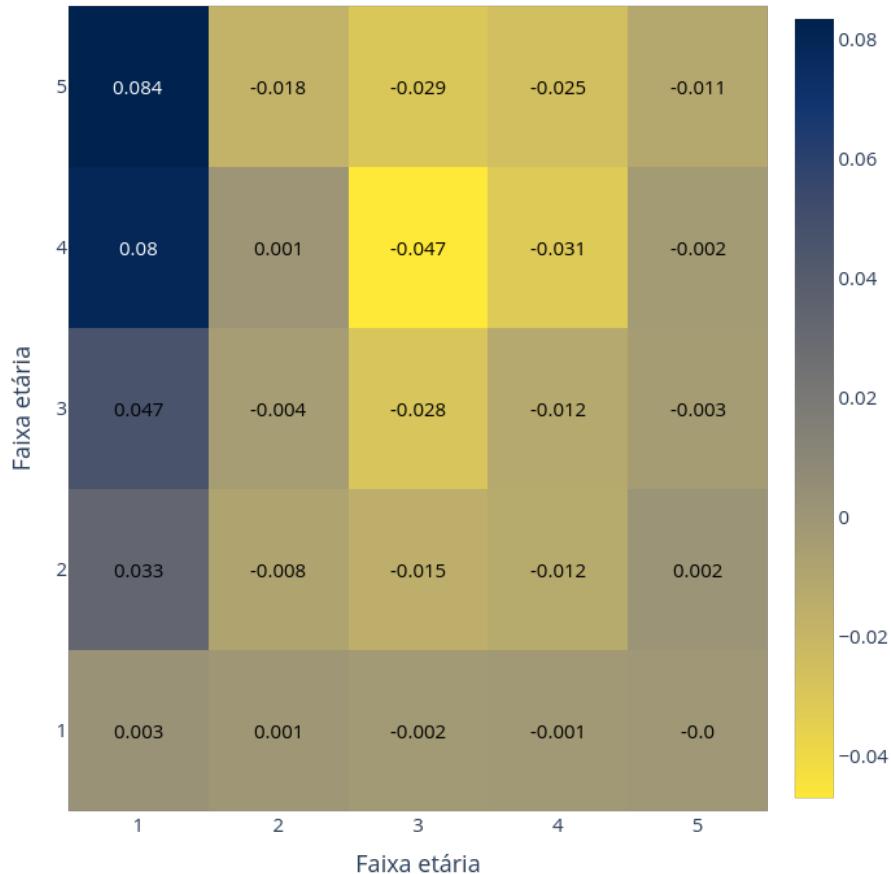
Figura 22 – Distribuição de Graus $p = 1.0$



Fonte: Elaborado pelo autor.

Nota: Comparação de como fica a distribuição de graus comparando modelo com $p = 1.0$ com dados do POLYMOD, é possível ver todas as faixas foram bem no teste de comparação.

Figura 23 – Resultados da diferença absoluta entre a matriz de proporções de conexões entre faixas etárias encontradas no modelo e no POLYMOD



Fonte: Elaborado pelo autor.

Nota: A figura mostra a diferença absoluta entre a matriz de proporções de conexões entre as faixas etárias do modelo e do encontrado no POLYMOD; essa diferença se mantém constante independe do valor de p .

4.3 Resultados do modelo de infecção

Utilizando o Modelo de Configuração Estratificado, com entradas a distribuição de graus empírica, na distribuição de faixas etárias brasileiras e a distribuição de ligações entre faixas etárias do banco de dados do POLYMOD, os parâmetros do modelo de contágio foram obtidos da literatura ou calculados a partir do OpenDataSUS. As médias e desvios do tempo de contato entre indivíduos, também derivadas do POLYMOD, foram incorporadas como pesos nas arestas do modelo utilizando uma distribuição normal com os parâmetros estimados. Com esses dados, iniciou-se a simulação da infecção.

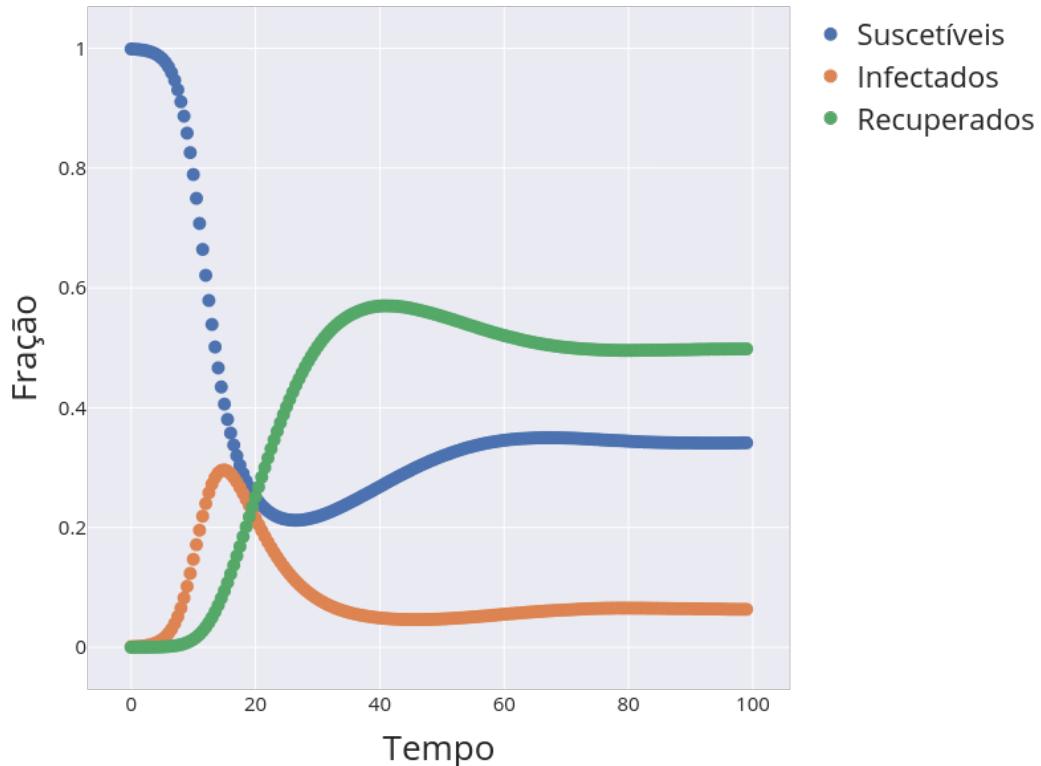
A simulação acontece do dia 0 até o dia 465 iniciando com 10 nós infectados para garantir que a simulação não termine cedo, caso tivesse menos nós infectados haveria a chance da infecção se extinguir no compartimento Exposto e todos os outros sítios como Suscetíveis. A vacinação ocorre no dia 100 (data que a infecção já entrou em regime estacionário ver Figura 24),

nesse dia são aplicadas vacinas para uma fração f da rede para sítios no compartimento Suscetível, Exposto, Recuperado ou Assintomático dado um critério de prioridade que é determinado pelo valor das centralidades introduzidas anteriormente. Se o indivíduo não estiver apto a ser vacinado no dia e após todos os aptos serem vacinados ainda existirem vacinas sobrando, então ao atingir um dos compartimentos aptos nos dias seguintes, ele será vacinado.

Tendo simulado o modelo epidemiológico na rede gerada pelo Modelo de Configuração adaptado, foi observado que após o dia 100 a infecção estabilizava e esse seria o dia da vacinação. Deste modo, simula-se a situação real na qual a vacina só se tornou disponível após a COVID-19 já ter se espalhado por todo o mundo. Anteriormente a esse dia a rede se comporta, para $p = 0.0$ (com agrupamento médio de 0.01), de acordo com as Figuras 24 e 25, nas quais foi feita uma média com 400 simulações.

A Figura 24 mostra a evolução da epidemia, em que pode-se observar que a proporção de indivíduos suscetíveis diminui rapidamente até o dia 25 para posteriormente aumentar e atingir um estado de estabilidade. A fração de indivíduos infectados (a soma da fração de assintomáticos com a de sintomáticos) aumenta até atingir um pico, com 30% da população total sendo infectada simultaneamente, seguido por um declínio. Finalmente, a fração de indivíduos recuperados aumenta até que 58% da população esteja nesse estado, estabilizando posteriormente.

Figura 24 – Fração de Suscetíveis, Infectados e Recuperados antes da aplicação de qualquer estratégia de vacinação

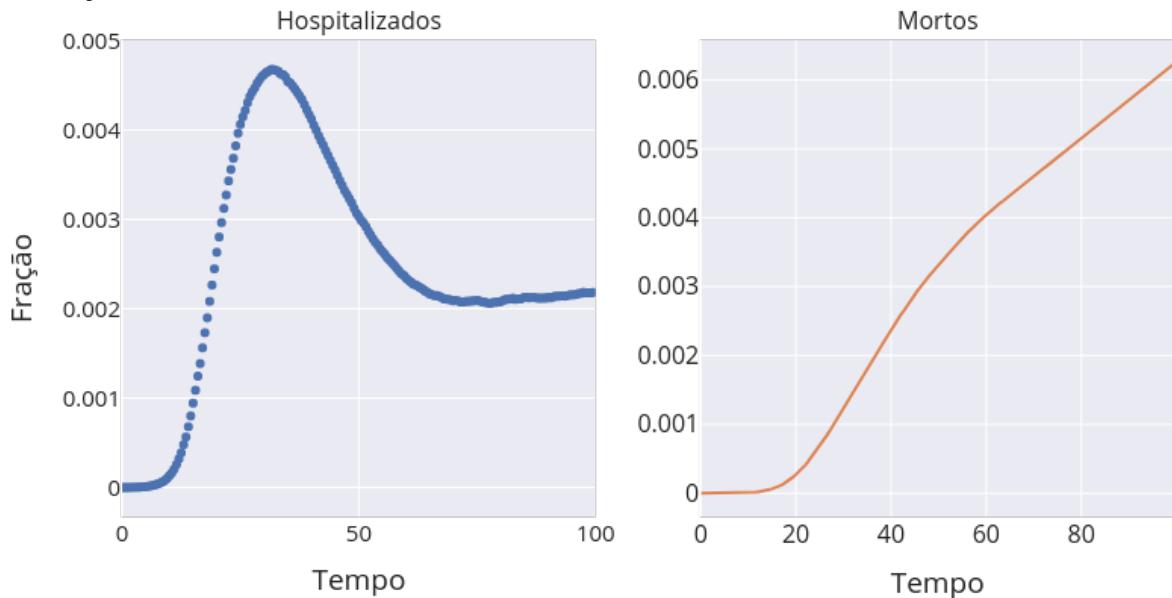


Fonte: Elaborado pelo autor.

Nota: Resultados antes da aplicação de qualquer estratégia de vacinação para uma rede $N = 10000$ com $p = 0$ após 400 realizações. Inicialmente, a proporção de suscetíveis diminui rapidamente, seguida por um aumento e estabilização. A fração de infectados atinge um pico, afetando 30% da população, e depois diminui. Paralelamente, a proporção de recuperados aumenta e estabiliza em 58% da população.

A Figura 25 mostra a evolução da fração de mortos e a fração de hospitalizados. Observa-se que o número máximo de hospitalizados observado foi cerca de 0.46% da rede, enquanto o número de mortos cresce indefinidamente. Neste trabalho, não consideramos a capacidade do sistema hospitalar, pois neste caso, precisaríamos de informações sobre a taxa de recuperação e de morte de sintomáticos que precisariam ser internados e não foram por falta de leitos, que não são facilmente determinadas.

Figura 25 – Fração de Hospitalizados e Mortos antes da aplicação de qualquer estratégia de vacinação

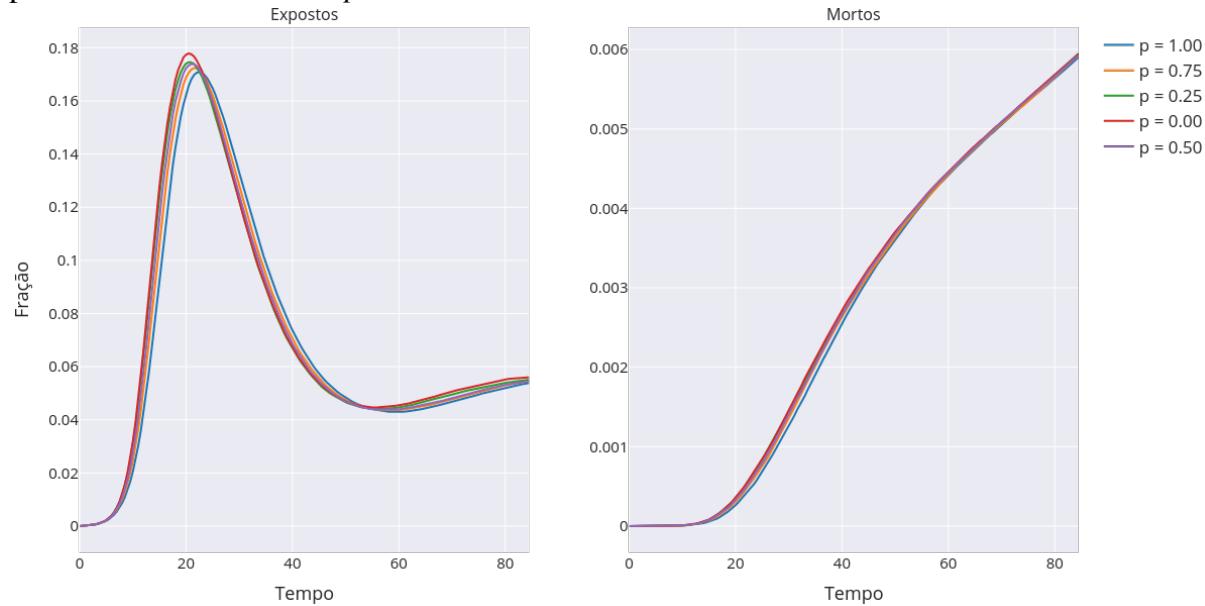


Fonte: Elaborado pelo autor.

Nota: Simulações foram feitas para uma rede $N = 10000$ com $p = 0$ após 400 realizações. O gráfico mostra que a Fração de Hospitalizados chega a um limite máximo de pessoas que podem estar no estágio ao mesmo tempo enquanto que a fração de mortos cresce indefinidamente até chegar em aproximadamente 0.46% da rede.

Com o aumento de p surgem diferenças com o comportamento da infecção; a Figura 26 mostra que a rede tem uma quantidade próxima de mortos, contudo a fração de expostos demora mais para crescer, decair e estabilizar, visto que com o maior agrupamento a infecção fica presa em regiões com alto agrupamento e tem muita dificuldade em avançar para outras regiões da rede, o que já é esperado devido ao aumento no diâmetro apresentado pela Figura 21.

Figura 26 – Fração de Expostos e Mortos antes da aplicação de qualquer estratégia de vacinação para diferentes valores de p

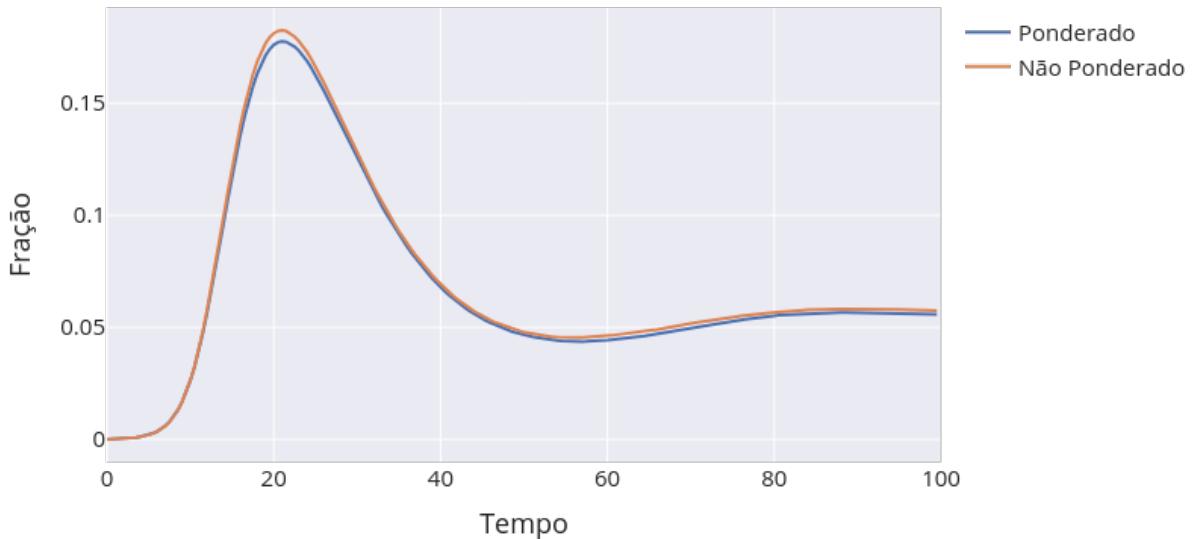


Fonte: Elaborado pelo autor.

Nota: Resultados para fração de Expostos e Mortos em função do tempo para diferentes valores de p . Com o seu incremento a epidemia fica presa em regiões muito agrupadas não conseguindo chegar em outras regiões mais distantes.

Para finalizar, a Figura 27 mostra comparação entre como a infecção se espalha nos primeiros 100 dias no caso ponderado nas arestas e não ponderado. Nesse caso não houve uma diferença significativa em nenhuma evolução temporal dos compartimentos. Apesar disso, foi feita a análise das centralidades comparando os dois casos.

Figura 27 – Comparação entre a ponderação nas arestas e sem ponderação Sintomáticos



Fonte: Elaborado pelo autor.

Nota: Nesse gráfico mostra a diferença não significativa entre o caso não ponderado nas arestas e ponderado nas arestas para os Sintomáticos.

4.4 Resultado das estratégias de vacinação

A simulação foi realizada considerando um período de 100 dias, a vacina foi aplicada em ordem de prioridade nós com maior valor de diversas centralidades em redes apresentadas anteriormente no Capítulo 2 e algumas com a sua versão individualista ou altruísta. Ao final foram coletadas três métricas de desempenho das estratégias: a soma de todos os tempos de hospitalização de todos os indivíduos da rede T_H , a fração de mortos F_M e a fração de indivíduos vacinados crítica para extinguir a doença f_c . Assim para cada uma dessas métricas foi feito um *ranking* para as melhores centralidades. Como os valores de T_H e F_M dependem de f , a fração de vacinados, a métrica utilizada foi a área abaixo do gráfico das curvas de T_H e F_M em função de f . Para determinar qual centralidade apresentou o melhor desempenho geral, foi calculada a média dos valores normalizados das três métricas de desempenho para cada centralidade. A normalização foi feita através de uma transformação linear dos valores de cada métrica para um intervalo entre 0 e 1, de modo que os valores mais baixos fossem transformados para 0 e os mais altos para 1. A qualidade do desempenho de cada centralidade depende da estrutura da rede, ou seja, com a alteração de p o *ranking* poderá ser diferente e podem existir diferentes melhores estratégias, sendo, neste caso, necessário estimar os valores de p da rede real para saber qual seria a melhor estratégia a ser utilizada na prática. Contudo será mostrado a posteriori que o valor de p não influencia no *ranking*.

Para o caso sem ponderação, a Figura 28 ilustra como ficaram as curvas de cada métrica em função de f e um gráfico de barras mostrando os valores de f_c para $p = 0.0$ e $p = 1.0$ com 4 estratégias: o caso de vacinação aleatória (Centralidade aleatória (CR)), o de vacinação prioritária dos mais idosos (Centralidade de Idade (CI)) a melhor em cada métrica e a melhor na média das três métricas. Quando o gráfico apresenta apenas 3 curvas significa que a melhor na métrica é a mesma para melhor na média. Esse esquema será repetido na Figura 29 para o caso com ponderação nas arestas. Para o leitor interessado, os resultados de todas as centralidades (37 para caso não ponderado e 63 para caso ponderado) estão no Apêndice A.

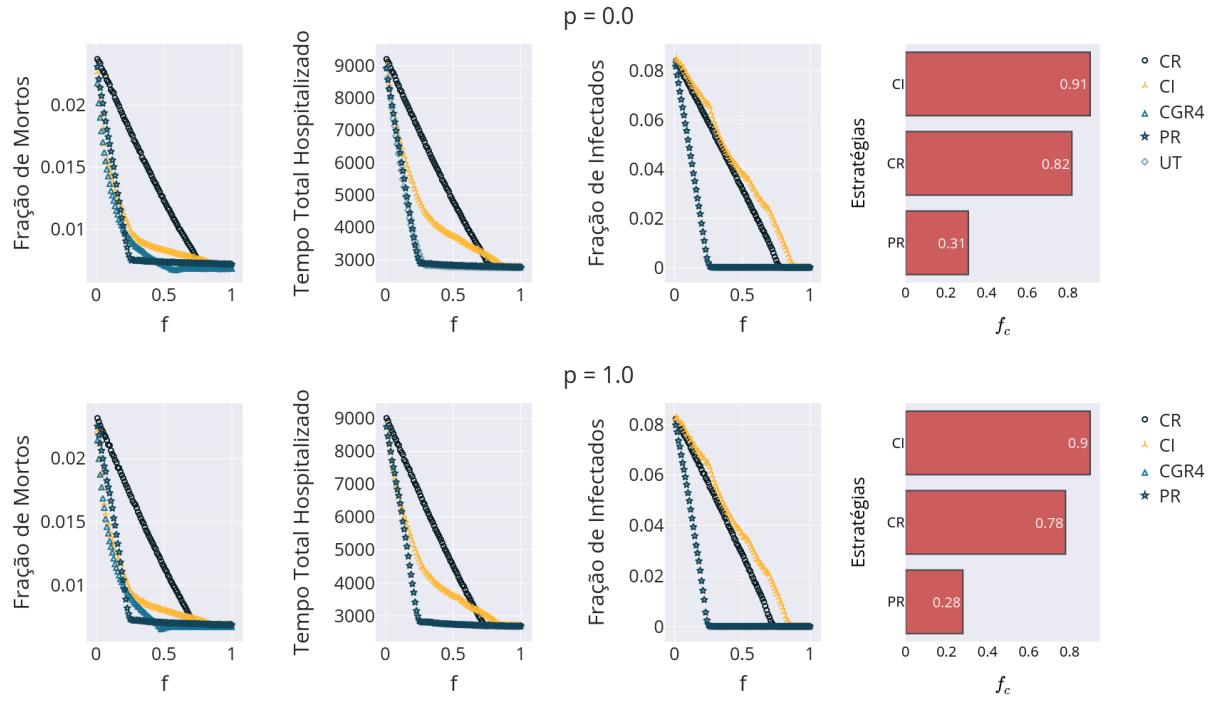
É possível perceber que em todas as métricas e mudando o valor de p a idade não foi a melhor estratégia em nenhuma, apesar de ter sido o principal critério utilizado em muitos países. Um resultado importante é que para extinguir a doença, usando o critério idade, seria necessário vacinar cerca de 90% da população, sendo inclusive pior que a vacinação aleatória. Vale ressaltar que usando a melhor estratégia nessa métrica, o PR, foram necessários apenas 31% para extinguir a doença para $p = 0.0$. Porém é notável que para uma fração menor de 15% da rede vacinada, a idade foi muito boa tanto no quesito de hospitalização como no de diminuição de óbitos, devido as complicações da doença estarem associadas à idade. Esse resultado se repete com o incremento de p e com a adição de ponderação nas arestas.

Embora o PR seja o melhor na média, o CGR4 apresentou ser o melhor na mortalidade para $p = 0$ e $p = 1$ no caso não ponderado, apesar de ter um decrescimento mais lento seus valores estabilizados conseguiram ser menores que o PR.

No caso ponderado nas arestas, com $p = 0.0$ e $p = 1.0$, respectivamente, a versão altruísta da utilidade, A-UT, e a CGR5 foram melhores na mortalidade; A-UT e A-CGR5 para o tempo hospitalizado; a centralidade Laplaciana que leva conta dos pesos nas arestas, WCL, e PR para f_c . Finalmente, o *PageRank* com (WPR) e sem pesos nas arestas (PR) foram os melhores na média. (Ver Figura 29.)

Isso mostra que tanto no caso não ponderado como ponderado houve uma grande presença da centralidade do PR e de centralidades que têm maior peso nos mais próximos e que também levam os pesos deles no cálculo.

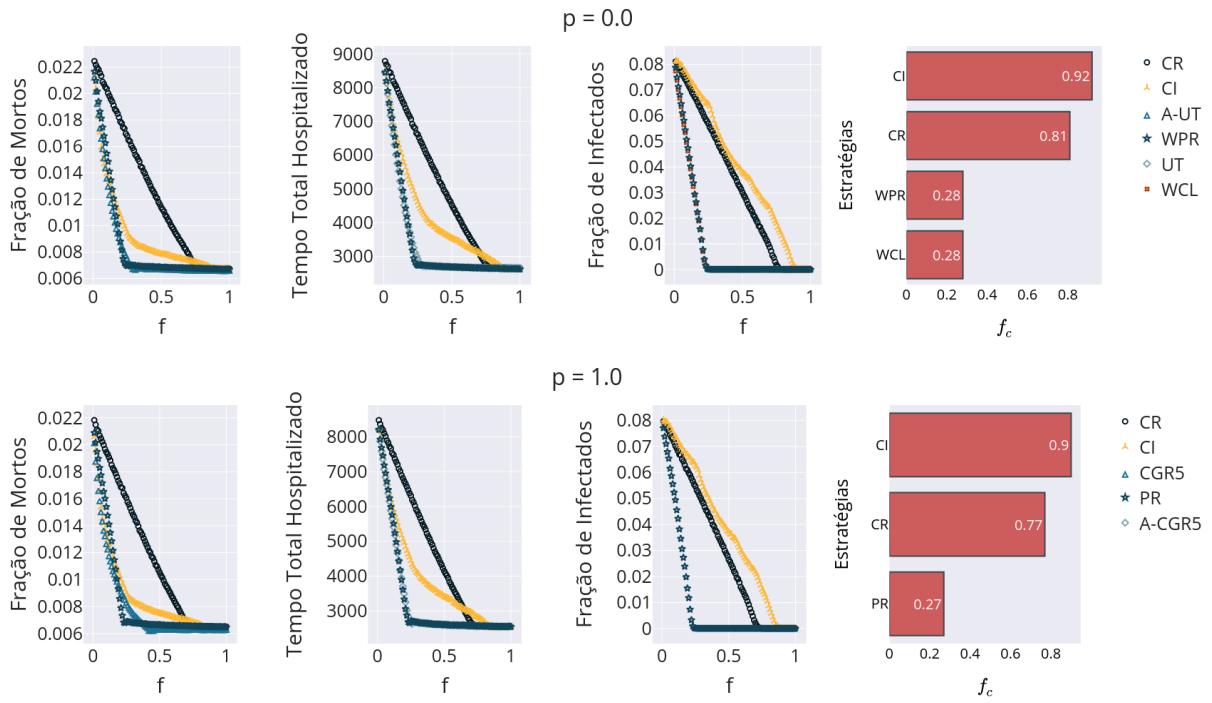
Figura 28 – Gráfico para cada métrica por fração de vacinados para $p = 0.0$ e $p = 1.0$ sem ponderação nas arestas



Fonte: Elaborado pelo autor.

Nota: Comparação de cada métrica, para os dois primeiros gráficos foi utilizado critério da integral enquanto na Fração de Infectados foi avaliado o valor de f quando a curva toca o 0 e o último mostra o valor de f_c para as curvas da Fração de Infectados. Comparando as melhores de cada métrica e a média contra a estratégia aleatória (CR) e idade (CI). Para $p = 0.0$, a melhor estratégia na média foi o *PageRank*, no caso da Fração de Mortos a melhor estratégia foi a Gravidade com expoente 4 (CGR4), na Fração de Infectados foi o *PageRank* (PR) e no Tempo Total Hospitalizado foi a Utilidade UT. No caso $p = 1.00$ obteve-se resultado muito parecido somente com a ausência da Utilidade. O último gráfico mostra o valor de f_c e que o PR precisou vacinar um terço do que o critério da idade.

Figura 29 – Gráfico para cada métrica por fração de vacinados para $p = 0.0$ e $p = 1.0$ com ponderação nas arestas

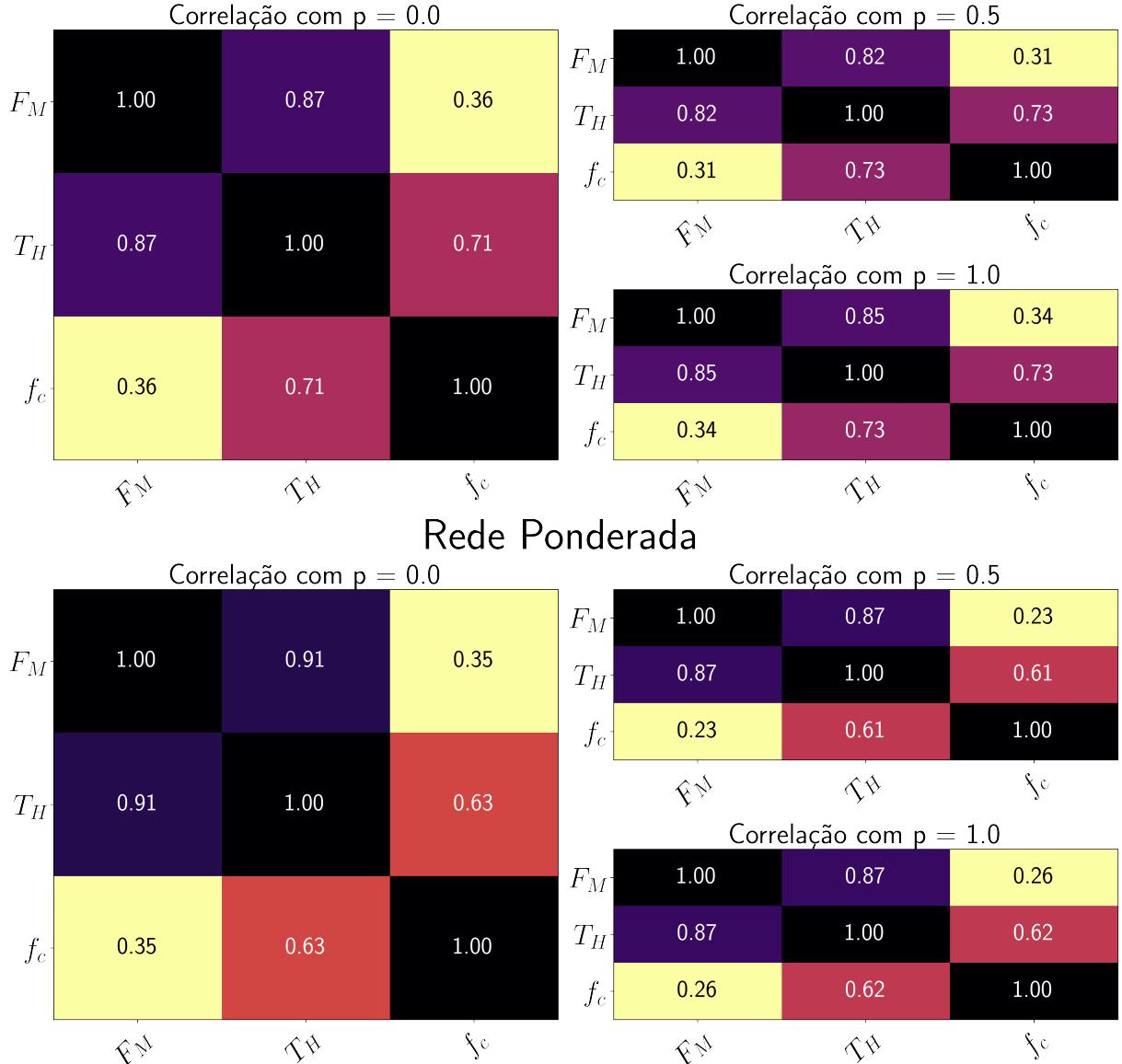


Fonte: Elaborado pelo autor.

Nota: Comparação de cada métrica, para os dois primeiros gráficos foi utilizado critério da integral enquanto na Fração de Infectados foi avaliado o valor de f quando a curva toca o 0 e o último mostra o valor de f_c para as curvas da fração de infectados. Comparando as melhores de cada métrica e a média contra a estratégia aleatória (CR) e idade (CI). Para $p = 0.0$, a melhor estratégia na média foi o PR com pesos nas arestas (WPR) e a versão altruísta da utilidade, A-UT, foi a melhor em fração de mortos e tempo hospitalizado, enquanto que em fração de infectados foi a Centralidade Laplaciana com ponderação nas arestas (WCL). No caso $p = 1.0$ a melhor na média foi o PR e a melhor na fração de infectados enquanto que a melhor na Fração de Mortos foi a CGR5, e no tempo total hospitalizado foi o altruísta da Centralidade de Gravidade com expoente 5, A-CGR5.

Embora os gráficos forneçam uma indicação preliminar da diferença entre a melhor estratégia em termos de uma determinada métrica de desempenho e do desempenho médio, é fundamental realizar uma avaliação para determinar se existe uma diferença significativa entre os *rankings* para as diferentes métricas. Nesse caso, foi feita a correlação de Spearman entre os *rankings* das métricas. A Figura 30 mostra o valor da correlação para diferentes valores de p para o caso não ponderado e ponderado. É possível ver que as duas métricas mais distintas entre si são F_M e f_c mostrando que a métrica que foca em diminuir a mortalidade não necessariamente previne o espalhamento da doença, enquanto que há uma alta correlação entre F_M e T_H , pois a única forma no modelo de morrer é sendo hospitalizado e ao diminuir o tempo na hospitalização acarreta numa diminuição de mortos. Todavia, que com o incremento de p , a matriz de correlação não se alterou significantemente tanto para o caso não ponderado como para o caso ponderado.

Figura 30 – Correlação entre o ranqueamento entre cada métrica
Rede Não Ponderada



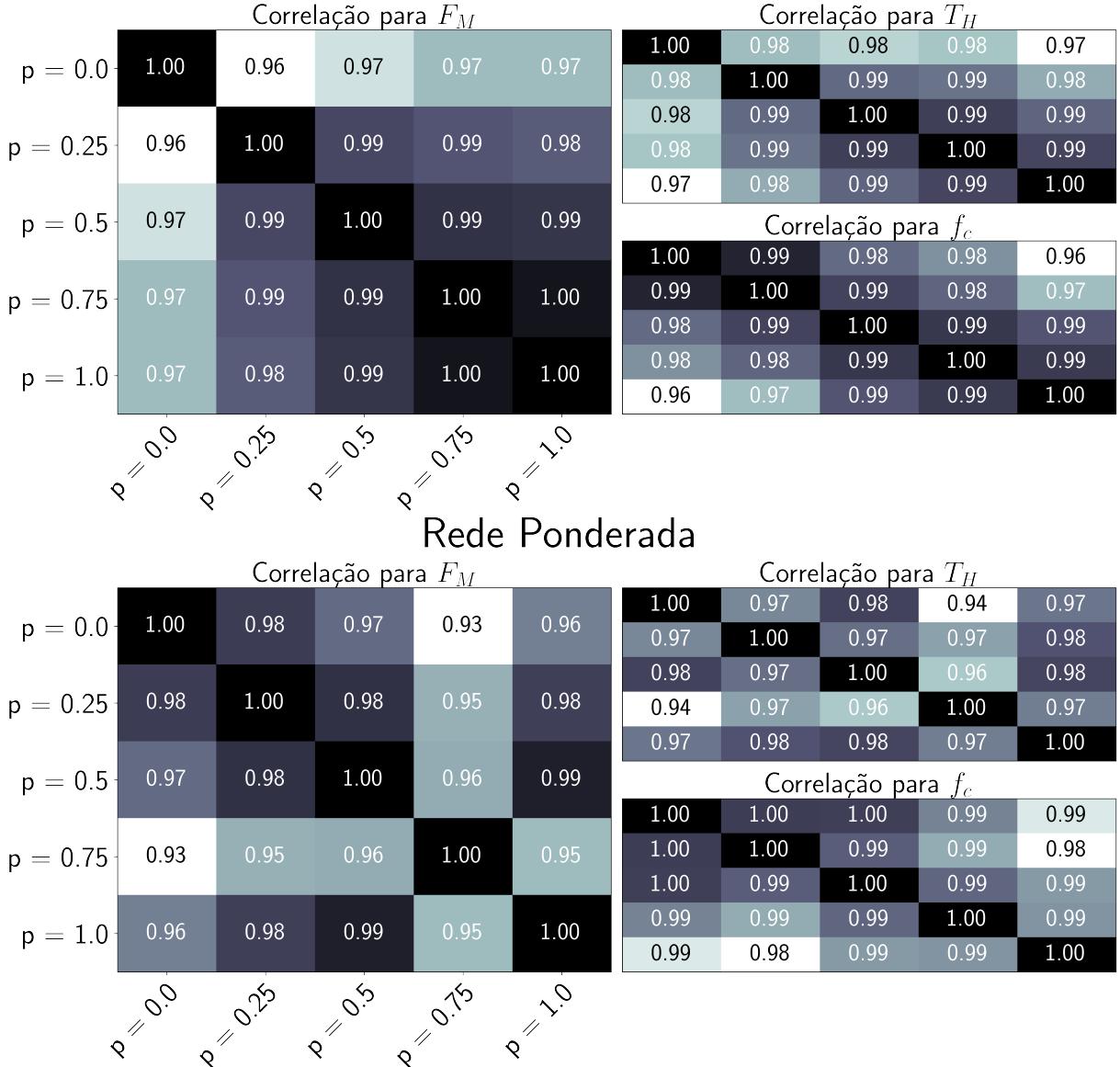
Fonte: Elaborado pelo autor.

Nota: A correlação entre o ranqueamento das métricas mostra que f_c e F_M são as mais diferentes e F_M e T_H são as mais próximas. Ademais, com o aumento de p não houve uma alteração significante da matriz.

Como mostrado pela Figura 30, as relações entre métricas mudam com o incremento de p , assim o próprio *ranking* das centralidades muda com a variação deste parâmetro. A Figura 31 mostra a correlação de Spearman do *ranking* das centralidades para cada par de valores de p , considerando rede não ponderada e ponderada. É possível ver que para o caso não ponderado no geral existe uma alta correlação entre os pares, tanto para $p = 0.0$ quanto para $p = 1.00$. Isso mostra que o posicionamento das estratégias é bastante estável com o aumento do agrupamento tanto no caso ponderado quanto não ponderado. Isso é bastante útil pois não foi medido o agrupamento da rede real de contatos entre pessoas e com essa estabilidade não se torna tão

necessário estimar esse valor.

Figura 31 – Correlação entre o ranqueamento de cada métrica para cada agrupamento
Rede Não Ponderada



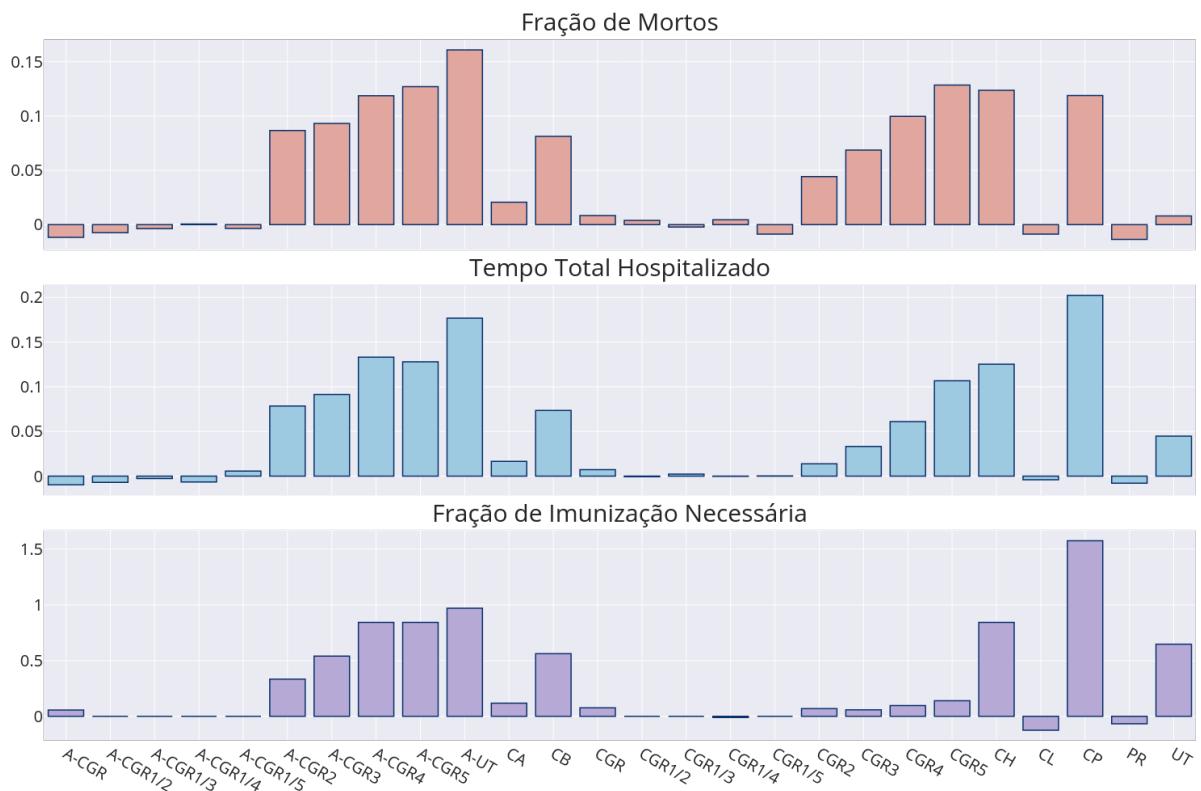
Fonte: Elaborado pelo autor.

Nota: Para cada métrica foi feita uma correlação de Spearman entre os 5 valores de p tanto para ponderação nas arestas quanto sem ponderação. Em ambos os casos mostra-se que a classificação se altera pouco o que é útil pois estimar o agrupamento de uma rede de contatos real é custoso.

Anteriormente, foram apresentadas as centralidades e suas versões com ponderação nas arestas. No entanto, a utilização da forma ponderada acarreta um custo computacional bastante elevado. Por isso, foi avaliado o desempenho da métrica ponderada em comparação com sua versão não ponderada, considerando a diferença relativa entre os valores das métricas com e sem ponderação $Z_w = \frac{M^w - M}{M}$ na qual M^w é o valor da métrica considerando ponderação nas arestas e M o valor sem ponderação nas arestas. Se $Z_w > 0$ significa que a métrica com

ponderação foi pior, pois é desejado uma minimização das métricas, já se $Z_w < 0$ quer dizer que houve melhora ao adicionar a ponderação. A Figura 32 superior mostra como ficou a comparação para o caso de $p = 0.0$ e pode se perceber que várias métricas não foram bem com o incremento da ponderação nas arestas, principalmente a Utilidade que para f_c teve uma perca de 100% e a CP que teve 150%. Algumas conseguiram se beneficiar, mas no geral a maioria teve uma queda de desempenho ao se considerar os pesos nas arestas. Isso mostra em conjunto com a Figura 33 superior a diferença relativa de tempo computacional, calculada da mesma forma que Z_w , com estratégias que demoram 5 vezes mais e não apresentam um desempenho que compensa o tempo adicional, outrossim na Figura 33 inferior mostra o tempo computacional de cada métrica, algumas como a CGR já tem um custo muito alto e ainda piora com a ponderação.

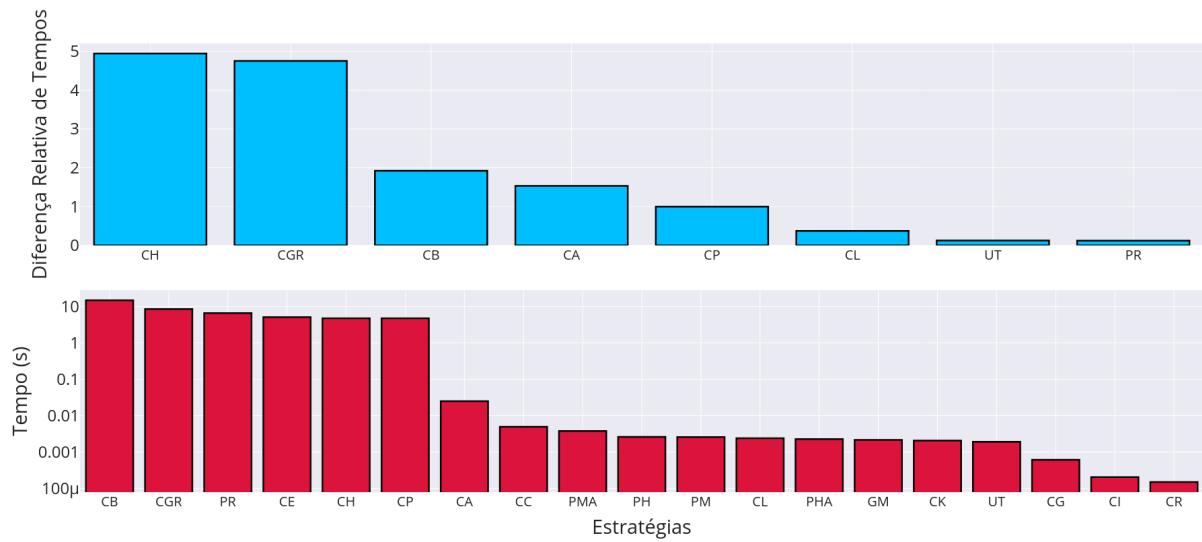
Figura 32 – Diferença relativa da centralidade com e sem peso nas arestas



Fonte: Elaborado pelo autor.

Nota: Diferença relativa entre ponderação nas arestas: em todas as métricas a diferença foi irrelevante ou houve uma piora considerável, não utilizá-las reduz o custo computacional do cálculo das centralidades.

Figura 33 – Diferença relativa da centralidade com e sem peso nas arestas

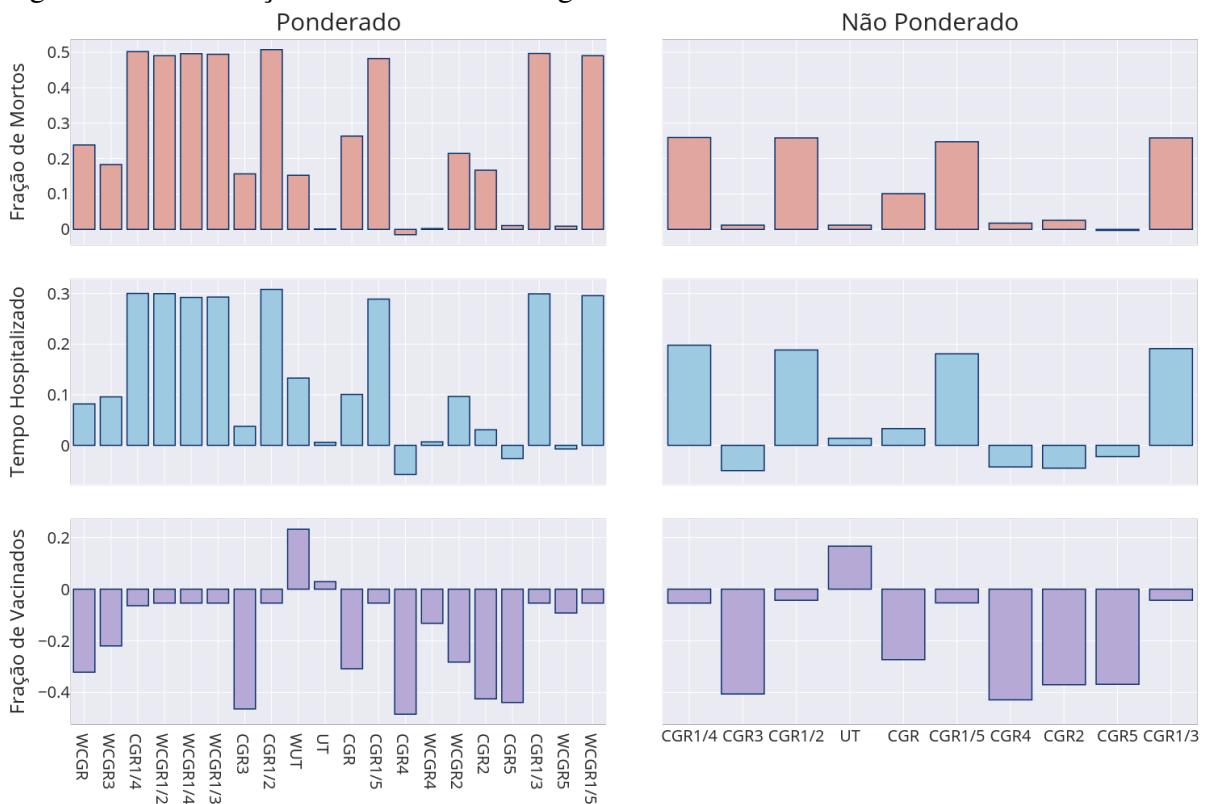


Fonte: Elaborado pelo autor.

Nota: A figura superior mostra que estratégias com tempos computacionais na qual foram obtidos de uma média de 5 simulações. Algumas chegam a 5 vezes maior não oferecem desempenho que justifique o tempo extra; a figura inferior exibe o logaritmo do tempo computacional de cada métrica, onde algumas, como a CGR, já possuem alto custo, agravado com a ponderação.

Assim como foi feita a comparação entre a centralidade na sua versão ponderada com arestas e sem ponderação será feito uma comparação entre a abordagem altruísta e individualista. A comparação entre Individualista e Altruísta foi pelo erro relativo $Z_A = \frac{M^I - M^A}{M^A}$, na qual M^A é o valor da métrica com abordagem altruísta e M^I o valor da métrica com abordagem individualista. A Figura 34 indica que para Fração de Mortos e Tempo Hospitalizado houve uma piora ao utilizar a abordagem individualista, entretanto para Fração de vacinados ocorreu o contrário. Além disso, nota-se que as melhorias em Z_A foram mais significativas nas centralidades que consideram o peso das arestas, em comparação com aquelas que não o levam em conta.

Figura 34 – Diferença relativa entre abordagem Individualista e Altruísta



Fonte: Elaborado pelo autor.

Nota: As estratégias individualistas tiveram pior desempenho no tempo de hospitalização e na fração de mortos, mas exigiram uma fração menor de vacinados que as altruístas.

As Tabelas 6 e 7 expõem as melhores estratégias para cada valor de p , métrica e ponderação na rede. No geral, as estratégias que são apresentadas são aquelas que têm característica local, ou seja, os vizinhos mais próximos têm um maior peso e também o *PageRank* na qual foi mostrado a sua utilidade em prever difusão (Masuda *et al.*, 2017) e que parece se mostrar como a melhor estratégia para prevenir a doença a partir dos dados coletados e do modelo proposto. Todos os resultados são encontrados no Apêndice A.

Por fim, foram identificadas as fronteiras de Pareto para cada valor de p , que representam as combinações de estratégias que oferecem o melhor compromisso entre as diferentes métricas de desempenho avaliadas. A fronteira de Pareto é uma representação gráfica que ilustra essa relação de trade-off: ao longo da curva, cada ponto representa uma configuração de estratégias onde melhorar uma métrica implica necessariamente piorar outra. Assim, as fronteiras de Pareto mostradas indicam as soluções eficientes para cada valor de p . Para o caso não ponderado a fronteira de Pareto é composta pelas seguintes estratégias:

- $p = 0.00$: A-UT, A-CGR, A-CGR2, A-CGR3, A-CGR4, A-CGR5, CB, UT, CG, CGR4, PR, PM;

- $p = 0.25$: A-UT, A-CGR, A-CGR2, A-CGR3, A-CGR4, A-CGR5, CB, UT, CG, CGR, CGR2, CGR3, CGR4, CGR5, PR;
- $p = 0.50$: A-CGR, A-CGR2, A-CGR3, A-CGR4, A-CGR5, CB, CG, CGR, CGR2, CGR3, CGR4, CGR5, PR;
- $p = 0.75$: A-UT, A-CGR, A-CGR2, A-CGR3, A-CGR4, A-CGR5, CB, CG, CGR3, CGR4, CGR5, PR, PH, PM;
- $p = 1.00$: A-CGR, A-CGR2, A-CGR3, A-CGR4, A-CGR5, CGR2, CGR3, CGR4, CGR5, PR.

Por outro lado, para o caso ponderado, a fronteira de Pareto é composta pelas seguintes estratégias:

- $p = 0.00$: A-UT, A-CGR, A-CGR2, A-CGR3, A-CGR4, A-CGR5, CB, UT, CG, PR, WCL, WPR
- $p = 0.25$: A-UT, A-CGR, A-CGR2, A-CGR3, A-CGR4, A-CGR5, CB, UT, CG, CGR4, CGR5, PR
- $p = 0.50$: A-UT, A-CGR, A-CGR2, A-CGR3, A-CGR4, A-CGR5, CB, UT, CG, CGR5, PR, WUT, WCL, WPR
- $p = 0.75$: A-UT, A-CGR, A-CGR2, A-CGR3, A-CGR4, A-CGR5, UT, CG, CGR4, CGR5, WCL, WPR
- $p = 1.00$: A-UT, A-CGR, A-CGR2, A-CGR3, A-CGR4, A-CGR5, UT, CG, CGR4, CGR5, PR, WPR

É notável a presença de estratégias que levam em consideração que utilizam tanto informação da estrutura da rede como também os pesos nos nós e ou nas arestas, além da predominância do altruísmo e da pouca sensibilidade perante o aumento de p , algo já discutido.

Tabela 6 – Melhores estratégias para cada valor de p na rede não ponderada. Há uma predominância de estratégias locais, ou seja, nós próximos do nó avaliado têm maior importância

Probabilidade	Fração de Mortos	Tempo Hospitalizado	Fração de Vacinados	Média
$p = 0.00$	CGR4	UT	PR	PR
$p = 0.25$	CGR4	PR	CG	PR
$p = 0.50$	CGR4	A-CGR5	PR	PR
$p = 0.75$	CGR4	A-CGR5	PR	PR
$p = 1.00$	CGR4	PR	PR	PR

Fonte: Elaborado pelo autor.

Tabela 7 – Melhores estratégias para cada valor de p na rede ponderada. Há uma predominância da WPR e PR em relação as outras centralidades e ainda sim como no exemplo passado uma predominância de estratégias locais.

Probabilidade	Fração de Mortos	Tempo Hospitalizado	Fração de Vacinados	Média
$p = 0.00$	A-UT	UT	WPR	WPR
$p = 0.25$	CGR5	CB	PR	PR
$p = 0.50$	WUT	CB	WPR	WPR
$p = 0.75$	CGR5	A-CGR5	WPR	WPR
$p = 1.00$	CGR5	A-CGR5	WPR	PR

Fonte: Elaborado pelo autor.

5 CONCLUSÕES E TRABALHOS FUTUROS

Neste trabalho, foram investigadas diferentes métricas de centralidade e quais delas são mais efetivas em identificar indivíduos chave na propagação do COVID-19 e nas hospitalizações e mortes decorrentes da COVID-19. Ao estudar o POLYMOD, realizamos uma análise exploratória que nos permitiu coletar dados fundamentais para a construção de um modelo de redes. Embora a pesquisa original não tenha sido direcionada especificamente para esse fim, propusemos um modelo de formação de rede a partir de uma sequência de graus, considerando as conexões entre indivíduos de diferentes faixas etárias e a distribuição etária na rede. Através dessa abordagem, foi possível obter dados sobre o tempo de contato entre as diferentes faixas etárias, permitindo ponderar as ligações entre os indivíduos. Como resultado, desenvolvemos um modelo de propagação da COVID-19 mais complexo em comparação aos usualmente presentes na literatura de redes. Esse modelo considera as interferências na rede e as alterações resultantes da vacinação dos indivíduos, oferecendo uma visão mais detalhada sobre os impactos da vacinação na dinâmica de propagação da doença.

A partir do modelo utilizado, da variação do agrupamento, da construção da rede, da ponderação ou não das arestas foi mostrado como o modelo se comporta, como ele é sensível em relação ao incremento do agrupamento e que ele é pouco alterado com a ponderação nas arestas. Com as métricas a disposição foi mostrado que centralidades que consideraram peso nos nós e a estrutura da rede foram muito efetivas para vacinação contra COVID na média e que como foi mostrado o *PageRank* ponderado apresenta uma grande efetividade contra a propagação, essas estratégias conseguiam reduzir a mortalidade em mais de 60%, o tempo total de hospitalização em 66% e foi necessário vacinar menos de 31% da rede com a melhor estratégia para inibir a proliferação do vírus. Ademais foi mostrado que utilizar peso nas arestas não teve ganho tão significativo perante o gasto computacional adicional ocorrendo muita vezes pioras nas métricas. Enquanto que uma abordagem altruísta dos valores se mostrou uma estratégia boa para reduzir a mortalidade e hospitalização mas a individualista foi melhor para mitigar o espalhamento da doença. Outrossim foi listado quais foram as melhores métricas, as métricas melhores em média e a fronteira de Pareto para os dados. Por fim, a classificação das melhores métricas mostrou ser bastante estável com o incremento do agrupamento da rede, o que sugere não ser tão necessário a estimação desse valor em uma rede de contágio real. Os resultados mostram que a idade utilizada por maior parte dos países não foi uma estratégia tão eficiente nos três parâmetros após a vacinação de 15% da população, a melhor escolha seria aquelas métricas que estão na fronteira

de Pareto.

Entretanto, essa pesquisa ainda apresenta limitações:

1. O POLYMOD apesar de obter vários dados não se sabe a abrangência total e até que ponto se pode aplicar para diferentes países e culturas;
2. O Modelo de Configuração Estratificado não consegue aumentar o agrupamento para nós com grau maiores, não se sabe se é uma limitação do algoritmo, dos dados, ou é um comportamento padrão;
3. O modelo epidemiológico que foi proposto abre margem para o estudo de superlotação de hospitais que foi uma característica importante na pandemia e não explorada neste trabalho.

Essas são algumas limitações do presente trabalho, que podem servir de propostas para futuras pesquisas nessa área.

REFERÊNCIAS

- ANDRADE, R. L. de; RÊGO, L. C. p-means centrality. **Communications in Nonlinear Science and Numerical Simulation**, United Kingdom, v. 68, p. 41–55, March 2019. ISSN 1878-7274.
- ANDRADE, R. L. de; RÊGO, L. C.; SILVA, T. L. Coelho da; MACÊDO, J. A. F. de; SILVA, W. C. Energy disruptive centrality with an application to criminal network. **Communications in Nonlinear Science and Numerical Simulation**, United Kingdom, v. 99, p. 105834, August 2021. ISSN 1878-7274.
- ARRUDA, G. F. de; RODRIGUES, F. A.; MORENO, Y. Fundamentals of spreading processes in single and multilayer complex networks. **Physics Reports**, Netherlands, v. 756, p. 159, October 2018. ISSN 0370-1573.
- BARABÁSI, A.-L. **Network Science**. [S. l.]: Cambridge University Press, 2016.
- BEAGLEHOLE, R. *et al.* **Epidemiologia Básica**. 2. ed. [S. l.]: Livraria Santos Editora, 2010.
- BERNOULLI, D.; BLOWER, S. An attempt at a new analysis of the mortality caused by smallpox and of the advantages of inoculation to prevent it. **Reviews in Medical Virology**, United Kingdom, v. 14, n. 5, p. 275–288, 2004. ISSN 1099-1654.
- BLOCK, P. *et al.* Social network-based distancing strategies to flatten the COVID-19 curve in a post-lockdown world. **Nature Human Behaviour**, United Kingdom, v. 4, n. 6, p. 588–596, June 2020. ISSN 2397-3374.
- BOGUÑÁ, M.; PASTOR-SATORRAS, R. Epidemic spreading in correlated complex networks. **Physical Review E**, United States, v. 66, n. 4, October 2002. ISSN 2470-0053.
- BOKÁNYI, E. *et al.* The anatomy of a population-scale social network. **Scientific Reports**, United Kingdom, v. 13, n. 1, June 2023. ISSN 2045-2322.
- BRANKE, J. *et al.* **Multiobjective Optimization**: interactive and Evolutionary Approaches. 1. ed. [S. l.]: Springer, 2008. (Lecture Notes in Computer Science 5252).
- BRASIL. Ministério da Saúde. **SRAG 2020**: - Banco de Dados de Síndrome Respiratória Aguda Grave - incluindo dados da COVID-19. Brasília, DF: Ministério da Saúde, 2020. Disponível em: <https://opendatasus.saude.gov.br/dataset/srag-2020>. Acesso em: 12 maio 2023.
- BRIN, S.; PAGE, L. The anatomy of a large-scale hypertextual web search engine. **Computer Networks**, Netherlands, v. 30, p. 107–117, 1998. ISSN 1389-1286.
- BUSTAMANTE-CASTAÑEDA, F. *et al.* Epidemic model on a network: analysis and applications to COVID-19. **Physica A: Statistical Mechanics and its Applications**, Netherlands, v. 564, p. 125520, February 2021. ISSN 1873-2119.
- CADEDU *et al.* Planning and organization of the covid-19 vaccination campaign: An overview of eight european countries. **Vaccines**, Switzerland, v. 10, n. 10, p. 1631, September 2022. ISSN 2076-393X.
- CARMI *et al.* A model of internet topology using k -shell decomposition. **Proceedings of the National Academy of Sciences**, United States, v. 104, n. 27, p. 1115011154, July 2007. ISSN 0027-8424.

- CHATTERJEE, A. **Studies on the structure and dynamics of urban bus networks in indian cities.** [S.l.]: [S. n.], 2015.
- CHEN *et al.* Age-stratified covid-19 spread analysis and vaccination: a multitype random network approach. **IEEE Transactions on Network Science and Engineering**, United States, v. 8, n. 2, p. 1862–1872, 2021. ISSN 2334-329X.
- CHIERICETTI *et al.* Rumor spreading in social networks. **Theoretical Computer Science**, Netherlands, v. 412, n. 24, p. 26022610, May 2011. ISSN 1879-2294.
- COSTA, L. da F. *et al.* Characterization of complex networks: a survey of measurements. **Advances in Physics**, United Kingdom, v. 56, n. 1, p. 167–242, January 2007. ISSN 1460-6976.
- CRAIG, B. R. *et al.* Improving epidemic modeling with networks. **Economic Commentary (Federal Reserve Bank of Cleveland)**, United States, p. 1–8, September 2020. ISSN 2163-3738.
- CRUZ, M. P. *et al.* COVID-19, a worldwide public health emergency. **Revista Clínica Española (English Edition)**, Brazil, v. 221, n. 1, p. 55–61, January 2021. ISSN 0103-4014.
- CSÁRDI *et al.* **igraph.** [S. l.]: Zenodo, 2024.
- DIETZ, K.; HESTERBEEK, J. Daniel Bernoullis epidemiological model revisited. **Mathematical Biosciences**, United States, v. 180, n. 1, p. 1–21, 2002. ISSN 0025-5564.
- DOLL, R.; HILL, A. B. Smoking and carcinoma of the lung. **BMJ**, United Kingdom, v. 2, n. 4682, p. 739–748, September 1950. ISSN 2044-6055.
- DOOSTMOHAMMADIAN, M. *et al.* Centrality-based epidemic control in complex social networks. **Social Network Analysis and Mining**, Austria, v. 10, n. 1, May 2020. ISSN 1869-5469.
- EAMES, K. T.; KEELING, M. J. Contact tracing and disease control. **Proceedings of the Royal Society of London. Series B: biological Sciences**, United Kingdom, v. 270, n. 1533, p. 2565–2571, 2003. ISSN 2053-9193.
- 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**, China, v. 5, p. 293–308, 2020. ISSN 2468-2152.
- ERDÖS, P.; RÉNYI, A. On random graphs i. **Publicationes Mathematicae Debrecen**, Hungary, v. 6, p. 290–297, 1959. ISSN 0033-3883.
- EULER, L. Solutio problematis ad geometriam situs pertinentis. **Commentarii Academiae Scientiarum Imperialis Petropolitanae**, Russian Federation, v. 8, p. 128–140, 1736. ISSN 2658-5065.
- FENG *et al.* Epidemiological models with non-exponentially distributed disease stages and applications to disease control. **Bulletin of Mathematical Biology**, United States, v. 69, n. 5, p. 1511–1536, 2007. ISSN 1522-9602.
- FILHO, R. *et al.* Optimization of covid-19 vaccination and the role of individuals with a high number of contacts: A model based approach. **PLOS ONE**, Public Library of Science (PLoS), United States, v. 17, n. 3, p. e0262433, mar. 2022. ISSN 1932-6203.

- GAN *et al.* Propagation of computer virus both across the internet and external computers: A complex-network approach. **Communications in Nonlinear Science and Numerical Simulation**, United Kingdom, v. 19, n. 8, p. 27852792, August 2014. ISSN 1878-7274.
- GILLESPIE, D. T. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. **Journal of Computational Physics**, Netherlands, v. 22, n. 4, p. 403–434, December 1976. ISSN 0021-9991.
- GOLDBARG, M. **Grafos:** conceitos, algoritmos e aplicações. [S. l.: s. n.], 2012.
- GONÇALVES, C. M. M.; RÊGO, L. C.; FIERENS, P. I. Estratégias de vacinação para COVID-19 usando redes complexas. In: ANAIS DO SIMPÓSIO BRASILEIRO DE PESQUISA OPERACIONAL, 55.,2023, São José dos Campos, **Anais eletrônicos** ..., Galoá, 2023. Disponível em: <https://proceedings.science/sbpo-2023/trabalhos/estrategias-de-vacinacao-para-covid-19-usando-redes-complexas?lang=pt-br>. Acesso em 25 de Janeiro de 2025.
- GONG *et al.* An efficient immunization strategy for community networks. **PLoS ONE**, United States, v. 8, n. 12, p. e83489, December 2013. ISSN 1932-6203.
- GONZALEZ, T.; RUBIA, M. A. de la; HINCZ, K. P.; COMAS-LOPEZ, M.; SUBIRATS, L.; FORT, S.; SACHA, G. M. Influence of covid-19 confinement on students performance in higher education. **PLoS ONE**, United States, v. 15, n. 10, p. 1–23, 10 2020. ISSN 1932-6203.
- HAAS, E. J. *et al.* Impact and effectiveness of mRNA BNT162b2 vaccine against SARS-CoV-2 infections and COVID-19 cases, hospitalisations, and deaths following a nationwide vaccination campaign in israel: an observational study using national surveillance data. **The Lancet**, Netherlands, v. 397, n. 10287, p. 1819–1829, May 2021. ISSN 0140-6736.
- HAGBERG, A.; SWART, P.; CHULT, D. S. **Exploring network structure, dynamics, and function using NetworkX**. [S. l.], 2008.
- HARRIS, C. R. *et al.* Array programming with NumPy. **Nature**, United Kingdom, v. 585, n. 7825, p. 357–362, September 2020. ISSN 1476-4687.
- HASSINE, I. H. Covid19 vaccines and variants of concern: a review. **Reviews in Medical Virology**, United Kingdom, v. 32, n. 4, November 2021. ISSN 1099-1654.
- HUH *et al.* Covid-19 vaccination campaign trends and challenges in select asian countries. **Asian Journal of Political Science**, United States, v. 29, n. 3, p. 274300, September 2021. ISSN 1559-0968.
- HUNTER, J. D. Matplotlib: a 2d graphics environment. **Computing in Science & Engineering**, United States, v. 9, n. 3, p. 90–95, 2007. ISSN 1521-9615.
- IBGE. **Mapas urbanos estatísticos [Censo 2022], por unidades da Federação**. Rio de Janeiro: IBGE, 2023. Disponível em: <https://www.ibge.gov.br/estatisticas/sociais/trabalho/22827-censo-demografico-2022.html>. Acesso em: 19 jul. 2024.
- IRAWAN, A.; ALAMSYAH, H. The COVID-19's economic crisis and its solutions: a literature review. **ETIKONOMI**, Indonesia, v. 20, n. 1, p. 77–92, February 2021. ISSN 2461-0771.
- JACKSON, M. O.; WOLINSKY, A. A strategic model of social and economic networks. **Journal of Economic Theory**, Belgium, v. 71, n. 1, p. 4474, October 1996. ISSN 1095-7235.

- 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**, International, v. 99, p. 266–268, October 2020. ISSN 1201-9712.
- JUNG, J. Preparing for the coronavirus disease (covid-19) vaccination: Evidence, plans, and implications. **Journal of Korean Medical Science**, Korea, Republic of, v. 36, n. 7, 2021. ISSN 1598-6357.
- JÚNIOR, F. P. de P. *et al.* **Modelagem matemática e epidemiologia com aplicações**. [S. l.: s. n.], 2021.
- KERMACK, W. O.; MCKENDRICK, A. G. A contribution to the mathematical theory of epidemics. **Proceedings of the Royal Society of London. Series A**, United Kingdom, v. 115, n. 772, p. 700–721, 1927. ISSN 2053-9169.
- KIRKCALDY, R. D. *et al.* COVID-19 and postinfection immunity. **JAMA**, United States, v. 323, n. 22, p. 2245, June 2020. ISSN 1538-3598.
- KITSAK *et al.* Identification of influential spreaders in complex networks. **Nature Physics**, United Kingdom, v. 6, n. 11, p. 888–893, 2010. ISSN 1745-2481.
- KLISE *et al.* Prioritizing vaccination based on analysis of community networks. **Applied Network Science**, Switzerland, v. 7, n. 1, p. 80, 2022. ISSN 2364-8228.
- LEE, C.; WILKINSON, D. J. A review of stochastic block models and extensions for graph clustering. **Applied Network Science**, Switzerland, v. 4, n. 1, December 2019. ISSN 2364-8228.
- LEYDESDORFF, L.; RAFOLS, I. Indicators of the interdisciplinarity of journals: Diversity, centrality, and citations. **Journal of Informetrics**, Netherlands, v. 5, n. 1, p. 87–100, January 2011. ISSN 1751-1577.
- LI, C.-L.; CHENG, C.-Y.; LI, C.-H. Global dynamics of two-strain epidemic model with single-strain vaccination in complex networks. **Nonlinear Analysis: real World Applications**, United Kingdom, v. 69, p. 103738, February 2023. ISSN 1468-1218.
- LIU, K.; LOU, Y. Optimizing COVID-19 vaccination programs during vaccine shortages. **Infectious Disease Modelling**, China, v. 7, n. 1, p. 286–298, March 2022. ISSN 2468-2152.
- LOYAL, J. D.; CHEN, Y. Statistical network analysis: a review with applications to the coronavirus disease 2019 pandemic. **International Statistical Review**, United Kingdom, v. 88, n. 2, p. 419–440, July 2020. ISSN 1751-5823.
- MAHESHWARI, P.; ALBERT, R. Network model and analysis of the spread of covid-19 with social distancing. **Applied Network Science**, Switzerland, v. 5, n. 1, December 2020. ISSN 2364-8228.
- MALLICK *et al.* **Hesitancy, Awareness, and Vaccination: A Computational Analysis on Complex Networks**. [S. l.]: Wiley, 2023. 101113 p.
- MANZO, G.; RIJT, A. van de. Halting SARS-CoV-2 by targeting high-contact individuals. **Journal of Artificial Societies and Social Simulation**, United Kingdom, v. 23, n. 4, 2020. ISSN 1460-7425.

- MASUDA *et al.* Random walks and diffusion on networks. **Physics Reports**, Netherlands, v. 716717, p. 158, November 2017. ISSN 0370-1573.
- MATSUMOTO *et al.* Mersenne twister: a 623-dimensionally equidistributed uniform pseudo-random number generator. **ACM Trans. Model. Comput. Simul.**, United States, v. 8, n. 1, p. 330, January 1998. ISSN 1049-3301.
- MILLER, J. C.; HYMAN, J. M. Effective vaccination strategies for realistic social networks. **Physica A : statistical Mechanics and its Applications**, Netherlands, v. 386, n. 2, p. 780–785, 2007. ISSN 1873-2119.
- MORENO *et al.* Dynamics of rumor spreading in complex networks. **Physical Review E**, United States, v. 69, n. 6, June 2004. ISSN 2470-0053.
- MOSSONG, J. *et al.* Social contacts and mixing patterns relevant to the spread of infectious diseases. **PLOS Medicine**, United States, v. 5, p. 1–1, 03 2008. ISSN 1549-1676.
- NEWMAN, M. **Networks**. 2. ed. [S. l.]: Oxford University Press, 2018.
- NIU *et al.* Modeling the COVID-19 pandemic using an seihr model with human migration. **IEEE Access**, United States, v. 8, p. 195503195514, 2020. ISSN 2169-3536.
- NOWZARI *et al.* Analysis and control of epidemics: A survey of spreading processes on complex networks. **IEEE Control Systems Magazine**, United States, v. 36, n. 1, p. 26–46, 2016. ISSN 2374-9385.
- PASTOR-SATORRAS, R.; VESPIGNANI, A. Immunization of complex networks. **Physical Review E**, United States, v. 65, n. 3, p. 036104, 2002. ISSN 2470-0053.
- PEI, H. *et al.* Impact of contact rate on epidemic spreading in complex networks. **The European Physical Journal B**, France, v. 96, n. 4, April 2023. ISSN 1434-6036.
- PELLIS, L. *et al.* Eight challenges for network epidemic models. **Epidemics**, Netherlands, v. 10, p. 58–62, March 2015. ISSN 1755-4365.
- PEREIRA, M. D. *et al.* A pandemia de COVID-19, o isolamento social, consequências na saúde mental e estratégias de enfrentamento: uma revisão integrativa. **Research, Society and Development**, Research, Society and Development, v. 9, n. 7, p. e652974548, June 2020.
- PETRIZZELLI, F. *et al.* Beyond covid-19 pandemic: Topology-aware optimization of vaccination strategy for minimizing virus spreading. **Computational and Structural Biotechnology Journal**, Sweden, v. 20, p. 2664–2671, 2022. ISSN 2001-0370.
- PSYCHIATRY of Pandemics: A Mental Health Response to Infection Outbreak. Springer International Publishing, 2019. ISBN 9783030153465. Disponível em: <http://dx.doi.org/10.1007/978-3-030-15346-5>.
- QI *et al.* Laplacian centrality: A new centrality measure for weighted networks. **Information Sciences**, Netherlands, v. 194, p. 240253, July 2012. ISSN 0020-0255.
- QUIROGA, C. de L.; SCHIMIT, P. H. T. A multi-city epidemiological model based on cellular automata and complex networks for the COVID-19. **Computational and Applied Mathematics**, Brazil, v. 42, n. 6, August 2023. ISSN 1807-0302.

- RÊGO, L. C.; SANTOS, A. M. dos. Co-authorship model with link strength. **European Journal of Operational Research**, Netherlands, v. 272, n. 2, p. 587594, January 2019. ISSN 1872-6860.
- ROSEN *et al.* Israels rapid rollout of vaccinations for covid-19. **Israel Journal of Health Policy Research**, United Kingdom, v. 10, n. 1, January 2021. ISSN 2045-4015.
- ROSSA, F. D. *et al.* A network model of italy shows that intermittent regional strategies can alleviate the covid-19 epidemic. **Nature Communications**, United Kingdom, v. 11, n. 1, October 2020. ISSN 2041-1723.
- SALATHÉ, M.; JONES, J. H. Dynamics and control of diseases in networks with community structure. **PLoS Computational Biology**, United States, v. 6, n. 4, p. e1000736, 2010. ISSN 1544-9173.
- SCABINI, L. F. *et al.* Social interaction layers in complex networks for the dynamical epidemic modeling of covid-19 in Brazil. **Physica A : statistical Mechanics and its Applications**, Netherlands, v. 564, p. 125498, 2021. ISSN 1873-2119.
- SILVA, D. H.; ANTENEODO, C.; FERREIRA, S. C. Epidemic outbreaks with adaptive prevention on complex networks. **Communications in Nonlinear Science and Numerical Simulation**, United Kingdom, v. 116, p. 106877, January 2023. ISSN 1878-7274.
- SINGHAL, T. A review of coronavirus disease-2019 (COVID-19). **The Indian Journal of Pediatrics**, India, v. 87, n. 4, p. 281–286, March 2020. ISSN 0973-7693.
- SZWARCFITER, J. **Teoria computacional de grafos: Os Algoritmos**. [S. l.]: ELSEVIER, 2018. ISBN 9788535288858.
- TETTEH *et al.* Network models to evaluate vaccine strategies towards herd immunity in covid-19. **Journal of Theoretical Biology**, Netherlands, v. 531, p. 110894, December 2021. ISSN 0022-5193.
- WANG *et al.* Coevolution spreading in complex networks. **Physics Reports**, Netherlands, v. 820, p. 151, August 2019. ISSN 0370-1573.
- WANG, C. *et al.* A novel coronavirus outbreak of global health concern. **The Lancet**, Netherlands, v. 395, n. 10223, p. 470–473, February 2020. ISSN 0140-6736.
- WATTS, D. J.; STROGATZ, S. H. Collective dynamics of small-world networks. **Nature**, United Kingdom, v. 393, n. 6684, p. 440442, June 1998. ISSN 1476-4687.
- 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**, China, v. 6, p. 324–342, 2021. ISSN 2468-2152.
- YANG *et al.* A study of the spreading scheme for viral marketing based on a complex network model. **Physica A: Statistical Mechanics and its Applications**, Netherlands, v. 389, n. 4, p. 859870, Feb. 2010. ISSN 1873-2119.
- YANG *et al.* Epidemics of computer viruses: A complex-network approach. **Applied Mathematics and Computation**, United States, v. 219, n. 16, p. 87058717, April 2013. ISSN 1873-5649.

ZHANG, Z.-K. *et al.* Dynamics of information diffusion and its applications on complex networks. **Physics Reports**, Netherlands, v. 651, p. 1–34, September 2016. ISSN 0370-1573.

APÊNDICE A – RESULTADOS PARA TODAS AS ESTRATÉGIAS

Tabela 8 – Tabela com todos os resultados para rede com arestas não ponderadas
 $p = 0.00$, em azul os menores valores e em vermelho os maiores.

Métricas	Fração de Mortos	Tempo Hospitalizado (dias)	Fração de Vacinados
A-UT	0.009	3648.921	0.42
A-CGR	0.012	4660.963	0.56
A-CGR1/2	0.014	5486.388	0.89
A-CGR1/3	0.014	5518.971	0.89
A-CGR1/4	0.014	5526.285	0.88
A-CGR1/5	0.014	5480.48	0.89
A-CGR2	0.011	4134.308	0.46
A-CGR3	0.01	3892.631	0.41
A-CGR4	0.01	3750.978	0.36
A-CGR5	0.009	3691.826	0.36
CB	0.01	3694.807	0.34
CP	0.01	3804.286	0.36
CC	0.013	4989.104	0.85
UT	0.009	3611.094	0.36
CE	0.012	4465.747	0.78
CA	0.01	3857.086	0.36
CG	0.009	3689.232	0.33
GM	0.012	4468.721	0.47
CGR	0.01	4151.916	0.77
CGR1/2	0.01	4322.697	0.93
CGR1/3	0.01	4344.113	0.93
CGR1/4	0.01	4330.174	0.93
CGR1/5	0.01	4344.356	0.94
CGR2	0.009	4127.597	0.73
CGR3	0.009	4017.729	0.69
CGR4	0.009	3912.571	0.63
CGR5	0.009	3813.301	0.57
CH	0.01	3790.328	0.36
CI	0.01	4318.644	0.91
CK	0.011	4095.622	0.39
CL	0.011	4188.341	0.34
PR	0.009	3642.825	0.31
PH	0.011	4158.684	0.33
PHA	0.016	5776.984	0.81
PM	0.011	4165.017	0.32
PMA	0.016	5786.542	0.82
CR	0.013	5234.572	0.82

Fonte: Elaborado pelo autor.

Tabela 9 – Tabela com todos os resultados para rede com arestas não ponderadas $p = 0.25$, em azul os menores valores e em vermelho os maiores.

Métricas	Fração de Mortos	Tempo Hospitalizado (dias)	Fração de Vacinados
A-UT	0.009	3752.509	0.44
A-CGR	0.012	4551.09	0.54
A-CGR1/2	0.014	5400.81	0.91
A-CGR1/3	0.014	5393.984	0.91
A-CGR1/4	0.014	5418.711	0.91
A-CGR1/5	0.014	5398.208	0.9
A-CGR2	0.011	4089.688	0.46
A-CGR3	0.01	3911.676	0.41
A-CGR4	0.01	3788.98	0.39
A-CGR5	0.01	3709.05	0.36
CB	0.01	3713.028	0.34
CP	0.01	3885.15	0.4
CC	0.015	5824.486	0.86
UT	0.009	3696.094	0.36
CE	0.011	4307.0	0.77
CA	0.01	3888.261	0.39
CG	0.01	3728.585	0.32
GM	0.012	4604.976	0.46
CGR	0.009	4087.133	0.77
CGR1/2	0.01	4277.773	0.94
CGR1/3	0.01	4254.297	0.94
CGR1/4	0.01	4274.417	0.94
CGR1/5	0.01	4281.586	0.94
CGR2	0.009	4041.585	0.71
CGR3	0.009	3969.95	0.67
CGR4	0.009	3857.801	0.58
CGR5	0.009	3785.124	0.52
CH	0.01	3874.235	0.39
CI	0.01	4276.754	0.91
CK	0.011	4073.086	0.4
CL	0.012	4293.609	0.36
PR	0.01	3690.279	0.32
PH	0.011	4228.272	0.33
PHA	0.016	5795.855	0.82
PM	0.012	4233.15	0.32
PMA	0.016	5776.343	0.8
CR	0.013	5137.396	0.82

Fonte: Elaborado pelo autor.

Tabela 10 – Tabela com todos os resultados para rede com arestas não ponderadas $p = 0.50$, em azul os menores valores e e em vermelho os maiores.

Métricas	Fração de Mortos	Tempo Hospitalizado (dias)	Fração de Vacinados
A-UT	0.009	3734.118	0.4
A-CGR	0.012	4574.054	0.54
A-CGR1/2	0.014	5363.338	0.92
A-CGR1/3	0.014	5359.573	0.92
A-CGR1/4	0.014	5354.557	0.89
A-CGR1/5	0.013	5363.209	0.9
A-CGR2	0.011	4121.688	0.46
A-CGR3	0.01	3877.933	0.39
A-CGR4	0.01	3769.114	0.36
A-CGR5	0.009	3687.969	0.35
CB	0.01	3734.788	0.33
CP	0.01	3880.418	0.42
CC	0.015	5816.934	0.86
UT	0.009	3715.138	0.36
CE	0.011	4428.814	0.78
CA	0.01	3909.133	0.42
CG	0.01	3699.483	0.32
GM	0.012	4559.273	0.44
CGR	0.009	4073.95	0.76
CGR1/2	0.01	4238.606	0.94
CGR1/3	0.01	4252.571	0.94
CGR1/4	0.01	4254.322	0.93
CGR1/5	0.01	4257.187	0.94
CGR2	0.009	4007.549	0.7
CGR3	0.009	3906.442	0.64
CGR4	0.009	3828.507	0.57
CGR5	0.009	3746.336	0.51
CH	0.01	3887.703	0.41
CI	0.01	4270.566	0.9
CK	0.011	4057.969	0.42
CL	0.011	4219.991	0.36
PR	0.01	3699.695	0.3
PH	0.011	4163.589	0.33
PHA	0.015	5705.304	0.81
PM	0.011	4160.892	0.34
PMA	0.015	5715.502	0.81
CR	0.013	5082.254	0.8

Fonte: Elaborado pelo autor.

Tabela 11 – Tabela com todos os resultados para rede com arestas não ponderadas $p = 0.75$, em azul os menores valores e e em vermelho os maiores.

Métricas	Fração de Mortos	Tempo Hospitalizado (dias)	Fração de Vacinados
A-UT	0.009	3682.116	0.4
A-CGR	0.011	4501.113	0.52
A-CGR1/2	0.013	5293.818	0.88
A-CGR1/3	0.013	5317.795	0.89
A-CGR1/4	0.013	5304.354	0.9
A-CGR1/5	0.013	5292.868	0.9
A-CGR2	0.01	4063.869	0.46
A-CGR3	0.01	3843.892	0.39
A-CGR4	0.009	3703.839	0.35
A-CGR5	0.009	3634.113	0.35
CB	0.009	3677.205	0.33
CP	0.01	3914.39	0.46
CC	0.015	5785.574	0.87
UT	0.009	3670.361	0.35
CE	0.011	4421.317	0.76
CA	0.01	4003.871	0.46
CG	0.009	3665.494	0.32
GM	0.012	4560.863	0.43
CGR	0.009	4088.88	0.75
CGR1/2	0.01	4237.772	0.93
CGR1/3	0.01	4239.523	0.93
CGR1/4	0.01	4221.207	0.93
CGR1/5	0.01	4237.295	0.93
CGR2	0.009	3998.6	0.69
CGR3	0.009	3896.531	0.62
CGR4	0.009	3786.979	0.56
CGR5	0.009	3704.953	0.5
CH	0.01	3894.006	0.43
CI	0.01	4214.7	0.9
CK	0.01	4041.844	0.43
CL	0.011	4195.887	0.34
PR	0.009	3649.881	0.29
PH	0.011	4138.824	0.31
PHA	0.015	5659.372	0.8
PM	0.011	4115.198	0.31
PMA	0.015	5674.675	0.78
CR	0.013	5041.591	0.79

Fonte: Elaborado pelo autor.

Tabela 12 – Tabela com todos os resultados para rede com arestas não ponderadas $p = 1.00$, em azul os menores valores e e em vermelho os maiores.

Métricas	Fração de Mortos	Tempo Hospitalizado (dias)	Fração de Vacinados
A-UT	0.009	3650.372	0.39
A-CGR	0.011	4414.776	0.56
A-CGR1/2	0.013	5214.517	0.88
A-CGR1/3	0.013	5224.52	0.89
A-CGR1/4	0.013	5227.647	0.88
A-CGR1/5	0.013	5218.44	0.89
A-CGR2	0.01	4041.925	0.48
A-CGR3	0.01	3819.671	0.43
A-CGR4	0.009	3660.669	0.35
A-CGR5	0.009	3601.401	0.31
CB	0.009	3642.685	0.34
CP	0.01	3911.712	0.5
CC	0.015	5759.512	0.87
UT	0.009	3612.155	0.34
CE	0.011	4318.053	0.73
CA	0.011	4098.075	0.52
CG	0.009	3621.08	0.31
GM	0.012	4507.041	0.43
CGR	0.009	4039.225	0.75
CGR1/2	0.01	4193.864	0.92
CGR1/3	0.01	4197.562	0.92
CGR1/4	0.01	4206.681	0.93
CGR1/5	0.01	4219.709	0.93
CGR2	0.009	3985.37	0.68
CGR3	0.009	3911.818	0.62
CGR4	0.009	3766.894	0.53
CGR5	0.009	3662.572	0.47
CH	0.01	3894.588	0.45
CI	0.01	4194.119	0.9
CK	0.01	4069.981	0.44
CL	0.011	4139.606	0.35
PR	0.009	3584.784	0.28
PH	0.011	4071.833	0.31
PHA	0.015	5581.888	0.8
PM	0.011	4067.04	0.32
PMA	0.015	5586.276	0.79
CR	0.013	5002.111	0.78

Fonte: Elaborado pelo autor.

Tabela 13 – Tabela com todos os resultados para rede com arestas ponderadas $p = 0.00$, em azul os menores valores e em vermelho os maiores.

Métricas	Fração de Mortos	Tempo Hospitalizado (dias)	Fração de Vacinados
A-UT	0.009	3424.777	0.35
A-CGR	0.013	4728.78	0.54
A-CGR1/2	0.015	5632.304	0.88
A-CGR1/3	0.015	5618.344	0.88
A-CGR1/4	0.015	5637.421	0.88
A-CGR1/5	0.015	5589.045	0.88
A-CGR2	0.011	4218.783	0.42
A-CGR3	0.011	4055.15	0.37
A-CGR4	0.009	3491.304	0.32
A-CGR5	0.009	3453.754	0.32
A-WUT	0.011	4378.506	0.69
A-WCGR	0.013	4694.657	0.57
A-WCGR1/2	0.015	5602.037	0.88
A-WCGR1/3	0.015	5605.862	0.88
A-WCGR1/4	0.015	5605.817	0.88
A-WCGR1/5	0.015	5611.971	0.88
A-WCGR2	0.013	4665.625	0.56
A-WCGR3	0.013	4659.857	0.57
A-WCGR4	0.011	4229.514	0.59
A-WCGR5	0.011	4226.683	0.59
CB	0.009	3442.372	0.32
CP	0.009	3534.845	0.33
CC	0.012	4643.295	0.85
UT	0.009	3408.482	0.34
CE	0.011	4183.84	0.72
CA	0.009	3588.969	0.34
CG	0.009	3425.236	0.29
GM	0.011	4217.292	0.45
CGR	0.009	4006.597	0.78
CGR1/2	0.009	4146.734	0.93
CGR1/3	0.009	4163.033	0.93
CGR1/4	0.009	4171.172	0.94
CGR1/5	0.009	4164.714	0.93
CGR2	0.009	3951.019	0.73
CGR3	0.009	3872.183	0.69
CGR4	0.009	3747.785	0.62
CGR5	0.009	3623.983	0.57

Métricas	Fração de Mortos	Tempo Hospitalizado (dias)	Fração de Vacinados
CH	0.009	3514.602	0.32
CI	0.009	4105.165	0.92
CK	0.01	3843.426	0.38
CL	0.011	3900.422	0.32
PR	0.009	3424.466	0.3
PH	0.01	3859.464	0.3
PHA	0.015	5435.107	0.8
PM	0.01	3872.367	0.3
PMA	0.015	5442.306	0.8
CR	0.013	4949.786	0.81
WCB	0.01	3894.029	0.5
WCP	0.01	4351.023	0.85
WUT	0.009	3657.961	0.56
WCA	0.01	3742.694	0.38
WCGR	0.009	4063.808	0.84
WCGR1/2	0.009	4148.56	0.93
WCGR1/3	0.009	4169.494	0.93
WCGR1/4	0.009	4173.168	0.93
WCGR1/5	0.009	4165.897	0.93
WCGR2	0.01	4087.971	0.78
WCGR3	0.01	4180.706	0.73
WCGR4	0.011	4190.044	0.68
WCGR5	0.011	4238.067	0.65
WCH	0.011	4199.002	0.59
WCL	0.01	3852.314	0.28
WPR	0.009	3415.234	0.28

Fonte: Elaborado pelo autor.

Tabela 14 – Tabela com todos os resultados para rede com arestas ponderadas $p = 0.25$, em azul os menores valores e em vermelho os maiores.

Métricas	Fração de Mortos	Tempo Hospitalizado (dias)	Fração de Vacinados
A-UT	0.009	3455.633	0.38
A-CGR	0.013	4657.919	0.51
A-CGR1/2	0.015	5572.02	0.89
A-CGR1/3	0.015	5521.751	0.9
A-CGR1/4	0.015	5555.186	0.91
A-CGR1/5	0.015	5536.02	0.9
A-CGR2	0.011	4220.512	0.41
A-CGR3	0.011	4071.62	0.37
A-CGR4	0.009	3464.23	0.33
A-CGR5	0.009	3426.166	0.35
A-WUT	0.011	4380.35	0.68
A-WCGR	0.012	4662.947	0.55
A-WCGR1/2	0.015	5543.06	0.91
A-WCGR1/3	0.015	5513.51	0.88
A-WCGR1/4	0.015	5526.69	0.91
A-WCGR1/5	0.015	5529.166	0.91
A-WCGR2	0.012	4646.729	0.57
A-WCGR3	0.013	4699.991	0.55
A-WCGR4	0.011	4106.059	0.58
A-WCGR5	0.011	4122.267	0.59
CB	0.009	3413.839	0.31
CP	0.009	3559.448	0.34
CC	0.014	5519.608	0.85
UT	0.009	3430.065	0.36
CE	0.01	3976.192	0.73
CA	0.009	3568.26	0.36
CG	0.009	3453.257	0.3
GM	0.012	4420.165	0.46
CGR	0.009	3913.884	0.77
CGR1/2	0.009	4096.159	0.94
CGR1/3	0.009	4091.956	0.94
CGR1/4	0.009	4082.347	0.95
CGR1/5	0.009	4093.142	0.93
CGR2	0.009	3852.63	0.73
CGR3	0.009	3765.439	0.66
CGR4	0.009	3638.568	0.6
CGR5	0.008	3564.607	0.53

Métricas	Fração de Mortos	Tempo Hospitalizado (dias)	Fração de Vacinados
CH	0.009	3552.338	0.34
CI	0.009	4055.7	0.92
CK	0.01	3771.326	0.38
CL	0.011	4053.154	0.32
PR	0.009	3419.579	0.29
PH	0.011	3996.453	0.32
PHA	0.015	5460.965	0.79
PM	0.011	3983.456	0.31
PMA	0.015	5483.671	0.8
CR	0.012	4803.213	0.81
WCB	0.01	3786.165	0.49
WCP	0.01	4275.297	0.83
WUT	0.009	3652.804	0.57
WCA	0.01	3700.837	0.39
WCGR	0.009	3943.281	0.81
WCGR1/2	0.009	4094.425	0.94
WCGR1/3	0.009	4061.437	0.93
WCGR1/4	0.009	4087.847	0.94
WCGR1/5	0.009	4081.64	0.94
WCGR2	0.009	3957.741	0.77
WCGR3	0.01	4043.713	0.72
WCGR4	0.01	4110.649	0.66
WCGR5	0.01	4122.584	0.64
WCH	0.01	4080.18	0.57
WCL	0.011	4019.485	0.31
WPR	0.009	3467.687	0.3

Fonte: Elaborado pelo autor.

Tabela 15 – Tabela com todos os resultados para rede com arestas ponderadas $p = 0.50$, em azul os menores valores e em vermelho os maiores.

Métricas	Fração de Mortos	Tempo Hospitalizado (dias)	Fração de Vacinados
A-UT	0.009	3440.745	0.38
A-CGR	0.012	4592.603	0.52
A-CGR1/2	0.015	5462.423	0.9
A-CGR1/3	0.015	5489.149	0.9
A-CGR1/4	0.015	5470.139	0.88
A-CGR1/5	0.015	5445.208	0.89
A-CGR2	0.011	4161.808	0.42
A-CGR3	0.011	4033.859	0.35
A-CGR4	0.009	3464.613	0.34
A-CGR5	0.009	3410.315	0.33
A-WUT	0.011	4331.283	0.67
A-WCGR	0.012	4622.722	0.54
A-WCGR1/2	0.015	5454.667	0.91
A-WCGR1/3	0.015	5473.906	0.88
A-WCGR1/4	0.015	5469.175	0.88
A-WCGR1/5	0.015	5494.575	0.89
A-WCGR2	0.012	4630.694	0.54
A-WCGR3	0.012	4628.755	0.55
A-WCGR4	0.01	4082.748	0.58
A-WCGR5	0.01	4101.739	0.58
CB	0.009	3400.204	0.31
CP	0.009	3589.032	0.38
CC	0.014	5460.18	0.87
UT	0.009	3432.515	0.35
CE	0.011	4114.433	0.7
CA	0.009	3576.659	0.38
CG	0.009	3422.498	0.3
GM	0.012	4449.249	0.45
CGR	0.009	3904.75	0.76
CGR1/2	0.009	4045.355	0.94
CGR1/3	0.009	4054.697	0.94
CGR1/4	0.009	4045.532	0.95
CGR1/5	0.009	4043.465	0.94
CGR2	0.009	3841.269	0.71
CGR3	0.009	3720.659	0.65
CGR4	0.009	3641.965	0.57
CGR5	0.008	3539.231	0.51

Métricas	Fração de Mortos	Tempo Hospitalizado (dias)	Fração de Vacinados
CH	0.009	3573.043	0.36
CI	0.009	4010.29	0.91
CK	0.01	3760.788	0.37
CL	0.011	3996.96	0.32
PR	0.009	3405.192	0.3
PH	0.011	3947.349	0.31
PHA	0.015	5441.06	0.79
PM	0.011	3945.847	0.3
PMA	0.015	5409.666	0.79
CR	0.012	4740.297	0.8
WCB	0.01	3751.958	0.48
WCP	0.01	4241.772	0.82
WUT	0.008	3619.796	0.54
WCA	0.01	3711.602	0.41
WCGR	0.009	3914.916	0.8
WCGR1/2	0.009	4036.869	0.93
WCGR1/3	0.009	4058.844	0.94
WCGR1/4	0.009	4046.758	0.94
WCGR1/5	0.009	4048.12	0.94
WCGR2	0.009	3916.507	0.77
WCGR3	0.01	4054.052	0.72
WCGR4	0.01	4053.718	0.66
WCGR5	0.01	4072.912	0.64
WCH	0.01	4066.455	0.58
WCL	0.011	3964.738	0.29
WPR	0.009	3419.76	0.28

Fonte: Elaborado pelo autor.

Tabela 16 – Tabela com todos os resultados para rede com arestas ponderadas $p = 0.75$, em azul os menores valores e em vermelho os maiores.

Métricas	Fração de Mortos	Tempo Hospitalizado (dias)	Fração de Vacinados
A-UT	0.009	3406.291	0.37
A-CGR	0.012	4510.977	0.5
A-CGR1/2	0.014	5393.464	0.88
A-CGR1/3	0.014	5396.322	0.89
A-CGR1/4	0.015	5401.999	0.9
A-CGR1/5	0.014	5379.187	0.9
A-CGR2	0.011	4139.817	0.43
A-CGR3	0.011	3975.789	0.36
A-CGR4	0.009	3425.762	0.32
A-CGR5	0.009	3355.054	0.3
A-WUT	0.011	4288.37	0.65
A-WCGR	0.012	4547.319	0.55
A-WCGR1/2	0.014	5374.054	0.86
A-WCGR1/3	0.014	5373.394	0.88
A-WCGR1/4	0.014	5360.752	0.89
A-WCGR1/5	0.014	5371.063	0.89
A-WCGR2	0.012	4551.284	0.53
A-WCGR3	0.012	4608.442	0.54
A-WCGR4	0.01	4038.795	0.56
A-WCGR5	0.01	4041.619	0.57
CB	0.009	3376.957	0.3
CP	0.009	3601.455	0.44
CC	0.014	5425.196	0.87
UT	0.008	3384.272	0.34
CE	0.011	4148.3	0.75
CA	0.009	3652.696	0.45
CG	0.009	3369.636	0.29
GM	0.012	4374.458	0.44
CGR	0.009	3882.782	0.76
CGR1/2	0.009	4040.024	0.93
CGR1/3	0.009	4026.168	0.93
CGR1/4	0.009	4032.729	0.94
CGR1/5	0.009	4027.889	0.94
CGR2	0.009	3807.227	0.71
CGR3	0.009	3707.496	0.63
CGR4	0.008	3605.936	0.54
CGR5	0.008	3484.701	0.5

Métricas	Fração de Mortos	Tempo Hospitalizado (dias)	Fração de Vacinados
CH	0.009	3567.602	0.41
CI	0.009	3978.745	0.9
CK	0.01	3756.786	0.4
CL	0.011	3964.62	0.31
PR	0.009	3383.27	0.29
PH	0.011	3907.41	0.29
PHA	0.014	5339.983	0.77
PM	0.011	3904.691	0.29
PMA	0.014	5358.855	0.79
CR	0.012	4701.007	0.79
WCB	0.009	3720.087	0.45
WCP	0.01	4225.17	0.81
WUT	0.008	3604.701	0.54
WCA	0.01	3773.118	0.43
WCGR	0.009	3896.158	0.79
WCGR1/2	0.016	5923.093	0.93
WCGR1/3	0.009	4020.735	0.92
WCGR1/4	0.009	4023.032	0.93
WCGR1/5	0.009	4043.593	0.94
WCGR2	0.009	3912.14	0.75
WCGR3	0.01	4019.73	0.7
WCGR4	0.01	4032.136	0.64
WCGR5	0.01	4035.849	0.6
WCH	0.01	4017.218	0.56
WCL	0.011	3924.01	0.28
WPR	0.009	3384.994	0.27

Fonte: Elaborado pelo autor.

Tabela 17 – Tabela com todos os resultados para rede com arestas ponderadas $p = 1.00$, em azul os menores valores e em vermelho os maiores.

Métricas	Fração de Mortos	Tempo Hospitalizado (dias)	Fração de Vacinados
A-UT	0.008	3370.725	0.37
A-CGR	0.012	4459.305	0.5
A-CGR1/2	0.014	5272.003	0.86
A-CGR1/3	0.014	5269.998	0.85
A-CGR1/4	0.014	5272.118	0.86
A-CGR1/5	0.014	5267.036	0.85
A-CGR2	0.011	4070.886	0.48
A-CGR3	0.01	3903.074	0.38
A-CGR4	0.009	3386.251	0.32
A-CGR5	0.008	3329.7	0.29
A-WUT	0.011	4226.58	0.64
A-WCGR	0.012	4473.489	0.55
A-WCGR1/2	0.014	5255.222	0.88
A-WCGR1/3	0.014	5260.268	0.88
A-WCGR1/4	0.014	5283.922	0.86
A-WCGR1/5	0.014	5256.038	0.87
A-WCGR2	0.012	4478.604	0.53
A-WCGR3	0.012	4542.163	0.55
A-WCGR4	0.01	4021.754	0.55
A-WCGR5	0.01	4007.248	0.56
CB	0.009	3345.877	0.31
CP	0.009	3596.557	0.5
CC	0.014	5412.409	0.86
UT	0.008	3349.457	0.32
CE	0.01	4001.191	0.73
CA	0.01	3781.089	0.49
CG	0.009	3357.337	0.28
GM	0.012	4345.141	0.43
CGR	0.009	3847.166	0.75
CGR1/2	0.009	3994.135	0.93
CGR1/3	0.009	3989.692	0.93
CGR1/4	0.009	4010.749	0.93
CGR1/5	0.009	3977.147	0.93
CGR2	0.009	3796.264	0.7
CGR3	0.008	3696.406	0.62
CGR4	0.008	3577.656	0.54
CGR5	0.008	3473.688	0.47

Métricas	Fração de Mortos	Tempo Hospitalizado (dias)	Fração de Vacinados
CH	0.009	3552.474	0.46
CI	0.009	3943.418	0.9
CK	0.01	3750.344	0.42
CL	0.011	3903.973	0.32
PR	0.009	3334.411	0.27
PH	0.01	3852.392	0.29
PHA	0.014	5265.469	0.78
PM	0.01	3880.232	0.29
PMA	0.014	5288.059	0.78
CR	0.012	4663.076	0.77
WCB	0.009	3675.024	0.43
WCP	0.01	4169.536	0.81
WUT	0.008	3571.838	0.53
WCA	0.01	3791.627	0.49
WCGR	0.009	3845.673	0.8
WCGR1/2	0.009	4000.327	0.93
WCGR1/3	0.009	3985.601	0.93
WCGR1/4	0.009	3983.626	0.93
WCGR1/5	0.009	3993.274	0.93
WCGR2	0.009	3884.92	0.74
WCGR3	0.01	3970.448	0.68
WCGR4	0.01	4007.267	0.63
WCGR5	0.01	4015.706	0.59
WCH	0.01	3998.115	0.59
WCL	0.01	3900.588	0.29
WPR	0.009	3332.773	0.27

Fonte: Elaborado pelo autor.