

# Simulation of the Effects of Targeted Immunization on the Peak Number of Infections in Complex Networks

Michael Emmerich<sup>1,\*</sup>, Yulian Kuryliak<sup>2,\*</sup> and Dmytro Dosyn<sup>2</sup>

<sup>1</sup>*Leiden Institute of Advanced Computer Science, LIACS Leiden University, Niels Bohrweg 1, 2333CA Leiden, The Netherlands*

<sup>2</sup>*Institute of Computer Science and Information Technologies, ICSIT of Lviv Polytechnic National University, 12 Stepan Bandera street, 79000 Lviv, Ukraine*

## Abstract

In this paper we study the effect of targeted immunization on the peak number of infections in an epidemic outbreak. For this we extend a previously developed python-based dashboard environment for the time efficient simulation-based study of SIR epidemics spread on complex network topologies, using realistic Continuous Time Markov Chain (CTMC) simulations by means of Gillespie's stochastic simulation algorithm. The new components make it possible to study targeted immunization by means of state-of-the-art methods and to visualize typical paths of infection during the temporal evolution of an epidemic. We show results obtained with different centrality measures (eigenvalue centrality of the adjacency and non-backtracking matrix, degree centrality, and average path length centrality), used in targeted immunization. In the results we focus on studying the peak number of infections (PNI). The PNI is very relevant when it comes to the practical management of an epidemic, as it determines, for instance, the number of intensive care units that are needed to offer an appropriate treatment of the disease in critical cases. However, the PNI has received much less attention in studies than the epidemic threshold, which is more relevant in the early stage of an epidemic. Our example study on classical network topologies reveal that the choice of the centrality measure for targeted immunization as well as the number of targeted nodes will have a strong impact on the drop of the PNI. Our simulation-based results on scale-free Barabasi-Albert networks show that the PNI reduction that can be achieved by using modern centrality metrics such as the non-backtracking eigenvalue drop, can lead to up to 40% lower peaks than those achieved with naïve methods such as degree based immunization (immunization of the biggest node(s)) in case of immunization of 2.5% nodes. These results underpins the crucial importance of the correct choice of the centrality metric in targeted immunization.

## Keywords

Targeted immunization, Non-backtracking matrix, COVID-19 epidemic, Network science, Peak number of infected nodes, Gillespie's algorithm, Continuous Time Markov Chains

## 1. Introduction

Effective resistance against virus disease spreading could be organized by taking into account all or at least most of the influenced conditions. The level of such influence should be estimated by

---

*MoMLeT+DS 2022: 4th International Workshop on Modern Machine Learning Technologies and Data Science, November, 25-26, 2022, Leiden-Lviv, The Netherlands-Ukraine*


\*These authors contributed equally.

EMAIL: m.t.m.emmerich@liacs.leidenuniv.nl (M. Emmerich); yulian.kuryliak.mnsam.2021@lpnu.ua (Y. Kuryliak); dmytro.h.dosyn@lpnu.ua (D. Dosyn)

ORCID: 0000-0002-7342-2090 (M. Emmerich); 0000-0002-6600-2637 (Y. Kuryliak); 0000-0003-4040-4467 (D. Dosyn)



© 2022 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

 CEUR Workshop Proceedings (CEUR-WS.org)

modeling and simulation as realistically and efficiently as possible. Of particular interest is the influence of the topology of the contact network and the impact of targeted immunization on the dynamics of an epidemic. In the COVID-19 pandemic, it became very clear that immunization should not just target the total spread of a virus in the population, or the epidemic threshold, but also seek to decrease peak numbers of infections (PNI)[1], because of the limited availability of intensive care units. If this indicator is neglected, there is a risk of the health system runs out of capacity resulting in unwanted 'triage' situations where not all patients can receive proper treatment.

However, in previous studies, the reduction of the epidemic threshold was mainly the focus of studies. The epidemic threshold should be as high as possible and it determines whether an epidemic is likely to invade the network, which is the case when the contagiousness exceeds the epidemic threshold, or, otherwise, tends to disappear before invading the network. Various measures have been suggested in the literature to increase the epidemic threshold by means of targeted immunization. As was argued in [2], the decreasing eigenvalue better increases the epidemic threshold than removing nodes with the highest degree. In addition the targeted immunization can be associated with the cost, which is given by the (weighted) costs for node-removal. Multiobjective Quadratic Programming Formulation for decreasing eigenvalue with the lowest cost for  $k$  nodes [3]. For the SIS model a good method of immunization is decreasing the maximum eigenvalue of the adjacency matrix[2], while for the SIR model it has been argued that better results can be achieved by replacing the eigenvalue of the adjacency matrix by that of the Hashimoto's non-backtracking (HNB matrix) of a network[4]. The latter is motivated by the impossibility of re-infection of the same node, once recovered. Ideally, the  $k$  non-cycling matrix[5] (of which the HNBM is a special case with  $k = 1$ ) models this situation of impossible re-infections better than the HNBM, but although it gives slightly better results, the computations get much more complex, especially for large  $k$  and networks with many cycles. Therefore, using the non-backtracking matrix is a good compromise in SIR epidemic targeted immunization, when the goal is to decrease the epidemic threshold. In [4] described the non-backtracking centrality as a computationally efficient proxy to the largest eigenvalue of the non-backtracking matrix.

Despite the big interest of the research community in methods that target the increase of the epidemic threshold, there are relatively few studies that focus directly on the peak number of infections [6]. Moreover, in existing studies, such as [6], the conclusions hold only asymptotically (in the size of the network and number of infections in the network) and small and moderate size networks where the exact number of infections matters were neglected. This is due to the techniques used that rely on differential equations and assume a continuous variable for the number of infections. In addition the structure of the complex networks is often oversimplified to a kind of random network where all nodes share the same number of contacts, which is known to be unrealistic, given that real world networks rather follow a power law in their degree distribution.

The PNI for such moderate size networks with a power law degree distribution can, however, be studied with Gillespie's stochastic simulation algorithm (SSA) which was originally designed for chemical reaction systems with a moderate, finite number of molecules [7]. It is based on a Continuous-time Markov chain that models the progression of the state of the network over time, while, as opposed to discrete-time Markov chains, modeling time durations in a

realistic manner. Moreover, it simulates efficiently and without idle time the exact numbers of infections at certain time-steps. That is the number of simulated time steps equals the number of state-changes in the network. The transfer of Gillespie's algorithm to the simulation of SIR epidemics and the underlying Markov chain models were explained in [8, 9].

To simulate the event's PNI, in previous work we developed an efficient method of stochastic simulation of epidemic outbreaks adopting Gillespie's stochastic simulation algorithm and exploiting specifics of the epidemic outbreak to avoid computing the full transition matrix. Whereas the dimension of the full transition matrix (or generator matrix) of the CTMC for the state of infections in complex networks grows exponentially with the number of nodes, it suffices to focus on non-zero elements in this matrix, the number of which is only linearly depending on the number of nodes in the network. This dashboard can simulate networks with several thousand nodes as outlined in [8] and visualize the virus propagation up to a few hundred. The previous version of the full dashboard software with the integrated simulation algorithm was described in [10]. In the current version, the possibility to suggest nodes for **targeted immunization** was added and the visualization of infection paths was improved and augmented to aggregate several simulation runs.

## 2. Method of the simulation

Next, we will briefly outline the details of the simulation method before proceeding with the topic of targeted immunization.

### 2.1. Epidemiological Model

There are three main epidemiological models: SI, SIS and SIR. The models allow every node of a network to be in one of the following states respectively:

- S - susceptible;
- I - infected. Such node has symptoms of a disease and is able to infect other nodes;
- R - removed (immunized). An immunized node is not a disease carrier.

In addition, our system allows taking into account simply infected (people who can be treated at home or who have no noticeable symptoms) and critically infected (people who need medical assistance). There are a number of other epidemiological models that take into account the incubation period, detected and undetected infected nodes, the possibility to return from immunized (removed) state to susceptible, etc. The main problem of such models to accurately simulate their impact on complex networks is lack of data about a disease. It is assumed here that the study of SIS and SIR dynamics can already reveal major trends in the epidemic dynamics effectiveness of targeted immunization. It remains, however, an important topic for future research whether or not the inclusion of further details of the disease progression have a major impact on the resulting dynamics.

Furthermore, in this study we show results on diseases that, after a successful recovery, result in the immunity of the infected individuals. Although in some diseases such as the COVID-19

outbreak (that was the motivating example for our study), this is debated among experts, we assume that for networks of moderate size the immunity to the disease after recovering is long enough to assume that nodes do not get infected again until the end of the epidemic outbreak. Therefore the SIR model is more suitable for this study than SIS.

## 2.2. Continuous-Time Markov Chain Simulation

For modeling the epidemic process we use Continuous Time Markov Chains (CTMCs). The advantage of using CTMCs over Markov Chains with discrete time is that CTMCs allow realistic modeling of time. Moreover, they are not based on asymptotical simplifications and stability of mean values as do the classical epidemiological models based on differential equations [6] and, respectively, mean-field models. One of the main reasons why CTMCs are not often used for such simulations is the explosion of the state space. To avoid it, we use the stochastic Gillespie algorithm [7], according to which the simulation of the time between the transitions of two states and the simulation of the next state is divided. During the transition from one state to another, only one additional node can be infected.

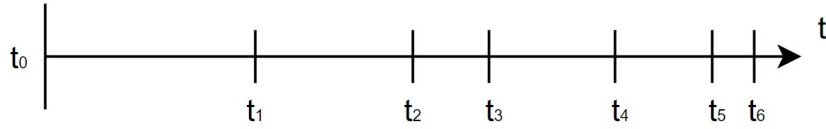
In a straightforward implementation of the CTMC model of the contact process all rates at which transitions occur between network states would be described in a generator matrix  $Q$ . Here the state of the entire population forms a row-column of that matrix. There are  $2^N$  possible population states ( $N$  nodes that can be infected (0) or non-infected/susceptible (1)), leading to a  $2^N \times 2^N$  matrix, where  $N$  denotes the number of nodes in the contact network. A matrix component  $q_{ij} \in \mathbb{R}^N$ ,  $i \neq j$  denotes the transition rates, and  $q_{ii} = -\sum_{i \neq j} q_{ij}$  is the rate of leaving state  $i$  (on the diagonal). To avoid exponential state-space, we proposed in [8] a way of avoiding the representation of the full generator matrix but to focus only on the  $O(N)$  non-zero rates in the matrix, that is rates to a state where a single additional node (as compared to the previous state) gets infected or where a node that was previously infected is treated or removed from the network. Let us denote with  $i$  the current state of the network and  $i_{+j}$  denote a network state where node  $j \in [1, N]$  is the index of a susceptible node that potentially gets infected in addition to the previously infected nodes in state  $i$ . Then the non diagonal state transitions  $q_{ii_{+j}}$ ,  $i \neq i_{+j}$  are computed as follows:

$$q_{i, i_{+j}} = \begin{cases} \lambda v(j), & \text{if node } j \text{ is susceptible;} \\ 0, & \text{if node } j \text{ is not susceptible.} \end{cases} \quad (1)$$

Here  $\lambda$  is the contagiousness of the virus,  $v(j)$  denotes the number of infected neighbors of the vertex  $j$  in state  $i$ . In the theory of the stochastic simulation of CTMC [7] it is a well known law that the transition time from one discrete state to the next discrete state of the system follows an exponential distribution with a probability density function (PDF) denoted with  $F$  defined as:

$$F(x) = \begin{cases} \theta e^{-\theta x} & \text{if } x \geq 0; \\ 0, & \text{otherwise} \end{cases} \quad (2)$$

where  $\theta = -q_{ii}$ . The expected value of the time until the next infection event is thus calculated by  $1/\theta$  for a single link, and in a given state of the network in the CTMC by  $\frac{1}{\lambda q_{ii}}$ .



**Figure 1:** Example of expected time

Next to the time duration between two events, we need to also simulate the type of event that occurs in the next transition, which is decoupled from the duration[7]. The probability of transition from state  $i$  to state  $j$ , that is  $p_{ii+j}$  is determined by the formula:

$$p_{ii+j} = \frac{q_{ii+j}}{\sum_{l,l \neq i} q_{ii+l}} \quad (3)$$

and it can be simulated by “roulette wheel” simulation [11].

### 3. Method of targeted immunization

Targeted immunization strategies select nodes that are to be quarantined, immunized or removed from the network in order to achieve a maximum desirable effect on the dynamics of the epidemic, in our case the maximum reduction of the peak number of infections.

#### 3.1. Degree-drop

The most common method of immunization is the immunization of nodes with the highest degree of a network  $max(k_i)$ . The degree of a node  $k$  is considered to be the number of its connections (edges) with other nodes. In the case of a directed graph, the degree of a node is divided into input  $k_{in}$  and output  $k_{out}$ , which can also be used for a certain type of contacts of network individuals.

Using this method of targeting, it is possible to select  $n$  nodes with the highest degrees, or apply iterative selection with the calculation of the highest degree after removing each of the nodes, since the degrees of the nodes change after each immunization of a node.

Relatively early, and based on differential equation models, it was argued that the number of contacts of a node in the network (degree) is not a reliable indicator for selecting nodes for targeted immunization[12]. Various alternative centrality measures for node-selection were subsequently suggested:

#### 3.2. Eigenvalue drop

The largest eigenvalue of a network is a measure that originated from the spectral theory of systems where it is also known as the spectral radius of a graph. From the point of view of the epidemiology, the eigenvalue characterizes the possibility of the successful occurrence of an epidemiological outbreak in a network and can be roughly characterized at the speed of expansion of a linear system along the direction of its principal eigenvector. The relation of the

dominant eigenvalue to the epidemic threshold  $\theta$  was described in [12] and it is determined by the following formula:

$$\frac{\beta}{\gamma} < \tau = \frac{1}{\lambda},$$

where  $\beta$  is the virus birth rate and  $\gamma$  is the virus death rate, and  $\lambda$  is the highest eigenvalue of the adjacency matrix. It is important to pay attention to the highest eigenvalue, because the measure takes into account not only the degree of a node, but also the entire connectivity of the network, including the degree of its neighbors. Therefore, immunization by the method increases the distance between hubs and breaks the network into clusters from which it is more difficult for the disease to escape. Depending on the type of epidemic outbreak - SIS or SIR - the eigenvalue is either computed for the adjacency matrix or, respectively, for the HNB matrix.

### 3.2.1. Adjacency matrix

In [12] it was shown that for SIS dynamics the maximum eigenvalue of the adjacency matrix of the contact network is inversely proportional to the epidemic threshold [12], [2]. Therefore it has been suggested to drop this eigenvalue maximally by removing nodes [3]. The Netshield proxy function was introduced to accomplish this efficiently (see, e.g., [4, 3]).

### 3.2.2. Non-backtracking matrix

The HNB matrix was considered first by Hashimoto for message-passing systems in complex networks and later found its application in the study of epidemic outbreaks of the SIR type [4]. We assume therefore that using Hashimoto's non-backtracking matrix might give a better approximation for targeted immunization in the case of the SIR model, whereas nodes during a simulation will not be able to infect a node by which they were infected.

The HNB matrix is defined by following formula:

$$B_{(i \rightarrow j), (k \rightarrow \ell)} = \begin{cases} 1, & \text{if } j = k \text{ and } i \neq \ell \\ 0, & \text{otherwise} \end{cases} \quad (4)$$

Thus the HNBM is the adjacency matrix of a network where the edges of the original network serve as nodes. A pair of edges  $u$  and  $v$  is adjacent (connected), if and only if the edges are adjacent that the node for which  $u$  is the incoming edge has the outgoing edge  $v$ , and  $v$  does not point back to the node from which  $u$  is the outgoing edge. In some cases the HNB matrix has no positive eigenvalues. Based on our observations this does not occur if every node of the original network has at least one incoming edge and one outgoing edge (the start node of the incoming edge is not equal end node of the outgoing edge in the case of the non-backtracking matrix). In the current software we demand the network to have positive non-backtracking matrix eigenvalues, but further analysis is needed to figure out when exactly this is the case.

## 3.3. Average Shortest Path Length

The shortest path length is the number of edges that need to be traversed to get from node  $i$  to node  $j$ . Average Shortest Path Length (ASPL) is the average of the shortest path lengths from each node to each.

The average shortest path length turned out to be an effective measure for reducing the PNI in COVID-19 dynamics, when considering multiple-node immunization strategies. This has been shown in our previous work on the COVID-19 outbreak in Ukraine [9]. This also seems plausible, since in networks with a small average shortest path length between pairs of nodes the propagation of the virus can proceed very fast. Therefore, we include increasing the average shortest path length here as an immunization strategy.

## 4. Dashboard

In order to better understand the process of the spread of a virus in the network we created a dashboard, an earlier version of which was described in detail in [10]. This dashboard provides a means to run a simulation with a user-defined network and parameters of the particular epidemic, to visualize the virus propagation process over time, and plot the (averaged) results of the simulation(s). Three types of settings are distinguished:

- Simulation parameters: maximum simulation time (simulation days), time(simulational) step for collecting data from a simulation, and the number of repeated simulations with given parameters. Repeated simulations are required to test for significant effects and understand the average, best-case and worst-case behavior of the stochastic process at hand.
- Network parameters: network model (complete graph, Erdős–Rényi, Watts-Strogatz, and Barabási–Albert models), number of nodes in the network, and number of edges per node
- Epidemiological parameters: epidemiological model, infection rate, period of being in infected and critically infected states. As a default data on demographic distribution and mortality statistics is the statistic on COVID-19 in Ukraine from 19 December 2020 and can be changed with replacing files by path `./scripts/datasets/`.

New in this version of the dashboard are options for nodes immunization. For this, a user is given the opportunity to enter the number of nodes for immunization and one of the following methods of immunization:

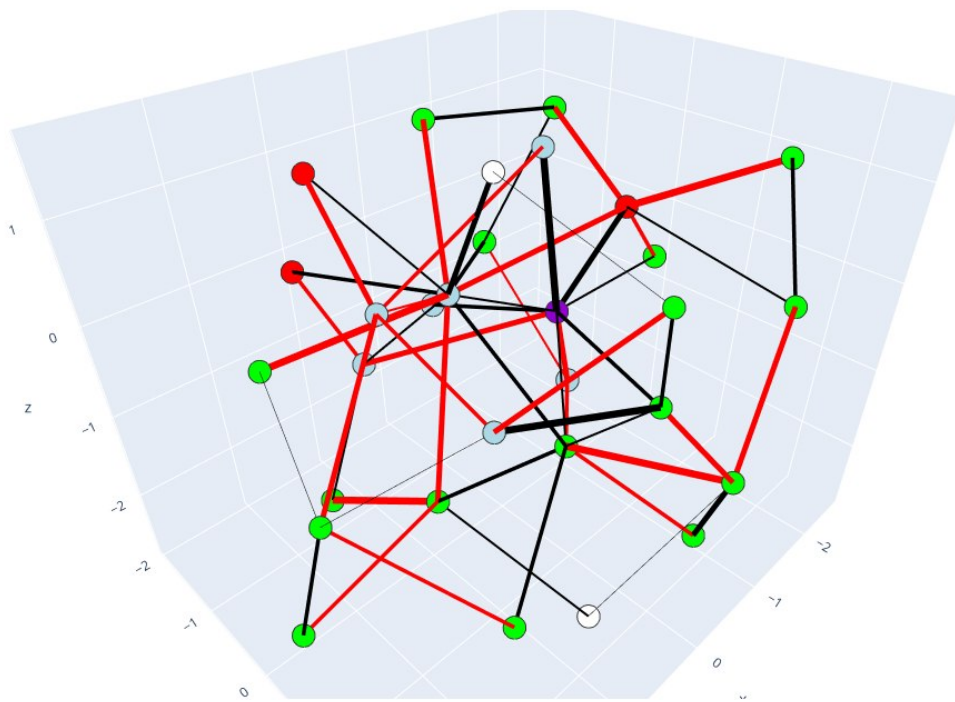
- Decreasing of the highest degree;
- Increasing of the average shortest path length;
- Maximum eigenvalue-drop of Adjacency matrix;
- Maximum eigenvalue-drop of Non-backtracking matrix;

According to chosen method, the user-defined number of nodes will be immunized using a greedy algorithm.

The network is presented in the form of an interactive three-dimensional graph (Fig.2), where each of the nodes is a person and each edge is a contact. The initial color of the nodes is white and the edges are black. When a node is infected, it turns green, and if the node is critically infected, it turns red. After recovery, the node turns blue, or purple if it was critically infected.

When infection occurs, the edge through which it was performed turns red. In the case of multiple simulations for given parameters, the thickness of the edge increases each time an infection is performed through it. The colors of all elements at each step are taken from the last simulation run. Also, it is possible to go to each of the steps using a slider, which is thinned under the graph.

Time : 16



**Figure 2:** Graph visualization

In addition to the visualization of the network, for the last simulation, a plot is built that displays the number of nodes in each of the states (susceptible, infected, critically infected, infectious and dead) during the simulation with a given time interval. The plot, like the graph is interactive, therefore user can move, change scale and disable curves that they do not need.

## 5. Experiments

For the experiments, 4 methods of immunization were chosen: decreasing the highest degree (HD), increasing the highest average shortest path length (ASPL), decreasing the highest eigenvalue of the adjacency matrix (EAM), and decreasing the highest eigenvalue of the non-backtracking matrix (ENBM). By every method of targeted immunization, the nodes are removed

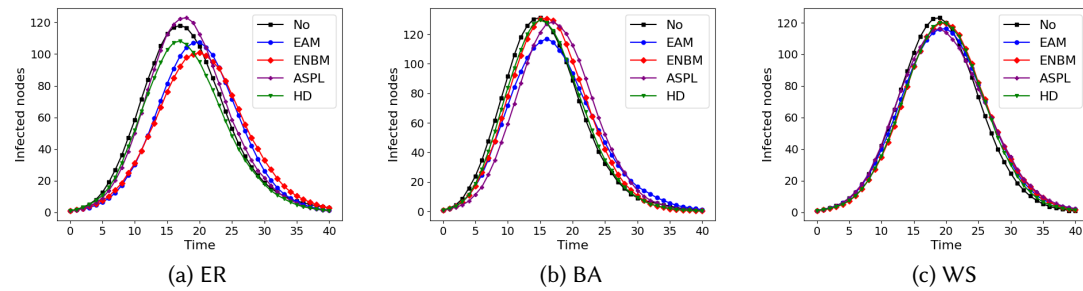


in a greedy manner that is nodes are removed subsequently and after each removal, the centrality measures are recomputed.

Note that greedy removal strategies do not always lead to the best result in multiple node removal when compared to an optimization of the removed subset. However, the optimization is usually very time expensive and limits the size of the networks that can be considered. The assessment of the gap between greedy and optimization-based subset selection remains however an interesting topic for future research.

As models of networks, three classical models of complex networks were chosen: Barabási-Albert, Watts-Strogatz (with reconnection rate of 0.3), and Erdős-Rényi model. Each network was undirected and contains 200 nodes and about 400 edges. The Erdős-Rényi model can be considered a base model in which each node has about the same number of contacts which is a common but unrealistic assumption in many previous studies on epidemic outbreaks. The Barabási-Albert network has a scale-free degree distribution and the Watts-Strogatz network features a small world structure with a small degree of separation [13].

For the infectious period, 10 simulation days were chosen. The results on the number of infected nodes on every timestep obtained and averaged from 50 simulations for each combination of the network model and method of nodes immunization on the Fig. 3 were compared with simulations in the corresponding network models without any immunization. The networks for immunization by each method were not exactly the same.

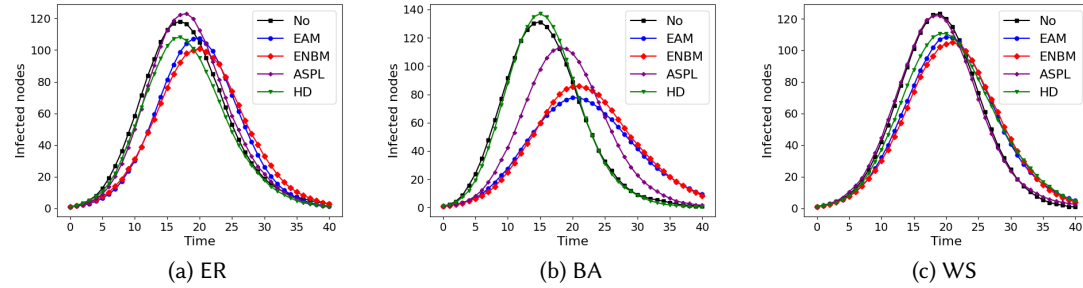


**Figure 3:** Comparison of plots of the number of infected nodes at each point of time for each immunization method

The observed in figure 3 results are quite controversial. Immunization of one node by none of the methods did not give significant results in reducing the PNI and delaying it. Only in the case of immunization by decreasing the eigenvalue of the non-backtracking matrix of ER network and also in the case of the eigenvalue of the adjacency matrix of the BA network we can observe reducing of PNI by about 10 percent. In addition, increasing the average shortest path length slightly decreased and postponed the PNI for all networks. According to facts that simulation is stochastic and for every simulation a new network was generated, there can be an error.

To understand the effect of immunization better, immunization of 5 nodes by greedy removal was performed on the same network models. The results collected on the same network models from at least 50 simulations are shown in figure 4. According to the results of the simulations,

decreasing the highest eigenvalue of an adjacency matrix provides a significant decrease and postpones the PNI in the case of degree-centralized, free-scale Barabas-Albert network and only in the case. Immunization by decreasing the highest eigenvalue of a non-backtracking matrix showed a considerable effect in the case of every network. On the highly clustered Watts-Strogatz network, the influence of immunization was the lowest. Also, we can make the conclusion that the greedy algorithm is not suitable for increasing the average shortest path length.



**Figure 4:** Comparison of plots of the number of infected nodes at each point of time for 5 immunized nodes by each method

The parameters of networks with every method of immunization and without any are given in table 1. The parameters are averaged from at least 50 networks in every case of immunization 1 and 5 nodes by each method and without any immunization.

**Table 1**  
Network properties

Network	nodes	nodes immunized	ASPL	EAM	ENBM	HD
Erdős–Rényi	200	0	3.896	5.191	3.954	10.472
Erdős–Rényi	200	1	3.902	5.024	3.804	10.38
Erdős–Rényi	200	5	3.916	4.581	3.416	10.523
Barabási–Albert	200	0	3.466	7.468	5.328	29.04
Barabási–Albert	200	1	3.693	6.496	4.512	27.62
Barabási–Albert	200	5	3.803	4.643	3.01	27.66
Watts–Strogatz	200	0	4.116	4.672	3.503	8.467
Watts–Strogatz	200	1	4.121	4.561	3.408	8.28
Watts–Strogatz	200	5	4.125	4.272	3.133	8.42

\*ASPL, EAM, ENBM, HD - network parameters.

\*\*The value of each property is obtained from the corresponding immunization method.

Time to a new event(infection)  $t_e$  generates randomly, according to the exponential distribution, the expected value of which is the number of contacts with infected nodes. The time is recalculated after every event(infection). Let's assume that there is a moment when in a network are not many contacts with infected nodes in the network (as usual, it happens at the beginning of the simulation), in this case, the time to an event  $t_e$  is relatively high, and might

be the case, that the highest time of recovering a node  $t_r$  is lower than  $t_e$ . So, there will be not any infected nodes at the time of happening a new event and the simulation stops. Let's call it failed outbreak.

*Condition of failed outbreak:* Total number of infected (TNI) nodes is relatively small,  $TNI \rightarrow 0$ .

Fractions of failed outbreaks for the experiments with 1 and 5 immunized nodes by each method are given in table 2 with the condition  $TNI < 0.2 * \text{number of nodes in the network}$ ).

**Table 2**

Fractions of failed outbreaks

Network	nodes	nodes immunized	No	ASPL	EAM	ENBM	HD
Erdős–Rényi	200	1	0.04	0.02	0.02	0.06	0.04
Erdős–Rényi	200	5	0.02	0.02	0.1	0.08	0.12
Barabási–Albert	200	1	0.04	0.02	0.04	0.0	0.04
Barabási–Albert	200	5	0.02	0.06	0.12	0.08	0.02
Watts–Strogatz	200	1	0.04	0.0	0.02	0.02	0.04
Watts–Strogatz	200	5	0.04	0.02	0.02	0.04	0.0

\*ASPL, EAM, ENBM, HD - immunization methods.

## 6. Conclusions and Outlook

This paper has reported progress in the simulative analysis of epidemic outbreaks following the SIR dynamics. The contributions reported in the paper were two-fold:

Firstly we provided a discussion of an extension of a dashboard for the stochastic simulation of epidemics complex networks. The extension includes the visualization of the trajectory of a single run, and the summarizing visualization for multiple runs. The latter highlights contacts in the network that are very likely to transmit the virus and the analysis of these could yield insights into some typical infection paths and how to block them.

Secondly we extended the dashboard with the possibility to assess targeted immunization strategies using state of the art centrality measures. This new feature was used to study on realistic network topologies the outbreak of an epidemic before and after targeted immunization. In contrast to existing studies that focus on the epidemic threshold, our analysis focused on the peak number of infections (PNI), which is a crucial indicator when it comes to the management of a health crisis caused by an epidemic. Interestingly, the choice of the targeted immunization strategy has a strong impact on the drop of the PNI. Naïve strategies that immunize merely by targeting "big" nodes with many contacts fall short in performance as compared to those that use Eigenvalue based immunization. The difference is up to 40% less infections when using the right immunization strategy as compared to no-immunization or degree based immunization in the averaged results in case of immunization of 2.5% nodes.

This paper is a rather theoretical study, but it has an important message for the practical research: Even if common sense might tell us that we should focus on nodes with many contacts first in immunization, the study shows that a significantly better effect can be achieved on the peak number of infections by using more sophisticated centrality measures than degree, namely

global centrality measures such as eigenvector based or path-based centralities that take into account the topology of the entire network. It will be an important topic for future research to validate this result also in real world studies. However, it is unethical to experiment with real situations, but instead we can make the simulation more realistic if data from real contact networks and on the virus/disease dynamics becomes available.

In this paper a shortcut is used to simulate the treatment of nodes, which does not use rates but a deterministic threshold duration after which nodes are considered to have been treated (and are removed from the network). It will be a topic of future research to make the treatment also a stochastic process governed by a rate in the CTMC model. Moreover temporal changes of the network, e.g., the establishment or disappearance of contacts in the network could provide a meaningful extension which also can be governed by using rates in the CTMC model. In all these extensions the CTMC modeling approach can be efficiently used, because of the state-space reduction that makes use of the fact that in a single iteration only a small number of state transitions occur at non-zero rates.

The source code of the system can be found by the link: <https://github.com/YulianKuryliak/EpidemicOutbreakPredictor>

## 7. Acknowledgements

Michael Emmerich acknowledges support for research on spread dynamics in complex networks in the scheme EU Horizon 2020 RISE SMA.

## References

- [1] P. Liu, P. Beeler, R. K. Chakrabarty, Covid-19 progression timeline and effectiveness of response-to-spread interventions across the united states, *MedRxiv* (2020).
- [2] Y. Wang, D. Chakrabarti, C. Wang, C. Faloutsos, Epidemic spreading in real networks: An eigenvalue viewpoint, in: *22nd International Symposium on Reliable Distributed Systems, 2003. Proceedings.*, IEEE, 2003, pp. 25–34.
- [3] M. Emmerich, J. Nibbeling, M. Kefalas, A. Plaat, Multiple node immunisation for preventing epidemics on networks by exact multiobjective optimisation of cost and shield-value, *arXiv* (2020). Doi:2010.06488.
- [4] L. Torres, K. S. Chan, H. Tong, T. Eliassi-Rad, Node immunization with non-backtracking eigenvalues, *arXiv preprint arXiv:2002.12309* (2020).
- [5] F. Arrigo, D. J. Higham, V. Noferini, Beyond non-backtracking: non-cycling network centrality measures, *Proceedings of the Royal Society A* 476 (2020) 20190653.
- [6] J. Huang, G. Qi, Effects of control measures on the dynamics of covid-19 and double-peak behavior in spain, *Nonlinear dynamics* 101 (2020) 1889–1899.
- [7] D. T. Gillespie, Stochastic simulation of chemical kinetics, *Annu. Rev. Phys. Chem.* 58 (2007) 35–55.
- [8] Y. Kuryliak, M. Emmerich, D. Dosyn, Efficient stochastic simulation of network topology effects on the peak number of infections in epidemic outbreaks, *arXiv preprint arXiv:2202.13325* (2022).

- [9] Y. Kuryliak, M. Emmerich, D. Dosyn, On the effect of complex network topology in managing epidemic outbreaks, in: 3rd International Workshop on Modern Machine Learning Technologies and Data Science Workshop, MoMLeT and DS 2021, 2021, pp. 1–15.
- [10] Y. Kuryliak, System of modeling the spreading of infectious diseases, Proceedings of the International Conference on Applied Informatics ICDD, Sibiu, Romania, November 4-6 (2021) 106–114.
- [11] A. Lipowski, D. Lipowska, Roulette-wheel selection via stochastic acceptance, *Physica A: Statistical Mechanics and its Applications* 391 (2012) 2193–2196. URL: <https://ideas.repec.org/a/eee/phsmap/v391y2012i6p2193-2196.html>. doi:10.1016/j.physa.2011.12.0.
- [12] D. Chakrabarti, Y. Wang, C. Wang, J. Leskovec, C. Faloutsos, Epidemic thresholds in real networks, *ACM Transactions on Information and System Security (TISSEC)* 10 (2008) 1–26.
- [13] D. J. Watts, *Six degrees: The science of a connected age*, WW Norton & Company, 2004.