

Estudo de critérios de distribuição de vacinas com doses limitadas - uma abordagem baseada em redes complexas

Ronaldo A. M. Júnior¹, Vinícius F. Vieira¹, Carolina R. Xavier¹

¹ Departamento de Ciência da Computação
Universidade Federal de São João del Rei (UFSJ)
São João del Rei – MG – Brasil

ronaldo.alves@ifsudestemg.edu.br, {carolinaxavier, vinicius}@ufsj.edu.br

Abstract. *The assertive choice of groups of individuals to be immunized can directly impact the contamination dynamics of infectious diseases, protecting a fragile part of the population as well as in minimizing the speed of contagion. Such a task is important and challenging, especially in scenarios with a reduced number of doses available, as it involves countless possible combinations for decision making. This research aims to propose a methodology for choosing individuals to be immunized using a network approach. We propose the use of Genetic Algorithms and network centrality measures to define the individuals to be vaccinated. The results indicate that the systematic choice of individuals is a correct decision for this type of problem.*

Resumo. *A escolha assertiva de grupos de indivíduos a serem imunizados pode impactar diretamente na dinâmica de contaminação, tanto na proteção de uma parte frágil da população quanto na minimização da velocidade de contágio. Tal tarefa é importante e desafiadora, especialmente em cenários com número reduzido de doses disponíveis, por envolver inúmeras combinações possíveis para tomada de decisão. O objetivo deste trabalho é propor uma metodologia para a escolha de indivíduos a serem imunizados utilizando uma abordagem em rede. Nós propusemos uma metodologia baseada em algoritmo genético, tendo como linha de base medidas de centralidade para definição dos indivíduos a serem vacinados. Os resultados indicaram que a escolha sistematizada de indivíduos é uma decisão correta para este tipo de problema.*

1. Introdução

Considera-se que a epidemiologia é a ciência que estuda o adoecimento de populações e que analisa os fatores e a distribuição dessas doenças com o objetivo de propor ações de prevenção, controle ou erradicação através de indicadores que auxiliem no planejamento e na administração de ações planejadas [Rouquayrol and Filho 2003]. O controle e acompanhamento de doenças contagiosas é um constante desafio que governos e agências de saúde precisam enfrentar. A definição de estratégias que garantam a prevenção de doenças dessa natureza normalmente envolve a utilização de mecanismos que garantam a imunização.

A imunização busca o aumento da resistência de um indivíduo a infecções. Um dos métodos utilizados para se atingir essa condição é a vacinação, principal mecanismo de promoção de saúde e prevenção de doenças.

O estudo realizado em [Li et al. 2021] concluiu que campanhas de vacinação para controle de doenças infectocontagiosas geradas por dez patógenos diferentes evitaram 37 milhões de mortes (intervalo de confiança de 95%) entre os anos de 2000 e 2019.

O interesse pela dinâmica e a necessidade de se compreender as características de proliferação de doenças motivaram o surgimento da área de epidemiologia matemática.

Dentre vários modelos que reproduzem a dinâmica de espalhamento de doenças, selecionou-se o SIR (Suscetível-Infetado-Recuperado) para as simulações deste trabalho. Este modelo não prevê a possibilidade de reinfeção de um indivíduo.

Trabalhos como [Ndiaye et al. 2020] e [Oliveira et al. 2021] utilizam modelos desta área de estudo para a caracterização da dinâmica de disseminação de doenças. Esta caracterização auxilia na elaboração de estratégias de imunização, como foi realizado no trabalho de [Tanner et al. 2008] que utilizou programação estocástica para encontrar políticas ótimas de vacinação e em [Shulgin et al. 1998] que a partir de modelo SIR propôs o princípio de vacinação em pulso, capaz de erradicar doenças com baixa vacinação.

O interesse também é frequente quando modelos epidemiológicos são aplicados em redes complexas. Alguns trabalhos como [Liu et al. 2017], [Li et al. 2019] e [Rodrigues et al. 2018] avaliaram e propuseram diferentes estratégias de imunização, utilizando uma rede para representar uma população.

Este trabalho realiza o estudo de diferentes critérios de escolhas de indivíduos a serem imunizados que podem ser adotados em cenários controlados onde existe a escassez do número doses para imunização das populações representadas por redes complexas. Assume-se a hipótese de que a escolha assertiva de um grupo de indivíduos para vacinação pode, devido aos seus relacionamentos, diminuir a taxa de contágio ou o número de indivíduos afetados pela doença. Este trabalho propõe resumidamente uma metodologia própria, baseada em algoritmos genéticos, para a seleção dos indivíduos a serem imunizados. Esta é comparada a outros critérios de seleção de indivíduos, que utilizam de medidas de centralidade e de aleatoriedade para a escolha. Os critérios são avaliados conforme o impacto que a imunização do grupo de indivíduos propostos causa no processo de disseminação da doença na rede, experimentados através do modelo epidemiológico SIR.

Foi observado que a escolha sistematizada de indivíduos é uma decisão acertada, embora a metodologia proposta não tenha apresentado vantagem competitiva ante aos critérios baseados em medidas de centralidade estudados.

2. Referencial teórico

2.1. Algoritmos genéticos

Algoritmo genético (AG) é um algoritmo evolutivo usado para resolver problemas de busca e otimização. Inspira-se em técnicas da teoria da evolução e busca reproduzir o comportamento da seleção natural e da genética

onde indivíduos mais adaptados ao meio tendem a perpetuar suas características.

Foca na intensificação de buscas em pontos para a melhora da aptidão (operador de cruzamento) e na diversificação em pontos aleatórios (operador de mutação). A técnica

foi proposta por [Holland 1975]. Um dos principais diferenciais dessa estratégia é a de se trabalhar com uma população de soluções candidatas.

2.2. Redes e Centralidade

Segundo [Barabasi 2016] a representação em rede cria uma linguagem comum para o estudo de sistemas complexos de naturezas diferentes. Formalmente, uma rede $G(V, E)$ é formada pelo conjunto V que contém n vértices e pelo conjunto E que contém m arestas. Cada aresta registra a interação entre dois vértices.

Em redes complexas, a centralidade é uma característica topológica relevante. Este conceito procura quantificar a importância de um vértice da rede ou de suas ligações em relação a toda rede. Existem diversas formas de se calcular a centralidade a depender da aplicação. Algumas das mais importantes medidas são:

- **Grau** - O grau é a medida mais intuitiva, caracterizando-se pela quantidade de conexões que um vértice possui.
- **Betweenness** - Para a obtenção do *betweenness* mede-se o número de vezes que um determinado vértice está no trajeto mais curto entre dois outros vértices da rede. Quanto mais vezes este vértice compor diferentes trajetos, maior será sua importância na rede [Freeman 1978]. É também conhecida como centralidade por intermediação.
- **Autovetor** - A medida de centralidade de autovetor foi proposta em [Bonacich 1987] e busca medir a importância de um vértice em função da importância de seus vizinhos.
- **Pagerank** - O *Pagerank*, proposto em [Page et al. 1999], foi inicialmente utilizado para medir a importância de páginas na Internet, considerando a qualidade e a quantidade dos links que se relacionam com ela. A centralidade que se propaga aos vizinhos é proporcional ao número de arestas de saída. Vértices propagam apenas parte da centralidade adiante aos outros.

2.3. O modelo SIR em redes complexas

A conexão entre os modelos epidemiológicos e redes complexas busca representar a natureza probabilística da disseminação de doenças e considera diversas características topológicas da rede. A intenção é representar melhor as interações aleatórias que acontecem em ambientes reais através de um modelo estocástico.

O SIR usado neste trabalho considera cada um dos vértices e suas conexões. Este é um método individual, baseado em pares e que considera os vértices e suas conexões para determinar a mudança de estado indivíduo a indivíduo. Assim, vértices suscetíveis podem ser contaminados por vizinhos infectados. Todo indivíduo infectado pode se recuperar dada a probabilidade de recuperação da doença.

Cada vértice i possui um estado s_i em um instante t , onde sua classe é definida $s_i(t) = 0$ se i está suscetível, $s_i(t) = 1$ se i está infectado e $s_i(t) = 2$ se i está recuperado.

Um vértice i tem seu estado alterado de $s_i(t) = 0$ para o estado $s_i(t + 1) = 1$ com probabilidade β_{ij} . Já a mudança de $s_i(t) = 1$ para $s_i(t + 1) = 2$ ocorre com a probabilidade γ_i .

A Figura 1 ilustra o comportamento de uma epidemia aplicada a uma rede complexa com 2530 vértices onde, inicialmente, se infectou o nó de maior grau.

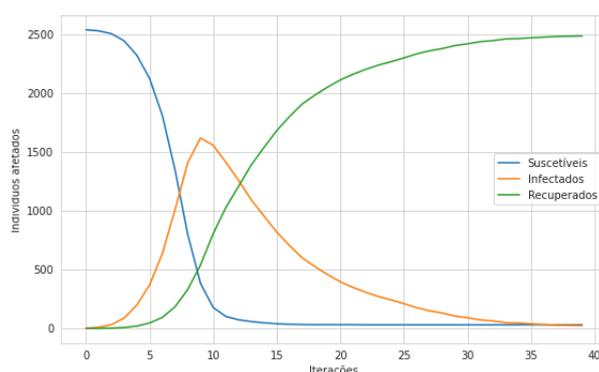


Figura 1. Comportamento epidêmico do modelo SIR aplicado em uma rede complexa. Utilizou-se $\beta = 0.2857$ e $\gamma = 0.1428$

3. Materiais e métodos

Esta seção avalia a efetividade dos diferentes critérios de vacinação na escolha de indivíduos ideais a serem vacinados na rede. São considerados critérios que utilizam um AG, medidas de centralidade ou a aleatoriedade para a escolha dos indivíduos. A ausência de vacinação também é utilizada como controle.

3.1. Redes utilizadas

Foram utilizadas quatro redes reais obtidas no repositório [Peixoto 2020]. A Tabela 1 apresenta as informações de cada uma delas, sendo $|V|$ o número de vértices, $|E|$ o número de arestas, \bar{k} o grau médio, k_{max} o grau máximo e k_{min} o grau mínimo. As colunas c e r representam o coeficiente de *clustering* e assortatividade respectivamente.

Tabela 1. Informações das redes utilizadas

Rede	$ V $	$ E $	\bar{k}	k_{max}	k_{min}	c	r
moreno highschool	70	274	7.83	19	2	0.40	0.01
moreno innovation	241	929	7.71	28	1	0.25	-0.01
moreno health	2539	10455	8.24	27	1	0.14	0.01
moreno oz	217	1839	16.95	56	2	0.30	0.00

[Newman 2002] descreve que redes assortativas ($r > 0$) são robustas enquanto as redes disassortativas ($r < 0$) possuem características contrárias. A robustez define o quanto a rede permanece interligada após a remoção de alguns vértices. Vê-se na Tabela 1 que o coeficiente de assortatividade das redes tende a zero. Este requisito foi considerado para que tal característica topológica não seja significativa na escolha dos vértices, seja no uso de um algoritmo heurístico ou no uso de medidas estáticas de centralidade.

3.2. Métodos

A avaliação da efetividade dos critérios de vacinação é realizada através de uma análise comparativa. Todos os experimentos aplicados baseiam-se em parâmetros epidêmicos estimados para uma epidemia de *Influenza* e obtidos através de informações dos trabalhos de [Bresee et al. 2011] e [Coburn et al. 2009]. Assim, a taxa de infecção β assumida foi de 0.2857 em todas as arestas das redes, a taxa de recuperação γ de 0.1428 para todos os

vértices. Simulações são realizadas com a execução do modelo SIR após a aplicação da imunização proposta pelo critério.

São experimentados diferentes percentuais de indivíduos vacinados (pv), valores entre 15% e 40%. O percentual de indivíduos infectados iniciais é sempre de 5%. Os critérios de vacinação considerados para este estudo são:

- vértices que possuem a maior centralidade de *grau* (*degree*) ou C_{pv}^D ;
- vértices que possuem a maior centralidade de *betweenness* ou C_{pv}^B ;
- vértices que possuem a maior centralidade de *pagerank* ou C_{pv}^P ;
- vértices que possuem a maior centralidade de autovetor (*eigenvector*) ou C_{pv}^E ;
- vértices propostos pelo AG ou C_{pv}^G ;
- vacinação randômica ou C_{pv}^R ;
- ausência de vacinação ou C^\emptyset .

As centralidades são calculadas pelos algoritmos disponíveis no *graph-tool* e o AG é de implementação própria. Utiliza-se de um AG tradicional, onde cada indivíduo é um vetor de índices dos vértices da rede candidatos a vacinação e a função de avaliação é a saída do SIR. O cruzamento é de dois pontos e a mutação se dá com substituição aleatória dos genes. A saída do AG é uma lista ranqueada dos vértices do grafo, ordenada de maneira decrescente pelo número de vezes que o vértice esteve em alguma das 10% melhores soluções de cada uma das 100 execuções feitas do AG para cada cenário.

Para a análise da similaridade entre os conjuntos de indivíduos selecionados pelos critérios, optou-se por utilizar o RBO, ou *Rank-Biased Overlap*. Tal medida foi proposta por [Webber et al. 2010] para atuar com classificações ordinais e indefinidas. A medida respeita um intervalo de zero a um onde maiores valores implicam em maior nível de similaridade entre as listas. O nível de similaridade dado pelo RBO é avaliado junto à intersecção dos indivíduos entre os critérios avaliados.

Os passos a seguir apresentam a dinâmica dos experimentos realizados para cada percentual de vacinados pv e que são posteriormente analisados:

1. Forma-se o grupo de indivíduos candidatos a serem infectados X constituído pela diferença de todos os indivíduos da rede V e os indivíduos selecionados por qualquer critério de vacinação, ou seja, $X = V - (C_{pv}^D \cup C_{pv}^B \cup C_{pv}^P \cup C_{pv}^E \cup C_{pv}^G \cup C_{pv}^R)$.
2. Um grupo de infectados Z_i é formado pela combinação de indivíduos contidos em X e que terá a cardinalidade correspondente a 5% do tamanho da rede, ou seja, igual a $0.05 * |V|$.
3. O conjunto de todos os grupos de infectados é constituído de 1000 grupos e dado por $Z = \{Z_i : i = 1, 2, 3, \dots, 1000\}$ tal que $Z_i \subset X$.
4. O SIR é utilizado para simular 100 vezes uma epidemia, utilizando-se de 60 iterações síncronas (passos) do modelo e aplicado para cada grupo de infectados Z_i e cada critério de vacinação da forma C_{pv} .
5. A vacinação é aplicada nos vértices selecionados pelo critério antes da simulação com o SIR.
6. É considerada apenas a simulação que atingiu o maior número de indivíduos da rede, o que registraria o pior caso possível dentre as simulações. Tal diretiva objetiva eliminar simulações que terminem precocemente.
7. São extraídas dessa simulação algumas variáveis dependentes.

É importante destacar que, para o mesmo percentual de vacinação, os grupos de infectados permanecem os mesmos e são testados para cada critério de vacinação discutido. Portanto, novos percentuais de vacinação geram novos grupos de infectados.

De cada simulação são extraídos as variáveis dependentes *pico*, *afetados*, *beta* e que significam, respectivamente o número máximo de todos os instantes de indivíduos infectados, a soma de indivíduos afetados (recuperados ou infectados) e o β inferido da curva de infecção.

Dentre essas variáveis, a inferência do β demanda um maior detalhamento. Com a probabilidade de recuperação γ mantida com o mesmo valor das simulações, esta inferência é aplicada à curva que representa o número de infectados por instante de tempo do experimento. Através de regressão com a técnica de mínimos quadrados não lineares, cada simulação estocástica tem seus dados ajustados à equação determinística do modelo SIR. A percepção numérica aproximada com a estimativa do β anseia captar o “achatamento” característico da curva de infectados.

A Figura 2b ilustra a curva de infectados de uma simulação estocástica com vacinação de 15% dos indivíduos na rede *moreno highschool* e a curva inferida ajustada aos dados. Pode ser observado na Figura 2a que os efeitos que a variação do β gera em uma curva de infecção quando aplicada em uma equação determinística.

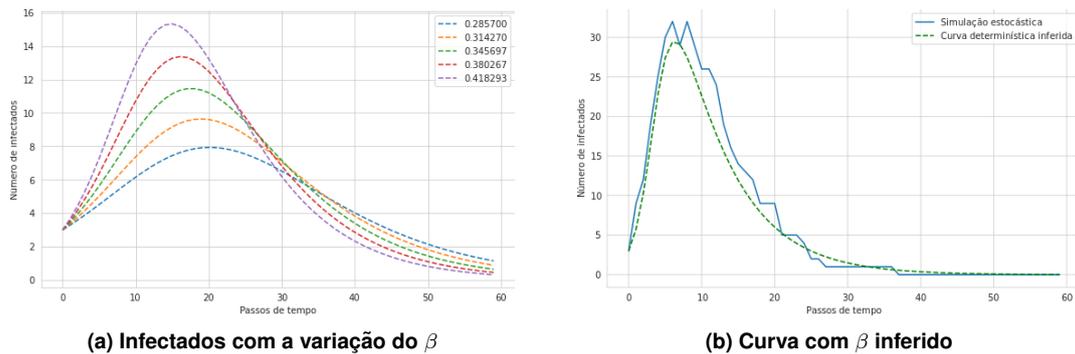


Figura 2. Curvas de infectados e os efeitos do β

Como o SIR utilizado nos experimentos é estocástico, espera-se uma alta variabilidade dos valores observados nas variáveis dependentes. Desta forma, uma série de testes estatísticos são aplicadas para que se avalie se existe diferença significativas na escolha dos critérios experimentados.

4. Resultados

4.1. Similaridades e intersecção

Observou-se que o número de indivíduos comuns selecionados pelos diferentes critérios varia bastante. Vê-se, na Figura 3, que entre os critérios C_{pv}^P e C_{pv}^D a intersecção dos indivíduos selecionados é alta, assim o nível de similaridade.

A similaridade do critério baseado em AG proposto, C_{pv}^G , é normalmente inferior a similaridade entre os demais critérios, apresentando resultados dados por $rbo < 0.6$ em todas as redes. É observada uma proximidade maior na contagem de indivíduos comuns com os critérios C_{pv}^P e C_{pv}^D .

Os níveis de similaridade e intersecção para os critérios C_{pv}^E e C_{pv}^B são mais sensíveis ao nível de complexidade da rede. Nas redes *moreno oz* e *moreno health* esses critérios apresentam menor nível de similaridade e menor contagens de indivíduos comuns quando comparados aos outros critérios. O critério C_{pv}^E é o critério baseado em centralidade que mais difere dos demais de mesma natureza. Seus valores de intersecção são comparáveis aos critérios C_{pv}^P e C_{pv}^D apenas na rede *moreno oz*.

Conforme esperado, os níveis de intersecção e de similaridade do critério C_{pv}^R são baixos quando comparados com qualquer outro critério.

Apenas nas redes *moreno oz* e *moreno health* foi observado que os níveis de intersecção entre os critérios subiu com o aumento do número percentual de vacinados.

A Figura 3 ilustra os níveis de similaridade (mapa de calor) e intersecção (valor no centro do quadrante) dos indivíduos selecionados considerando a vacinação de 40% da população. As diferenças observadas entre as redes indicam que a topologia impacta diretamente na escolha dos indivíduos.

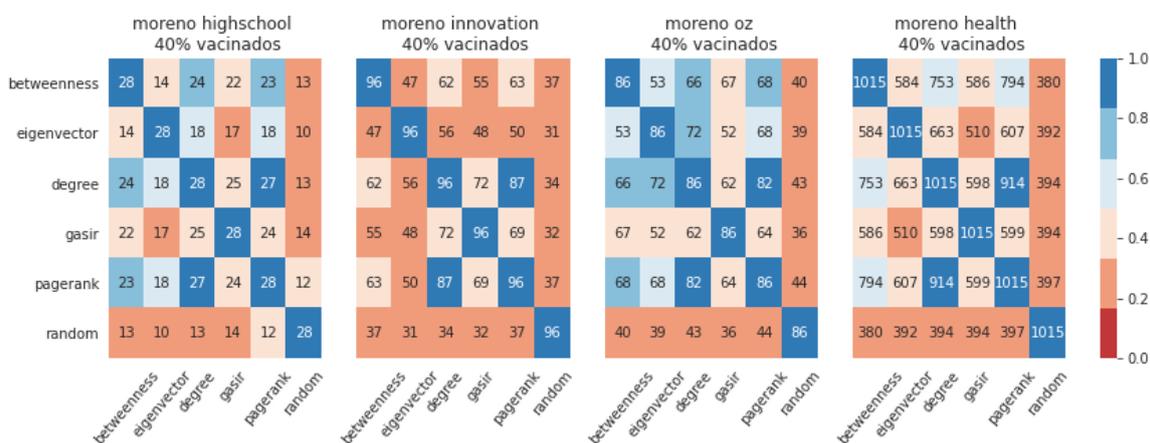


Figura 3. Similaridade e intersecção dos critérios de vacinação nas redes.

É possível concluir que o critério C_{pv}^G possui uma natureza própria na escolha dos indivíduos caracterizada pela diferença ordinal da escolha dos vértices. Tal interpretação é consequência da menor similaridade *rbo* quando comparada a outros critérios. Embora haja menor similaridade, espera-se que os resultados de epidemias simuladas sejam semelhantes aos dos critérios C_{pv}^D e C_{pv}^P , exceto na rede *moreno health*, dado o número de indivíduos comuns.

Os critérios C_{pv}^P e C_{pv}^D são os que mais se assemelham quando comparados em todas as redes experimentadas. Espera-se que o desempenho desses critérios seja semelhante.

4.2. Análise estatística das variáveis dependentes

Testes preliminares indicaram que, dentre as 504 amostras, em menos de 40% delas os dados das variáveis dependentes tinham aderência à normalidade. A avaliação foi realizada através do teste de *Shapiro-Wilk*, onde a rejeição da hipótese nula H_0 indica que as amostras de dados não seguem uma distribuição normal.

Dado tal comportamento observado nas amostras, decidiu-se pela utilização da abordagem estatística não paramétrica para avaliação da equivalência entre os diferen-

tes critérios de vacinação. Assumiu-se que a mediana e a amplitude interquartil são as medidas consideradas na comparação das amostras.

O teste de Kruskal-Wallis foi aplicado sobre as amostras agrupadas por rede, percentual de vacinação p_v e critérios de vacinação na forma C_{p_v} de cada uma das três variáveis dependentes estudadas. Ele verifica a hipótese nula H_0 de que as medianas das amostras agrupadas possuem equivalência estatística para uma mesma variável dependente. Para esta análise foram removidos os critérios C^\emptyset e $C_{p_v}^R$ por estes existirem apenas para fins comparativos de quão é acertada a decisão de se aplicar algum critério sistematizado para vacinação.

Obeve-se como resultado a rejeição da hipótese nula H_0 a um nível de significância de 0.05, com grau de liberdade 4 para todos os 100 agrupamentos (4 redes, 5 C_{p_v} , 5 p_v) de cada uma das três variáveis dependentes. Isso representa que existe diferença estatística entre os critérios de vacinação propostos baseados em centralidade e o AG. A checagem dessa hipótese sinaliza que pelos menos dois critérios de vacinação adotados apresentam diferença estatística. Para avaliação dos critérios aos pares, aplicou-se um teste de *Mann-Whitney*.

Não é observada uma relação uniforme entre as variáveis dependentes analisadas. Alguns resultados da variável *afetados* geraram altos valores de amplitude interquartil, deixando praticamente indefinido qual o melhor critério. As diferenças foram então observadas através das variáveis *pico* e *beta* que indicam o achatamento da curva e a diminuição da taxa de contágio.

O comportamento observado nas variáveis dependentes consideradas sugere similaridade entre os critérios $C_{p_v}^P$ e $C_{p_v}^D$. Tal observação advém das frequentes indicações de equivalência estatística entre os critérios nos testes de *post-hoc*, especialmente nas redes *moreno highschool*, *moreno innovation* e *moreno oz*. O critério $C_{p_v}^B$ é semelhante somente na rede *moreno highschool*.

O número de indivíduos que se mantiveram suscetíveis nessas redes varia bastante e as maiores taxas observadas foram de 40% na rede *moreno highschool* e de 41.49% na rede *moreno innovation*.

Na rede *moreno health* observou-se que os experimentos não apresentaram indícios de equivalência estatística entre os critérios aplicados. A maior fração da população mantida como suscetível foi de 15% e a maioria dos melhores resultados foram entregues pelo critério $C_{p_v}^P$. Vale ressaltar que os critérios $C_{p_v}^B$, $C_{p_v}^D$ e $C_{p_v}^G$, embora não sejam equivalentes, também são competentes na contenção e desaceleração da epidemia nesta que é a maior rede experimentada.

A Figura 4 ilustra que, nas redes *moreno oz* e *moreno health*, o percentual de indivíduos que se mantiveram suscetíveis após a simulação da epidemia são visualmente inferiores quando comparados às outras duas redes. Esta observação sugere que o aumento do grau médio da rede impacta inversamente no número de indivíduos protegidos com os critérios adotados.

A Figura 5 mostra que na rede *moreno oz*, a variável de *afetados* apresentou uma contagem elevada de indivíduos, igual ou muito próxima do tamanho da rede. A rede, que possui o mais alto grau médio e um dos maiores coeficientes de agrupamento, cria

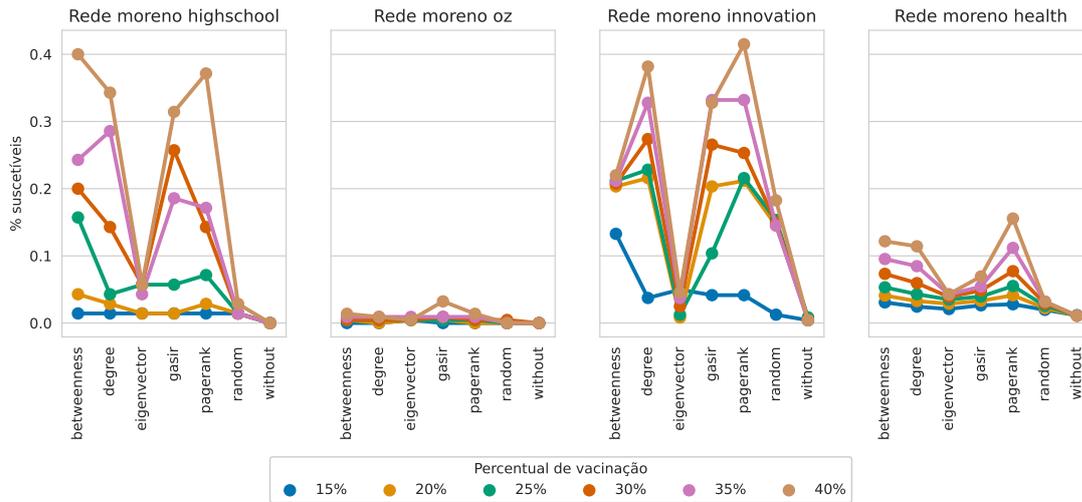


Figura 4. Percentual da população que se manteve suscetível após a aplicação dos critérios de vacinação. Valor estimado a partir da mediana dos experimentos.

a tendência de se atingir mais facilmente todos os indivíduos, o que dificulta em muito a contenção da epidemia. Contudo, as medianas e intervalos interquartis das variáveis *beta* e *pico* vistos nos *boxplots* demonstram que houve um decréscimo da intensidade da epidemia.

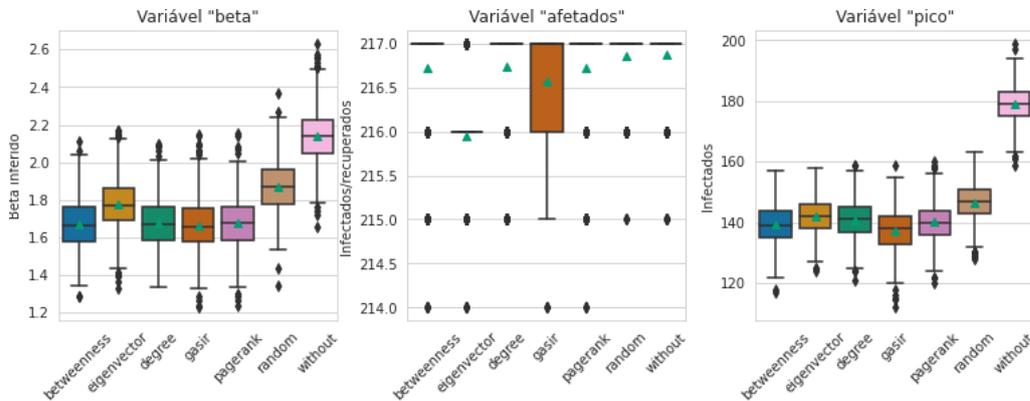


Figura 5. Boxplots com resultados das três variáveis dependentes com vacinação de 15% na rede *moreno oz*.

Na rede *moreno innovation* foi possível verificar que na variável *afetados* o critério C_{pv}^B apresenta *outliers* com valores inferiores às medianas observadas nos critérios C_{pv}^G , C_{pv}^D e C_{pv}^P , como ilustra a Figura 6. Tal observação indica que o ponto inicial de contágio na rede pode determinar que o critério C_{pv}^B seja mais ou menos efetivo.

Os diferentes níveis de complexidade topológica das redes demonstram que as estruturas impactam no sucesso dos critérios de vacinação, fator que determina a importância da análise de outras características do espalhamento, como a taxa de contágio e o pico de infectados.

A Tabela 2 exibe os melhores critérios nas redes avaliados por sua mediana e classificados segundo cada uma das variáveis dependentes consideradas neste estudo.

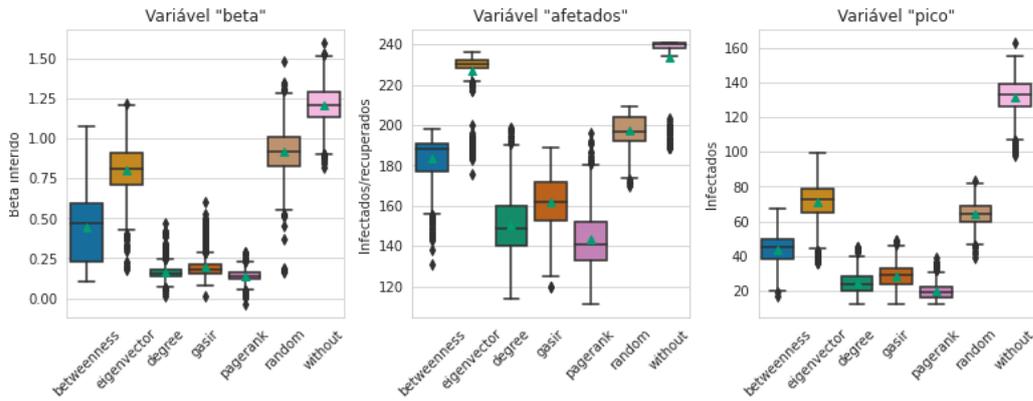


Figura 6. Boxplots com resultados das três variáveis dependentes com vacinação de 40% na rede *moreno innovation*.

É possível concluir que, em todas as redes estudadas, os resultados obtidos pelo critério C_{pv}^G são bons e estão normalmente a margem dos critérios C_{pv}^P e C_{pv}^D . É possível observar sua seleção como o melhor critério em algum momento, mas em muitos desses cenários existe equivalência estatística com outro critério baseado em centralidade. Outra observação importante é que C_{pv}^E é o critério sistemático de pior desempenho para as redes experimentadas, em muitas vezes seu desempenho se assemelha ao critério C_{pv}^R .

5. Considerações finais

Este trabalho analisou um conjunto de experimentos que buscou averiguar a eficácia de diferentes critérios de vacinação em diferentes redes complexas.

A primeira investigação analisou o quão equivalentes são os conjuntos de indivíduos selecionados pelos critérios em cada percentual de vacinação avaliado. Utilizou-se a medida de similaridade *rbo* e o número de indivíduos comuns entre os conjuntos. Os indivíduos selecionados pela centralidade *pagerank* e grau apresentaram constantemente um alto número de indivíduos comuns, assim como níveis de similaridade superiores a 0.5. Diversos indivíduos selecionados pela metodologia AG são comuns aos escolhidos por esses critérios de centralidade, mas o nível de similaridade é mais próximo de zero. Isso indica que o AG avaliou a importância dos seus indivíduos de forma distinta dos critérios *pagerank* e grau. É possível concluir que, normalmente, o AG apresenta resultados equivalentes a essas duas centralidades em simulações de contágio na rede.

Viu-se que os critérios *betweenness* e *eigenvector* possuem intersecção e níveis de similaridades maiores apenas em algumas redes, evidência de que estas centralidades podem apresentar resultados mais situacionais, conforme a estrutura topológica da rede.

A estrutura topológica da rede possui um papel importante no processo de disseminação, visto que as variáveis dependentes estudadas podem apresentar melhores resultados com critérios diferentes para a mesma rede e percentual de vacinação. Em redes de maior complexidade viu-se que minimizar o número de indivíduos afetados pode ser um trabalho difícil, mas que é possível reduzir a taxa de contágio, como visto nos resultados providos pelas variáveis dependentes *beta* e *pico*.

Por fim, observa-se que o critério que utiliza a centralidade *pagerank* apresenta com maior frequência os melhores resultados na avaliação das variáveis dependentes

Tabela 2. Melhores critérios avaliados pela mediana em redes reais.

Rede	pv	Pico	Equivalentes	Beta	Equivalentes	Afetados	Equivalentes
health	0.15	pagerank		pagerank		betweenness	
health	0.20	pagerank		pagerank		pagerank	C^B
health	0.25	pagerank		pagerank		pagerank	
health	0.30	pagerank		pagerank		pagerank	
health	0.35	pagerank		pagerank		pagerank	
health	0.40	pagerank		pagerank		pagerank	
highschool	0.15	degree	C^E, C^G, C^P	gasir	C^P, C^D	betweenness	C^D, C^G, C^P
highschool	0.20	pagerank		pagerank		betweenness	
highschool	0.25	betweenness	C^D, C^G, C^P	gasir	C^P, C^D	betweenness	
highschool	0.30	gasir	C^B	betweenness	C^G	gasir	C^B
highschool	0.35	degree	C^B	degree	C^B	degree	
highschool	0.40	betweenness	C^P	betweenness	C^P	betweenness	C^P
innovation	0.15	gasir	C^P, C^D	degree	C^G, C^P	betweenness	C^D, C^G, C^E, C^P
innovation	0.20	degree	C^P	degree	C^P	degree	C^P
innovation	0.25	gasir		gasir		degree	C^B
innovation	0.30	degree	C^G, C^P	degree	C^G, C^P	degree	C^G, C^P
innovation	0.35	pagerank		gasir	C^P, C^D	gasir	C^P, C^D
innovation	0.40	pagerank		pagerank		pagerank	
oz	0.15	gasir		betweenness	C^D, C^G, C^P	eigenvector	
oz	0.20	gasir	C^B	gasir	C^P, C^B, C^D	betweenness	
oz	0.25	gasir		gasir	C^P	betweenness	C^E
oz	0.30	gasir		pagerank	C^G	gasir	
oz	0.35	betweenness	C^G, C^P	gasir	C^P, C^B	betweenness	C^E
oz	0.40	gasir		pagerank	C^B	gasir	

na maioria das situações experimentadas. Contudo, os resultados equivalentes entre os critérios *pagerank* e grau demonstram que o último, que demanda baixo custo computacional para cálculo, pode ser um alternativa viável para redes muito grandes.

A análise comprovou que, sob o ponto de vista das três variáveis dependentes analisadas, os resultados obtidos pela metodologia utilizando AG são, muitas vezes, similares aos critérios que utilizavam a centralidade por grau e *pagerank*. Tal comportamento corrobora com a expectativa gerada pela análise da similaridade e da intersecção dos indivíduos selecionados pelos critérios. Com isso não se viu vantagem competitiva para a metodologia baseada em um AG com esse conjunto de redes, devido ao seu alto custo computacional para a escolha dos indivíduos.

A principal limitação deste trabalho é que o ambiente experimentado, com redes pequenas, dista o experimento de uma aplicação prática. O aumento da rede gera um custo computacional alto para análise e experimentação, mas é desejável para uma melhor avaliação do critério proposto e poderia fornecer conhecimento para aplicações reais.

Referências

- Barabasi, A.-L. (2016). Network Science. Cambridge University Press.
- Bonacich, P. (1987). Power and centrality: A family of measures the american journal of sociology, vol. 92. American Journal of Sociology.
- Bresee, J. S., Fiore, A. E., Fry, A., Gubareva, L. V., Shay, D. K., and Uyeki, T. M. (2011). Antiviral agents for the treatment and chemoprophylaxis of influenza: Recommendations of the advisory committee on immunization practices (acip). Morbidity and Mortality Weekly Report, 60:1–24.

- Coburn, B. J., Wagner, B. G., and Blower, S. (2009). Modeling influenza epidemics and pandemics: Insights into the future of swine flu (h1n1). BMC Medicine, 7:30.
- Freeman, L. C. (1978). Centrality in social networks conceptual clarification. Social Networks, 1(3):215–239.
- Holland, J. H. (1975). Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence. University of Michigan Press, Ann Arbor.
- Li, X., Mukandavire, C., Cucunubá, Z. M., Londono, S. E., Abbas, K., Clapham, H. E., Jit, M., Johnson, H. L., Papadopoulos, T., and Vynnycky, E. (2021). Estimating the health impact of vaccination against ten pathogens in 98 low-income and middle-income countries from 2000 to 2030: A modelling study. The Lancet, 397(10272):398–408.
- Li, Z., Zhu, P., Zhao, D., Deng, Z., and Wang, Z. (2019). Suppression of epidemic spreading process on multiplex networks via active immunization. Chaos: An Interdisciplinary Journal of Nonlinear Science, 29(7):073111.
- Liu, L., Luo, X., and Chang, L. (2017). Vaccination strategies of an sir pair approximation model with demographics on complex networks. Chaos, Solitons & Fractals, 104:282–290.
- Ndiaye, B. M., Tendeng, L., and Seck, D. (2020). Comparative prediction of confirmed cases with covid-19 pandemic by machine learning, deterministic and stochastic sir models. arXiv:2004.13489 [math, q-bio, stat].
- Newman, M. E. (2002). Assortative mixing in networks. Physical review letters, 89(20):208701.
- Oliveira, R. S., Xavier, C. R., Vieira, V. d. F., Reis, R. F., dos Santos, Quintela, B. d. M., Rocha, Martins, B., and Weber, R. (2021). How fast vaccination can control the covid-19 pandemic in brazil? In International Conference on Computer Science 2021.
- Page, L., Brin, S., Motwani, R., and Winograd, T. (1999). The pagerank citation ranking: Bringing order to the web. Technical report, Stanford InfoLab.
- Peixoto, T. P. (2020). The netzschleuder network catalogue and repository.
- Rodrigues, R. F., da Silva, A. R., da Fonseca Vieira, V., and Xavier, C. R. (2018). Optimization of the choice of individuals to be immunized through the genetic algorithm in the sir model. In ICCSA, pages 62–75. Springer.
- Rouquayrol, M. Z. and Filho, N. d. A. (2003). Epidemiologia & saúde. MEDSI, 6th edition.
- Shulgin, B., Stone, L., and Agur, Z. (1998). Pulse vaccination strategy in the sir epidemic model. Bulletin of Mathematical Biology, 60(6):1123–1148.
- Tanner, M. W., Sattenspiel, L., and Ntaimo, L. (2008). Finding optimal vaccination strategies under parameter uncertainty using stochastic programming. Mathematical Biosciences, 215(2):144–151.
- Webber, W., Moffat, A., and Zobel, J. (2010). A similarity measure for indefinite rankings. ACM Transactions on Information Systems (TOIS), 28(4):1–38.