

Um estudo sobre árvores geradoras como alternativas conexas para o controle de doença em um grafo de infecção*

Pedro Guimarães Caninas¹

Bacharelado em Sistemas de Informação - UFSC

Florianópolis - SC - Brasil

pedro.caninas@grad.ufsc.br

Priscila Cardoso Calegari, Álvaro Junio Pereira Franco

Departamento de Informática e Estatística – UFSC

Florianópolis - SC - Brasil

priscila.calegari@ufsc.br, alvaro.junio@ufsc.br

RESUMO

Neste trabalho utilizamos um modelo epidemiológico sobre grafos. As pessoas estão representadas e classificadas de acordo com os grupos definidos pelo modelo epidemiológico utilizado. Uma certa movimentação de pessoas é considerada pelo modelo, além de uma medida de distanciamento social. Chamamos um grafo com todas essas características de *grafo de infecção*. Realizamos experimentos que ilustram a dinâmica do espalhamento de uma doença sobre um grafo de infecção. Os experimentos consideram dois grafos de infecção que capturam a topologia de duas cidades do Brasil, considerando a vizinhança entre bairros. Os experimentos também consideram árvores geradoras desses grafos de infecção buscando entender o espalhamento de uma doença nesses subgrafos minimamente conexos. Os resultados sugerem estratégias para redução de pico de infectados em uma epidemia.

PALAVRAS CHAVE. Grafo de infecção, árvore geradora, modelo epidemiológico. PO na

Área de Saúde, Simulação, Teoria e Algoritmos em Grafos

ABSTRACT

In this work we use an epidemiological model on graphs. The people are represented and classified according to defined groups by the epidemiological model used. The movement of people is considered by the model, also a measure of social distancing. We call a graph with all these characteristics an *infection graph*. We performed experiments that illustrate the dynamics of disease spreading on infection graph. The experiments consider two infection graphs that capture the topology of two cities in Brazil, taking in account the proximity between neighborhoods. The experiments also consider spanning trees of these infection graphs in order to understand the spread of a disease in these minimally connected subgraphs. The results indicate strategies to reduce an epidemic infected peak.

KEYWORDS. Infection graph, spanning tree, epidemiological model.

OR in the Health Area, Simulation, Theory and Algorithms in Graphs

*Este trabalho recebeu apoio financeiro da Fundação de Amparo à Pesquisa e Inovação do Estado de Santa Catarina - FAPESC e da Universidade Federal de Santa Catarina - UFSC.

¹Bolsista do CNPq - Brasil

1. Introdução

Em 2020, um modelo epidemiológico sobre um grafo com restrições na circulação de pessoas foi desenvolvido [Franco, 2020] tendo como base um modelo epidemiológico sobre um autômato celular [White et al., 2007]. O modelo epidemiológico sobre um grafo considera de fato a aplicação de um modelo epidemiológico em cada vértice do grafo; considera uma quantidade de pessoas em cada vértice; e considera uma certa movimentação das pessoas entre vértices adjacentes. O trabalho de [Tomassini, 2006] descreve um modelo sobre *redes de autômato generalizado*, semelhante à ideia de aplicar um modelo epidemiológico sobre um grafo. Um modelo epidemiológico qualquer pode ser aplicado a cada vértice. Na ocasião, foi escolhido o modelo clássico SIR em que as pessoas são classificadas como *suscetíveis*, *infectadas* e *recuperadas* (veja este e outros modelos clássicos, por exemplo, em [Hethcote, 1989]). Dessa forma, cada pessoa de um vértice está no correspondente grupo SIR. Além disso, as pessoas podem aderir ou não a um *distanciamento social*, uma medida dada por um índice de distanciamento por vértice.

Em seguida, utilizamos a notação de conjuntos para descrever o modelo. O grupo SIR de pessoas no vértice i e no tempo t é dado por $N_i^t = S_i^t \cup I_i^t \cup R_i^t$. O subconjunto $\bar{S}_i^t \subseteq S_i^t$ ($\tilde{S}_i^t \subseteq S_i^t$) destaca as pessoas suscetíveis do vértice i no tempo t que atendem (não atendem) ao distanciamento social. Da mesma forma temos isso para os infectados \bar{I}_i^t (\tilde{I}_i^t) e recuperados \bar{R}_i^t (\tilde{R}_i^t). A constante α_i^t corresponde ao índice de pessoas do vértice i que atendem ao distanciamento social no tempo t . Com isso, o número de suscetíveis que atendem e não atendem ao distanciamento é, respectivamente $|\bar{S}_i^t| = \lfloor \alpha_i^t |S_i^t| \rfloor$ e $|\tilde{S}_i^t| = |S_i^t| - |\bar{S}_i^t|$. O mesmo é aplicado para obter os números de infectados e de recuperados que atendem e não atendem ao distanciamento social. Portanto, as pessoas de i e no tempo t que atendem ao distanciamento social pertencem ao conjunto $\bar{N}_i^t = \bar{S}_i^t \cup \bar{I}_i^t \cup \bar{R}_i^t$; e as pessoas que não atendem ao distanciamento social pertencem ao conjunto $\tilde{N}_i^t = \tilde{S}_i^t \cup \tilde{I}_i^t \cup \tilde{R}_i^t$.

Neste modelo, a movimentação das pessoas acontece da seguinte forma. Existem pessoas circulando em cada vértice i do grafo e em cada tempo t . Essas pessoas são pessoas do próprio vértice i (de \bar{N}_i^t e de \tilde{N}_i^t) e pessoas de vértices adjacentes de i que não atendem ao distanciamento social (de $\bigcup_{j \in \chi(i)} \tilde{N}_j^t$, sendo $\chi(i)$ o conjunto dos vértices adjacentes de i). Um fator $\beta_{j \rightarrow i}^t$ é utilizado para representar a proporção de pessoas do vértice j que vão para o vértice i sendo $j \in \chi(i)$. Em outra seção, logo depois de apresentar a construção de uma instância do problema, vamos detalhar um pouco mais a circulação de pessoas nos vértices de um grafo.

O conjunto de suscetíveis que encontram com um infectado no vértice i e no tempo t é denotado por X_i^t . Considerando o distanciamento social, temos \bar{X}_i^t (\tilde{X}_i^t) o conjunto de pessoas $p \in \bar{S}_i^t$ ($p \in \tilde{S}_i^t$) que encontram com um infectado no vértice i (em vértices adjacentes a i). Note que o modelo considera uma movimentação de pessoas que atendem ao distanciamento social. Tais pessoas precisaram se movimentar por algum motivo. Dessa forma, temos $X_i^t = \bar{X}_i^t \cup \tilde{X}_i^t$. Finalmente, o modelo sobre o grafo considerando o modelo epidemiológico SIR em cada vértice é o seguinte:

$$\begin{aligned}
 |I_i^{t+1}| &\geq |I_i^t| - \lceil \varepsilon \mathbb{E}[|I_i^t|] \rceil + \lfloor v(\mathbb{E}[|\bar{X}_i^t|] + \mathbb{E}[|\tilde{X}_i^t|]) \rfloor =: |I_i^{t+1}|, \\
 |S_i^{t+1}| &\leq |S_i^t| - \lfloor v(\mathbb{E}[|\bar{X}_i^t|] + \mathbb{E}[|\tilde{X}_i^t|]) \rfloor =: |S_i^{t+1}|, \\
 |R_i^{t+1}| &\leq |R_i^t| + \lceil \varepsilon \mathbb{E}[|I_i^t|] \rceil =: |R_i^{t+1}|,
 \end{aligned}$$

sendo v é uma taxa de virulência, ε é a taxa de recuperação; o valor esperado de $|\bar{X}_i^t|$ definido como $\mathbb{E}[|\bar{X}_i^t|] = \bar{S}_i^t \mathbb{P}\{\bar{Y}_{pi}^t\}$ com \bar{S}_i^t sendo o número de pessoas suscetíveis de i que atendem ao distanciamento social mas que por algum motivo circulam somente em i e \bar{Y}_{pi}^t o valor esperado de $|\tilde{X}_i^t|$

definido como sendo $E[|\tilde{X}_i^t|] = \tilde{S}_i^t P\{\tilde{Y}_{pi \rightarrow i}^t\} + \sum_{j \in \chi(i)} \tilde{S}_j^t P\{\tilde{Y}_{pi \rightarrow j}^t\}$, com \tilde{S}_i^t sendo o número de pessoas suscetíveis de i que não atendem ao distanciamento social e que circulam em i e \tilde{S}_j^t sendo o número de pessoas suscetíveis de i que não atendem ao distanciamento social e que circulam em um vértice adjacente $j \in \chi(i)$. As probabilidades $P\{\tilde{Y}_{pi}^t\}$, $P\{\tilde{Y}_{pi \rightarrow i}^t\}$ e $P\{\tilde{Y}_{pi \rightarrow j}^t\}$ são todas iguais ao número de pessoas infectadas circulando em um vértice dividido pelo total de pessoas circulando em tal vértice. Dessa forma, temos para cada tempo e para cada vértice do grafo, limitantes inferiores para o número de infectados e limitantes superiores para os números de suscetíveis e recuperados. Para a execução dos experimentos, os números de infectados, suscetíveis e recuperados são fixados nos limitantes encontrados.

Na época, foram desenvolvidos experimentos sobre grafos com topologias específicas como, por exemplo, a topologia *cíclica* (o grafo é um ciclo), a topologia *completa* (todos os vértices do grafo estão conectados entre si), e a topologia *grade* (o grafo conecta os vértices em forma de uma matriz bidimensional). Realizamos ainda experimentos em outras topologias fixas tais como a topologia *roda* (um ciclo com um vértice no centro do ciclo e todos os vértices do ciclo ligados a este vértice central); a topologia *estrela* que é semelhante à roda exceto pelo fato que não há a formação do ciclo neste caso; e a topologia *linha* que consiste de um grafo que é um caminho simples. Naquele momento, constatamos experimentalmente que o espalhamento de uma doença se comporta de maneira distinta sobre grafos com diferentes topologias mas com as mesmas condições iniciais e mesmas restrições de circulação de pessoas. As redes utilizadas naqueles experimentos foram geradas por um programas de computador. Cada rede tinha 100 vértices, cada vértice possuía 3.000 pessoas e inicialmente haviam 400 pessoas infectadas em um único vértice da rede em questão.

Um modelo epidemiológico pode ser utilizado para estudar e prever a dinâmica de uma doença em uma população. O trabalho de [Barbosa et al., 2021] faz uma análise da pandemia na cidade de Marabá utilizando o modelo SEIR, com dados reais da cidade e da doença (Covid-19). Outros trabalhos, como o [Manrique-Abril et al., 2020], consideram uma abordagem SIR puramente matemática para analisar e prever o desenvolvimento da Covid-19 no território colombiano. O trabalho [Visbal e Pedraza, 2020] também utiliza o modelo SIR e realiza uma conexão entre parâmetros do modelo e motivações em torno das medidas de contenção adotadas na cidade de Santa Marta, Colômbia. Os modelos epidemiológicos também podem ser aplicados sobre grafos, como feito no trabalho [Keeling e Eames, 2005]. Nele, cada pessoa é representada como um vértice da rede e a proximidade/contato entre pessoas são representadas como arestas, individualizando de fato as pessoas. No nosso trabalho, consideramos o modelo epidemiológico SIR aplicado sobre dois grafos, com os vértices representando um conjunto de pessoas que pertencem a regiões adjacentes de duas cidades do Brasil. O objetivo aqui é analisar a dinâmica de uma doença sobre grafos depois de alterar a sua topologia. Os números de pessoas em cada vértice das redes de entrada foram obtidos através de dados oficiais (IBGE e Data.Rio). Apresentamos os resultados experimentais comparando o desenvolvimento de uma doença ao longo do tempo sobre os grafos originais com subgrafos minimamente conexos (árvores). Com base nos resultados experimentais observados, podemos dizer que na grande maioria dos casos, há maior controle sobre o espalhamento de uma doença sobre um grafo conexo quando a topologia do grafo é alterada para uma árvore.

Na próxima seção, comentamos como foram construídas as instâncias utilizadas neste trabalho. Podemos também chamar cada instância de *grafos de infecção*. Dedicamos as outras seções para apresentação dos experimentos realizados e dos resultados obtidos.

2. Sobre a construção das instâncias utilizadas - os grafos de infecção

Neste trabalho são consideradas duas instâncias para o problema. As instâncias representam de certa forma as cidades de Florianópolis e Rio de Janeiro. Nesta seção, daremos maior destaque para a construção do grafo de infecção da cidade Rio de Janeiro. Para o grafo de infecção da cidade Florianópolis, a construção foi feita de forma similar e descrita no seguinte trabalho de conclusão de curso [Ataide, 2022]. No entanto, é importante pontuar que para o caso da cidade de Florianópolis, os vértices correspondem a grupos de bairros vizinhos, simplificando a construção deste grafo de infecção. No total, foram desenvolvidos 16 grupos de bairros e 25 conexões interligando os grupos. Os grupos combinam bairros que estão em regiões próximas uns dos outros e foram assim definidos: Grupo 1: Cachoeira do Bom Jesus, Ingleses, Canasvieiras, Pontas das Canas, Santinho; Grupo 2: Vargem Grande, Rio Vermelho; Grupo 3: Vargem Pequena, Jurerê; Grupo 4: Santo Antônio de Lisboa, Ratones, Saco Grande, Costa da Lagoa; Grupo 5: Itacorubi, João Paulo, Lagoa da Conceição, Barra da Lagoa; Grupo 6: Balneário, Jardim Atlântico, Estreito, Novo Continente; Grupo 7: Coloninha, Sapé, Monte Cristo, Capoeiras; Grupo 8: Vila Aparecida, Abraão, Coqueiros; Grupo 9: Agronômica, Prainha, Centro, Monte Serrat; Grupo 10: Trindade, Pantanal, Saco dos Limões; Grupo 11: Rio Tavares, Costeira do Pirajubaé; Grupo 12: Canto da Lagoa, Córrego Grande; Grupo 13: Tapera, Carianos, Fazenda do Rio Tavares, Campeche; Grupo 14: Morro das Pedras, Armação; Grupo 15: Ribeirão da ilha, Alto Ribeirão; Grupo 16: Caieira da Barra do Sul, Pântano do Sul. A Figura 1 (à esquerda) ilustra o desenho do grafo simples.

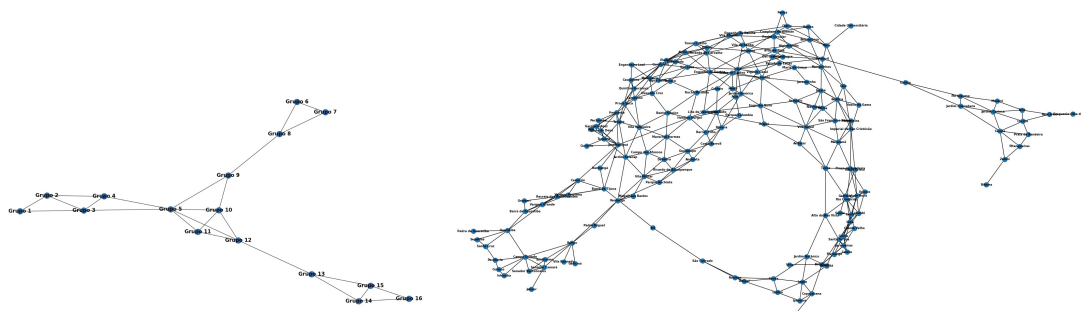


Figura 1: Grafos simples das instâncias de Florianópolis (à esquerda) e do Rio de Janeiro (à direita).

Já para a cidade do Rio de Janeiro, a construção do grafo de infecção não considera essa simplificação. Os vértices dos grafos da cidade do Rio de Janeiro representam bairros da cidade, as arestas representam uma vizinhança entre bairros. Na Figura 2, temos o mapa do município do Rio de Janeiro com destaque para os bairros da cidade (números indo-arábicos).

Note que a cidade do Rio está dividida em 16 *regiões de planejamento*. Cada região possui vários bairros. Para esta instância, consideramos uma construção *mais fina*, representando cada bairro em um vértice do grafo. Então, os vértices adjacentes ao vértice correspondente ao bairro 024 (Copacabana) são os vértices 023 (Leme), 020 (Botafogo), 027 (Lagoa) e 025 (Ipanema). Dessa forma, criamos uma lista de adjacências para cada vértice e construímos o grafo simples desta instância que está ilustrado na Figura 1 (à direita).

A população de cada bairro foi obtida a partir de dados oficiais do Instituto Municipal de Urbanismo Pereira Passos, referência na divulgação de dados geográficos para o planejamento urbano e produção cartográfica do Rio. Este instituto mantém o portal chamado Data.Rio que disponibiliza a população do Rio de Janeiro por bairro, extraídos a partir do Censo Demográfico

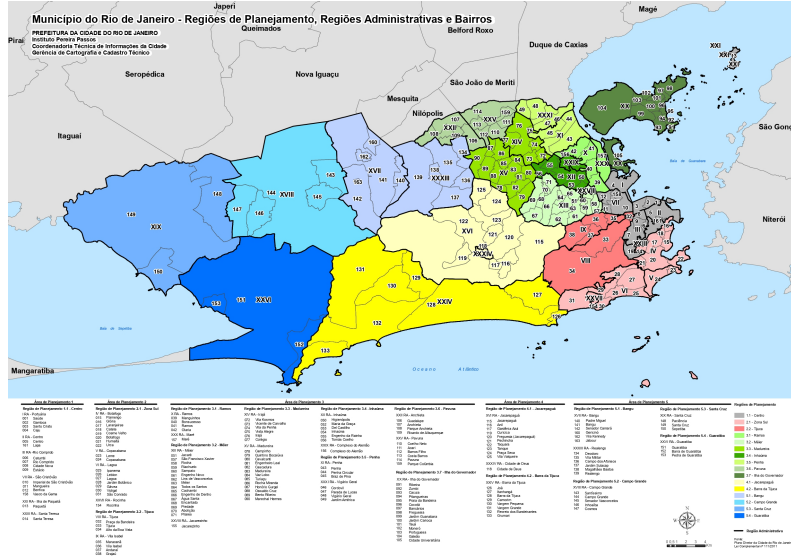


Figura 2: Município do Rio de Janeiro, Fonte: Plano Diretor da Cidade do Rio de Janeiro Lei complementar nº 111/2011.

do Brasil de 2010 realizado pelo Instituto Brasileiro de Geografia e Estatística (IBGE). Algumas alterações foram necessárias considerar. Os dados do Data.Rio não possuem a população dos bairros Lapa, Vila Kennedy e Jabour, pois em 2010 tais bairros não existiam. Por causa disso estes bairros tiveram que ser removidos e as adjacências atualizadas para refletir como os bairros eram em tal ano. Nos nossos experimentos, o bairro Paquetá também não foi considerado por se tratar de uma ilha sem acesso por ruas, apenas por barcas que transportam uma quantidade de pessoas que pode ser no máximo 0,0005% da população total da cidade.

Ainda foram consideradas as seguintes simplificações. Toda aresta $i - j$ possui fluxo de pessoas de $i \rightarrow j$ e de $j \rightarrow i$. Para o caso do Rio, a quantidade de pessoas do vértice i que circula no próprio vértice e nos vértices adjacentes é igualmente proporcional ao número total de pessoas de i , ou seja, fazemos $\beta_{i \rightarrow j}^t = \frac{1}{|X(i)|+1}$ para todo vértice i e todo tempo t . Para o caso de Florianópolis, cada valor $\beta_{i \rightarrow j}^t$ foi definido como sendo a proporção entre a população de j e a soma das populações de i e de todos os vértices adjacentes a i . Para ambas as instâncias, o fator de distanciamento social foi fixado em 40% para todo vértice i e todo tempo t . As pessoas que atendem ao distanciamento social mas que ainda saem para circular no próprio vértice são as seguintes frações $\frac{2}{3}$ para os suscetíveis, $\frac{1}{10}$ para os infectados e $\frac{3}{5}$ para os recuperados.

Aproveitamos a apresentação da construção da instância do Rio para voltar ao assunto da movimentação de pessoas que o modelo considera e detalhar um pouco mais a circulação de pessoas em um vértice do grafo. A Figura 3 destaca o número de pessoas que estão circulando no Flamengo no dia 30.

Existem pessoas do próprio Flamengo que atendem ao distanciamento social (\bar{S} , \bar{I} e \bar{R}); pessoas do próprio Flamengo que não atendem ao distanciamento social (\tilde{S} , \tilde{I} e \tilde{R}); e pessoas de bairros vizinhos ao Flamengo que não atendem ao distanciamento social ($\tilde{\tilde{S}}$, $\tilde{\tilde{I}}$ e $\tilde{\tilde{R}}$). A figura destaca também o número de pessoas suscetíveis circulando no Flamengo mas pertencentes a bairros vizinhos ao Flamengo. Esse número é obtido através da distribuição de pessoas suscetíveis desses

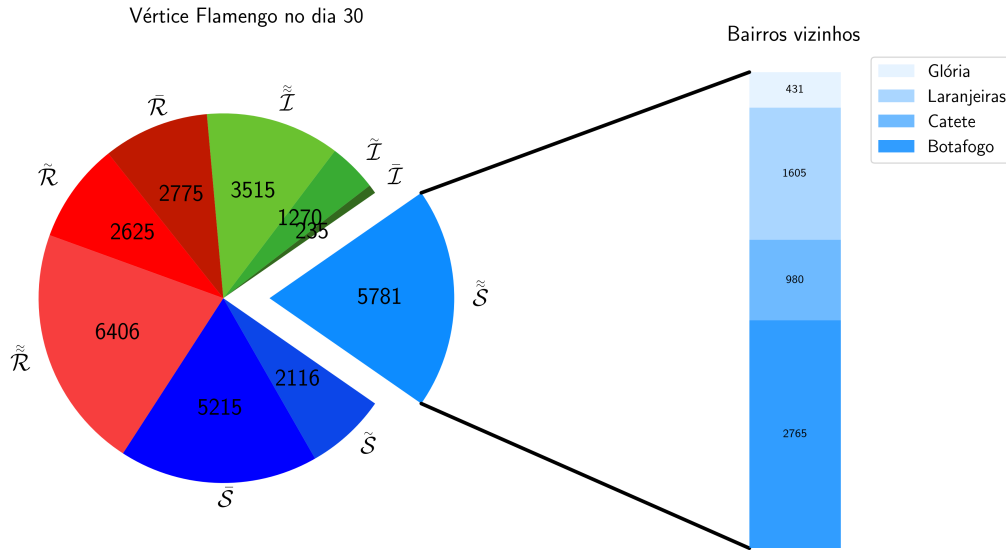


Figura 3: Estão em destaque somente as pessoas que circulam no Flamengo no tempo $t = 30$. Os números totais de suscetíveis, infectados e recuperados que atendem ao distanciamento são 11.594, 2.635 e 5.787, respectivamente.

bairros (usando $\beta_{i \rightarrow j}^t$ com j igual a Flamengo e i um bairro vizinho).

Na próxima seção iniciamos a apresentação dos resultados obtidos quando consideramos a topologia original do grafo de infecção e quando alteramos a topologia do grafo para uma árvore geradora.

3. Algumas árvores geradoras de um grafo de infecção

Nesta seção descrevemos como encontramos alguns subgrafos conexos de um grafo de infecção original. A ideia é encontrarmos subgrafos minimamente conexos envolvendo todos os vértices do grafo original para assim manter a característica de que no grafo obtido ainda há caminhos entre quaisquer vértices do grafo original. Com isso, seria possível movimentar mercadorias por todo o grafo. De fato, o modelo considera uma restrição sobre o movimento das pessoas, mas não sobre o movimento das mercadorias.

Quando temos um subgrafo que é uma árvore contendo todos os vértices do grafo original, então essa árvore é chamada de *geradora* do grafo original [Bondy e Murty, 1976]. Utilizamos *árvores geradoras* obtidas através do processo de busca em largura (em profundidade) sobre o grafo original. Tais processos, devolvem *árvores de busca em largura* ou (*em profundidade*). O livro [Dasgupta et al., 2009] é uma boa referência de leitura sobre as buscas em largura (em profundidade) e as características das correspondentes árvores devolvidas pelas buscas. Na Figura 4, temos um exemplo de uma árvore de busca em profundidade considerando o grafo de infecção do Rio de Janeiro.

Agora podemos comparar o comportamento do espalhamento de uma doença sobre o grafo de infecção original e sobre uma árvore geradora, com as mesmas condições iniciais de pessoas infectadas para ambos os casos. No primeiro experimento, observamos os resultados de duas simulações do espalhamento de uma doença considerando infectados somente no bairro Flamengo

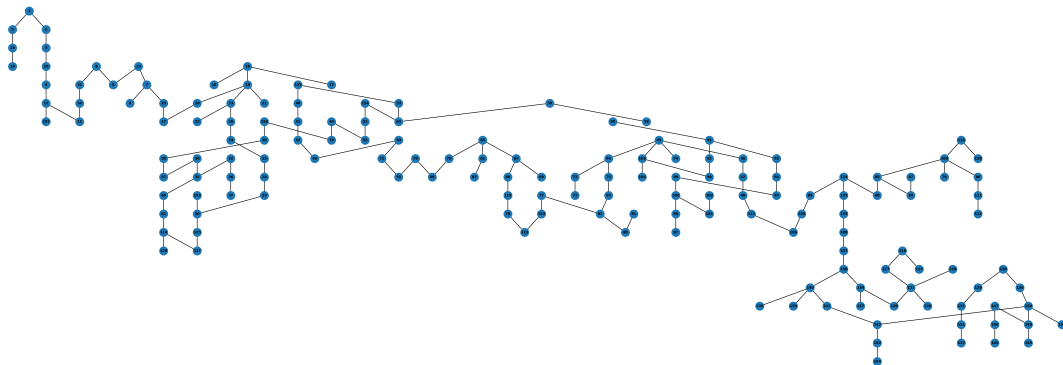


Figura 4: Árvore de busca em profundidade do grafo de infecção do Rio de Janeiro considerando como raiz da árvore o bairro Flamengo (vértice com rótulo 14 localizado na região de planejamento: Zona Sul).

- Figura 5 (na linha cheia, o experimento considerado foi aplicado sobre o grafo de infecção original da cidade do Rio de Janeiro, enquanto que na linha triangulada, o experimento considerado foi aplicado sobre uma árvore geradora (em largura) com raiz no Flamengo). No início da simulação, haviam 2.502 pessoas infectadas no bairro Flamengo (5% da população do bairro). Note que houve uma certa amortização na curva dos infectados quando comparamos os resultados considerando o grafo original e uma árvore geradora (em largura). Além disso, o número de infectados no grafo original atingiu o pico com 810.106 pessoas, enquanto que na árvore geradora, o pico foi de 527.026 pessoas.

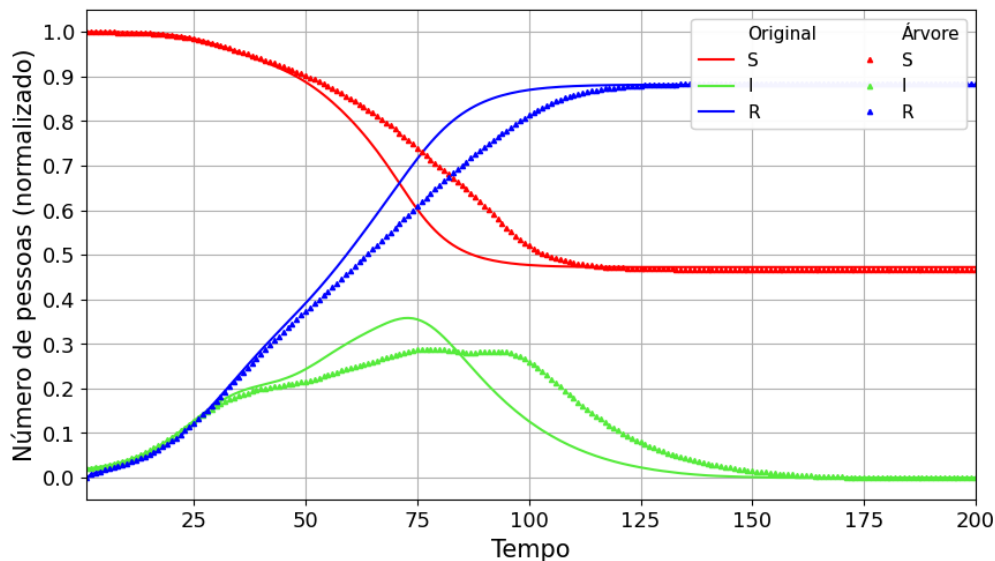


Figura 5: Rio de Janeiro - Resultado da simulação sobre o grafo de infecção original (linha cheia) e sobre uma árvore geradora com raiz no Flamengo (linha triangulada) - a condição inicial considera que 5% do total de pessoas que vivem no bairro Flamengo estão infectadas.

Os resultados iniciais obtidos nos motivou a verificar o espalhamento de uma doença sobre um grafo considerando inicialmente infectados em cada um dos bairros da cidade. Como

existem 159 vértices no grafo do Rio de Janeiro, foram construídas 159×2 árvores geradoras (159 árvores de busca em largura e 159 árvores de busca em profundidade - cada bairro sendo uma raiz de exatamente duas árvores). Obtemos os seguintes resultados ilustrados na Figura 6.

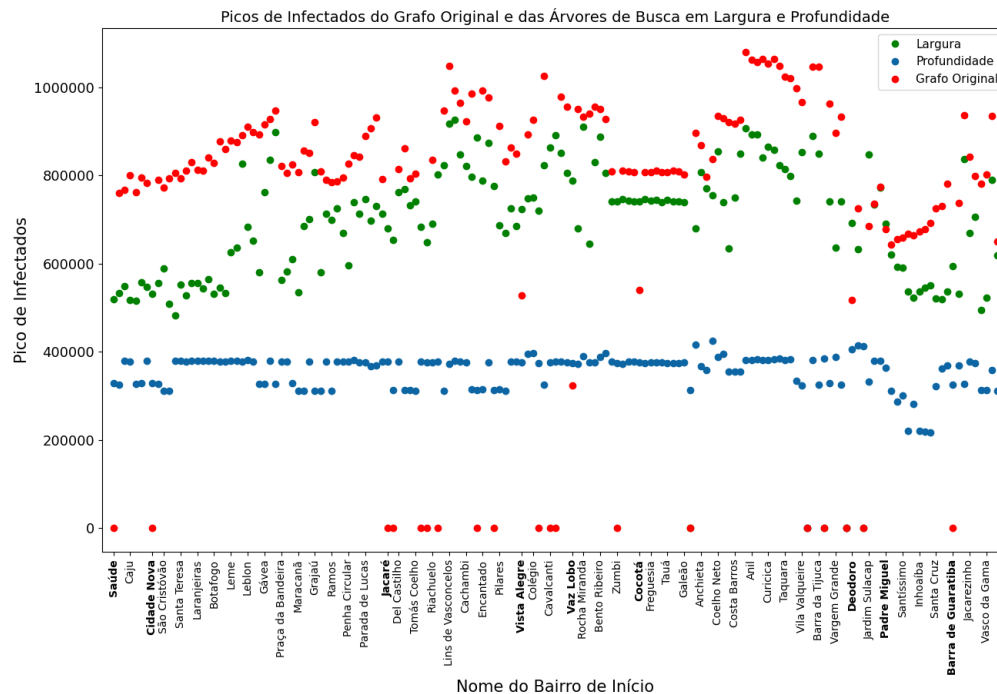


Figura 6: Rio de Janeiro - Observe três curvas para os picos de infectados para cada uma das árvores de busca em largura, em profundidade e para o grafo de infecção original. Descrevemos alguns bairros da cidade no eixo x e aqueles citados no texto aparecem em negrito.

Os experimentos nos indicam que manter o grafo de infecção original proporciona, para quase todos os casos, um pico de pessoas infectadas maior que as topologias das árvores geradoras testadas. As árvores de busca em profundidade possuem grande destaque no que diz respeito ao número de pessoas infectadas no pico, tal número muito próximo de uma constante. No entanto, vale a pena citar que em alguns casos, o grafo de infecção original é a melhor opção de topologia pois a doença não conseguiu se espalhar (a distribuição das pessoas da forma como foi considerada no modelo implicou no número esperado de infectados igual a 0 (zero) para todo vértice logo no início da simulação). Isso acontece quando os infectados estão somente nos bairros Saúde, Cidade Nova, Barra de Guaratiba ou Jacaré, por exemplo. Para poucos casos, temos o grafo original tendo menos infectados no pico que uma árvore de busca em largura, mas não que uma árvore de busca em profundidade (como no caso dos bairros Vista Alegre, Cocotá, Deodoro ou Padre Miguel). Ainda temos o caso de infectados no bairro Vaz Lobo, cujo pico de infectados não é igual a zero para qualquer topologia considerada, mas a melhor opção é o grafo de infecção original.

Realizamos os experimentos também para o grafo de infecção da cidade de Florianópolis. Para este caso, aplicamos o modelo sobre o grafo de infecção original e também sobre árvores de busca em largura e em profundidade - cada grupo de bairros sendo a raiz de uma árvore de busca em largura e em profundidade totalizando 16×2 árvores geradoras. Os resultados relacionados ao número de infectados no pico podem ser observados na Figura 7. Vale a pena lembrar que neste caso, são conhecidos os números reais de suscetíveis, infectados, e recuperados de cada vértice em

um determinado dia (dados reais). A linha horizontal que cruza esta figura está sobre o número 83.271, o número de infectados no pico considerando o grafo de infecção original. Neste caso, os experimentos nos indicam que há muitas árvores geradoras (daquelas consideradas) cujo o número de infectados no pico é maior que o número de infectados no pico do grafo de infecção original. A maioria das árvores de busca em profundidade possuem pico de infectados maior que o pico de infectados do grafo de infecção original, enquanto que a maioria das árvores de busca em largura possuem pico de infectados menor. Das árvores geradoras consideradas, aquela que obteve o menor número de infectados no pico foi uma árvore de busca em largura com raiz no Grupo 11.

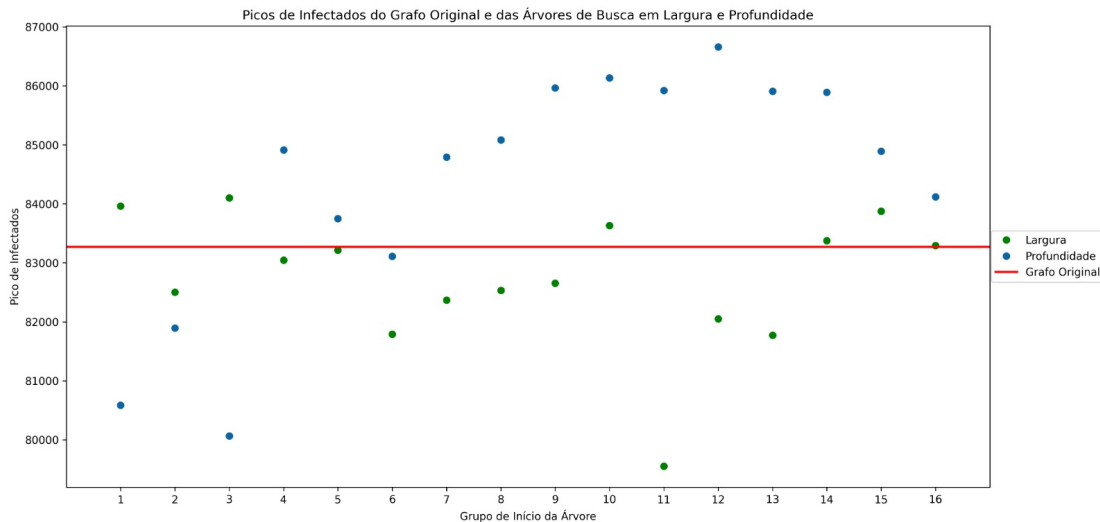


Figura 7: Florianópolis - Observe duas curvas para os picos de infectados para cada uma das árvores de busca em largura e em profundidade, além do pico de infectados considerando o grafo de infecção original.

Podemos então concluir que, na maioria dos experimentos realizados, a restrição da topologia nos leva a um menor pico de infecção no grafo. Isso acontece em decorrência da restrição de movimento que essas mudanças topológicas proporcionam. A movimentação das pessoas é dificultada e os experimentos ilustram que o espalhamento da doença no grafo é reduzido. Com essas restrições, pode ser impossível a movimentação direta de pessoas infectadas entre vértices adjacentes do grafo original. Com isso, a infecção pode levar mais tempo para atingir tais vértices. Vale ressaltar a importância de controlar o espalhamento de uma doença. Uma menor quantidade de pessoas infectadas simultaneamente implica em menor ocupação em hospitais e em UTIs, por exemplo, o que nos leva a possibilidade de tratar eficientemente cada um dos pacientes de acordo com a gravidade dos sintomas.

Na próxima seção, descreveremos uma heurística que, a partir de uma árvore geradora (em largura ou em profundidade), passa por outras árvores geradoras do grafo de infecção, vizinhas da árvore de partida, e destaca aquela com o menor número de infectados.

4. Uma heurística que encontra novas árvores geradoras

Nesta seção, descrevemos um procedimento que recebe uma árvore geradora T de um grafo de infecção G e devolve uma lista de outras árvores geradoras de G , caso o conjunto de arestas de G seja diferente do conjunto de arestas de T . Sejam $E(G)$ e $E(T)$ os conjuntos de arestas de G e de T , respectivamente. Definimos uma árvore vizinha T' de T se T' difere de T em uma aresta. Dessa forma, podemos construir uma lista com árvores vizinhas de T da seguinte maneira. Seja e

uma aresta de $E(G) - E(T)$, isto é, uma aresta de G que não está na árvore T . Ao adicionar e em T , criamos um grafo com exatamente um ciclo $(T + e)$. Quase todas as arestas deste ciclo estão na árvore T , exceto por e . Considerando e_1, e_2, \dots, e_k todas as arestas do ciclo que estão na árvore T , ao removermos exatamente uma aresta e_i , obtemos uma nova árvore (geradora de G) vizinha de T . Dessa forma, construímos k árvores vizinhas de T ao remover uma aresta e_i por vez. Para construir uma lista L com todas as árvores vizinhas de T , basta repetimos o procedimento descrito anteriormente para cada aresta em $E(G) - E(T)$. O modelo epidemiológico pode ser executado considerando todas as árvores vizinhas de uma árvore geradora T , encontrando assim, um mínimo local.

O tempo de execução da heurística por árvore geradora é da ordem de $O(nm)$ sendo n e m são os números de vértices e de arestas de G , respectivamente. Isso porque para cada uma das $m - n + 1$ ($O(m)$) arestas de $E(G) - E(T)$, adicionar e em T e encontrar o ciclo consome tempo $O(n)$. Para encontrar o ciclo, basta seguirmos pelos pais das pontas de e na árvore T até chegarmos em um ancestral comum. Então, se a lista de árvores vizinhas de T possui tamanho k , encontrar um mínimo local considerando uma árvore T leva tempo $O(knm)$. No pior caso, teremos $k = O(nm)$. Isso porque para cada $O(m)$ aresta inserida em T , podemos ter ciclos de tamanho $O(n)$ e assim, podemos criar $O(n)$ árvores vizinhas para cada inserção de aresta em T . Logo, o tempo de execução no pior caso é $O(n^2m^2)$. Para o nosso caso específico, essa heurística é executada $2n = O(n)$, uma execução para cada árvore de busca em largura e em profundidade com raiz em cada vértice do grafo de infecção G com n vértices. Então, para o nosso caso específico, o tempo de execução do procedimento total fica igual a $O(n^3m^2)$. Note que a lista de árvores vizinhas não precisa ser construída explicitamente pois podemos manter apenas aquela que possui o melhor resultado. Então, o espaço consumido para guardar uma árvore geradora vizinha é suficiente, ou seja, $O(n)$.

Em seguida apresentamos os resultados experimentais da aplicação dessa heurística sobre algumas árvores geradoras (de busca em largura e em profundidade) do grafo de infecção da cidade de Florianópolis. Os resultados estão em gráficos *boxplot*. Uma linha horizontal no gráfico está sobre o número 83.271, o pico da curva de infectados do grafo de infecção original. Cada *boxplot* corresponde aos resultados do pico da curva de infectados das árvores vizinhas de uma árvore geradora com raiz em um vértice do grafo de infecção original. Podemos ver na Figura 8 que a maioria das árvores vizinhas de árvores de busca em profundidade possuem resultados piores que o resultado do grafo de infecção original. Já observando a Figura 9, essa conclusão se inverte. A maioria das árvores vizinhas de árvores de busca em largura possuem resultados melhores que o resultado do grafo original. O melhor resultado foi obtido através de uma árvore vizinha de uma árvore de busca em largura com raiz no Grupo 7. Neste caso, o pico atingiu 70.647, ou seja, o pico de infectados da árvore vizinha é 15,16% menor que o pico de infectados do grafo original. É importante também dizer que os experimentos indicam que há muitas árvores vizinhas de árvores geradoras que possuem resultados piores quando comparados com o resultado do grafo de infecção original. Há árvores vizinhas de árvores de busca em profundidade com mais de 6% de infectados no pico (Grupos 7 e 15) quando comparados com o pico do grafo de infecção original. Note na Figura 9 que pelo menos $\frac{3}{4}$ das árvores vizinhas de árvores de busca em largura com raiz nos Grupos 4, 6, 7, 11 e 12 possuem resultados melhores que o resultado do grafo de infecção original. Em contrapartida, note na Figura 8 que nem mesmo $\frac{1}{4}$ das árvores vizinhas de árvores de busca em profundidade com raiz nos Grupos 12 e 13 possuem resultados melhores que o resultado do grafo de infecção original.

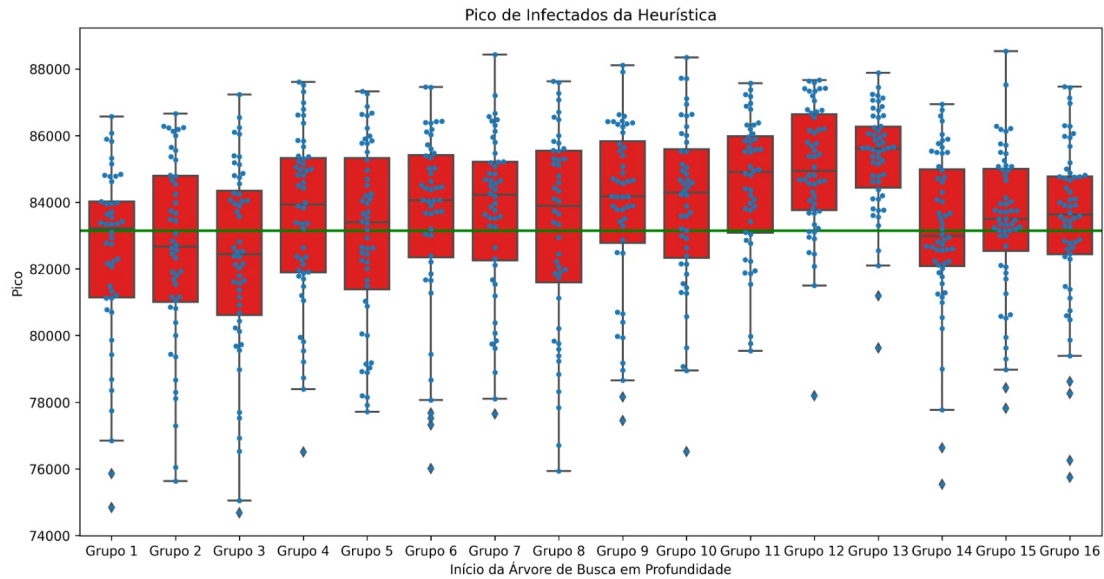


Figura 8: Florianópolis - *Boxplot* com os resultados do pico de infectados da vizinhança sobre árvores de busca em profundidade.

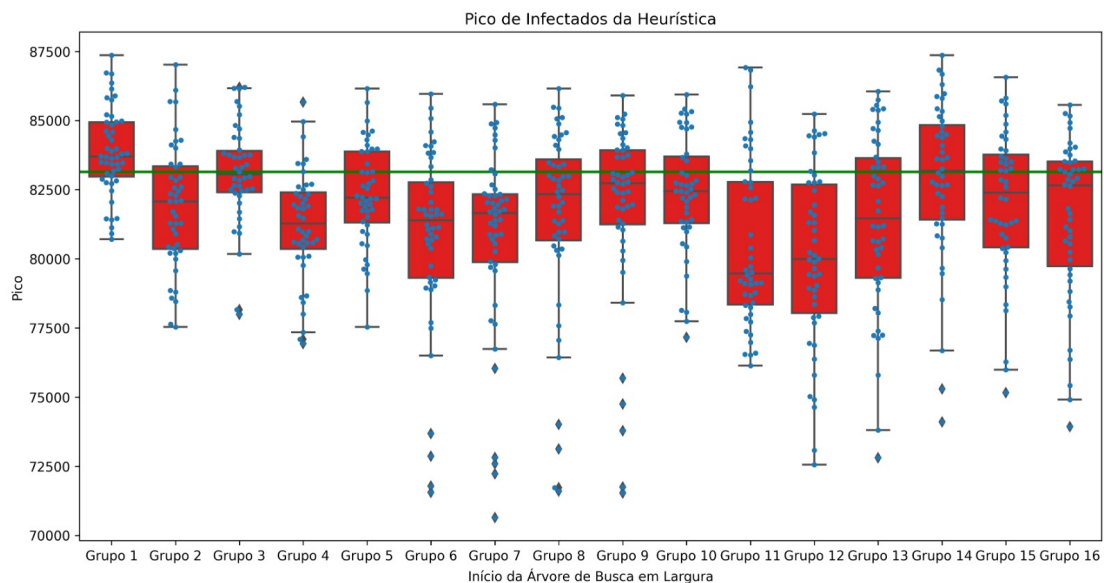


Figura 9: Florianópolis - *Boxplot* com os resultados do pico de infectados da vizinhança sobre árvores de busca em largura.

5. Considerações Finais

Neste trabalho, aplicamos um modelo epidemiológico sobre grafos de infecção. Utilizamos dois grafos de infecção correspondentes a duas cidades do Brasil: Rio de Janeiro e Florianópolis. Cada bairro (ou grupo de bairros para o caso de Florianópolis) é representado como um vértice neste grafo e bairros vizinhos (ou grupos de bairros vizinhos) são conectados por uma aresta do grafo de infecção. Foram utilizados dados conhecidos e reais sobre o número de pessoas

que vivem nos bairros. Para o caso de Florianópolis, utilizamos ainda dados reais do número de pessoas suscetíveis, infectadas e recuperadas, considerando a doença COVID-19. Apresentamos uma heurística que encontra mínimo local dada uma definição de vizinhança de árvore geradora. Foram apresentados os resultados experimentais obtidos pela aplicação do modelo epidemiológico sobre algumas árvores geradoras dos grafos de infecção das duas cidades.

6. Agradecimentos

Este trabalho recebeu apoio financeiro da Fundação de Amparo à Pesquisa e Inovação do Estado de Santa Catarina - FAPESC e da Universidade Federal de Santa Catarina - UFSC.

Referências

- Ataide, W. C. (2022). Minimização do contágio em uma topologia com Programação Matemática. Trabalho de conclusão de curso, Universidade Federal de Santa Catarina. Acessado: 02-06-2022.
- Barbosa, W. F., da Costa Moreira, E. B., de Araújo, J. M., Pazin-Filho, A., e Bueno, C. D. F. (2021). Modelo seir para avaliação do comportamento da pandemia de covid-19 em marabá-pa. *Revista de Medicina USP*, 100(4):322–328.
- Bondy, J. A. e Murty, U. S. R. (1976). *Graph theory with applications*, volume 290. Macmillan London.
- Dasgupta, S., Papadimitriou, C., e Vazirani, U. (2009). *Algoritmos*. Mc Graw Hill.
- Franco, Á. J. P. (2020). Epidemic model with restricted circulation and social distancing on some network topologies. In *International Conference on Cellular Automata for Research and Industry*, p. 261–264. Springer.
- Hethcote, H. W. (1989). Three basic epidemiological models. In *Applied Mathematical Ecology*, p. 119–144. Springer.
- Keeling, M. J. e Eames, K. T. (2005). Networks and epidemic models. *Journal of the royal society interface*, 2(4):295–307.
- Manrique-Abril, F. G., Agudelo-Calderon, C. A., González-Chordá, V. M., Gutiérrez-Lesmes, O., Téllez-Piñerez, C. F., e Herrera-Amaya, G. (2020). Modelo sir de la pandemia de covid-19 en colombia. *Revista de Salud Pública*, 22:123–131.
- Tomassini, M. (2006). Generalized automata networks. In *International Conference on Cellular Automata*, p. 14–28. Springer.
- Visbal, J. H. W. e Pedraza, M. C. C. (2020). Aproximación matemática del modelo epidemiológico sir para la comprensión de las medidas de contención contra la covid-19. *Rev. esp. salud pública*.
- White, S. H., Del Rey, A. M., e Sánchez, G. R. (2007). Modeling epidemics using cellular automata. *Applied Mathematics and Computation*, 186(1):193–202.