



Eidgenössische Technische Hochschule Zürich  
Swiss Federal Institute of Technology Zurich

*Distributed  
Computing*



# Infection Spreading in Graphs

Semester Thesis

Emanuel Jampen

`emjampen@student.ethz.ch`

Distributed Computing Group  
Computer Engineering and Networks Laboratory  
ETH Zürich

## **Supervisors:**

Pál András Papp, Béni Egressy  
Prof. Dr. Roger Wattenhofer

January 25, 2021

# Acknowledgements

I thank my supervisors Pál András Papp and Béni Egressy for their valuable feedback and interesting ideas, and Simea Jampen for helping me discuss some medical details of virus spreading.

# Abstract

In this thesis we implement virus spreading behaviour in social graphs based on the SIR Model. We show that reducing the number of social contacts as well as reducing infection risks in meetings can help to flatten the infection curve. Further we found that using limited vaccination resources, say less than 25% of the total population can be vaccinated, vulnerable individuals can be better protected by a vaccination campaign aimed at individuals with high numbers of social contacts or high infectiousness rather than vaccinating the most vulnerable directly.

To our surprise we found that forming small, highly connected communities within the population does not seem to have big impacts on the infection dynamics if the average degree is kept constant.

# Contents

<b>Acknowledgements</b>	<b>i</b>
<b>Abstract</b>	<b>ii</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Related Work . . . . .	1
<b>2 Model</b>	<b>3</b>
2.1 Graph . . . . .	3
2.1.1 Topology . . . . .	3
2.1.2 Edge Weights . . . . .	5
2.2 SIR Model . . . . .	5
2.2.1 Virus Initialization . . . . .	6
2.2.2 Infection . . . . .	6
2.2.3 Recovery . . . . .	7
2.2.4 Super Spreader . . . . .	8
2.2.5 Very Vulnerable individuals (VVI) . . . . .	8
2.2.6 Vaccination . . . . .	9
<b>3 Experimental Setup</b>	<b>11</b>
3.1 Simulation . . . . .	11
3.2 Figure Interpretation . . . . .	12
<b>4 Results</b>	<b>13</b>
4.1 SI-Method Comparison . . . . .	13
4.2 IR-Method Comparison . . . . .	16
4.3 Graph Type Comparison . . . . .	18
4.4 Weight Methods . . . . .	20
4.5 Flatten the Curve . . . . .	22

CONTENTS	iv
4.6 Super Spreaders in Community Graphs . . . . .	24
4.7 Vaccination in Community Graphs . . . . .	25
4.8 Vaccination in Erdős–Rényi Graphs . . . . .	27
<b>5 Conclusion</b>	<b>30</b>
5.1 Future Work . . . . .	30
<b>Bibliography</b>	<b>32</b>

# Introduction

---

During the Covid-19 pandemic probably every government in the world had to find some answers to questions like: How can we protect our population? Which measures are effective, which are not? How should we select individuals for our vaccination campaigns? Can the population reduce overall infectiousness by forming smaller communities within the population?

We try to answer questions like these based on simulations in our simplified model for a general viral disease. To this end we apply the SIR-Model of infection spreading to social network graphs. Different methods for infection spreading and recovering as well as different types of social graphs are explained in Chapter 2 and compared in Chapter 4. Further we will explore different vaccination strategies based on the objective of minimizing infections to the most vulnerable individuals. In the case of limited vaccination resources (less than 25% of the population can be vaccinated) we found that there are better strategies than to vaccinate the most vulnerable individuals directly.

We found that both reducing the number of social contacts as well as minimizing the risk of infection during a social meeting have huge effects on the rate at which the infection spreads. Those measures can even lead to extermination of the disease before every individual in the fixed-size population gets infected.

## 1.1 Related Work

The basis of our work was done already in 1927 by Kermack and McKendrick [1] when they first proposed the SIR-Model. A lot of work has been done based on this model, for example [2], [3] and [4]. Where [4] focuses on analytical solutions to the differential equations, [3] on the specific application regarding the measles vaccination and [2] on public health. All these differ from our work by the fact that we consider the social structures instead of the random mixing in the general model.

Application of the SIR-Model to graphs is for example done in [5] and [6]. Apart from other things the latter has a focus on the adjacency matrix of the

network and its eigenvalues, a topic that we did not consider in our work.

In [7] the concept of evolving Erdős–Rényi random graphs in combination with SIR was explored, in contrast to our work where graphs were fixed during simulation.

Not only viral infections of humans are considered as [8] shows epidemic spreading in cattle trade networks. An interesting alternative problem was treated in [9] where the spreading of information instead of viral infections in a graph was modeled by SIR. The goal was to locate the source of the information based only on knowledge of the graph topology and which individuals are infected at one arbitrary time-step. The extension to locate multiple sources was worked on in [10].

A lot of work was done regarding SIR-Models specific to the Covid-19 pandemic. For example [11], [12] and [13]. Where the first two, apart from other questions, look into the effect of viral detection tests not being 100% accurate. The third models population size that changes in surges. The work by [11] also claims to be able to match and predict actual data from the Covid-19 pandemic in China. Also the turning point, where infection rates drop below 1 could be accurately predicted. In our work we do not consider a specific viral infection and therefore also do not try to predict or match real-world data.

Our model consists of two main parts, the graph, representing individuals and their social contacts, and the SIR-Model. In this chapter we explain these two parts as well as the different optional settings we used for some of the simulations.

We represent individuals in a fixed population of  $N$  individuals as nodes in a graph. Edges between nodes represent social relations between the respective individuals such as friends, work colleagues and family members. As we will see, edges are weighted which represents some measure of closeness and infectiousness between individuals. By assuming that infectiousness is not necessarily symmetric between two individuals we end up with a symmetric directed graph i.e. every directed edge has a corresponding directed edge in the opposing direction with possibly a different weight.

An infectious viral disease is introduced via initial conditions to the graph and the spreading of the virus is modeled by the application of the SIR-Model [1] to the graph.

## 2.1 Graph

### 2.1.1 Topology

Three different random graph models were implemented to model social interactions, *Erdős–Rényi Graph*, *Small World Graph*, and *Community Graph*.

- **Erdős–Rényi Graph**

This is an Erdős–Rényi  $G(n, p)$  model [14]. It is implemented with probability  $p$  and number of nodes  $n$ . An example topology with  $N$  (we use  $N$  instead of  $n$ ) set to 20 and  $p = 0.65$  is displayed in Figure 2.1.

- **Small World Graph**

This topology is a Watts–Strogatz graph [15] based on the implementation from networkx [16]. To create a Small World Graph the  $N$  nodes are placed

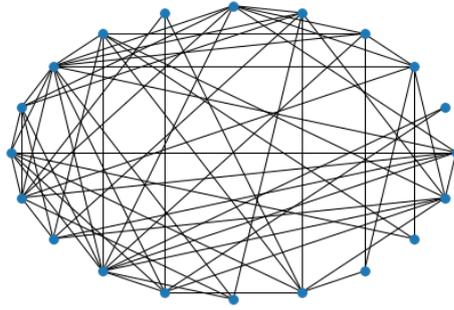


Figure 2.1: Erdős-Rényi Graph

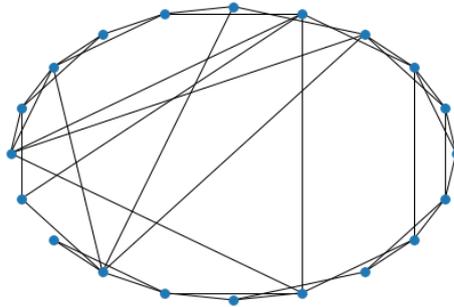


Figure 2.2: Small World Graph

in a circle. Each node is then connected to  $k$  of its nearest neighbours, where  $k$  is an even integer. Afterwards each edge is deleted and replaced with an edge to an uniformly random selected other node with probability  $p$ . In Figure 2.2 an example is displayed with  $N = 20$ ,  $k = 4$  and  $p = 0.15$ .

We did use Small World Graphs mostly in combination with the following Community Graph.

- **Community Graph**

The Community Graph is a two leveled version of the other topologies. Similar types of graphs are featured for instance in [17]. In the first level a Small World Graph is created with the number of nodes equal to the desired number of communities  $c$ . Each of this first level nodes is replaced by an Erdős-Rényi Graph with the number of nodes equal to  $N/c$  in order to create the second level. When  $N/c$  is not an integer, one of the communities has fewer nodes so that the total number of nodes is still  $N$ . When two level one nodes were directly by the connected Small World Graph we create multiple connections between the two corresponding communities. The number of those new connections is proportional to the number of nodes in the communities.

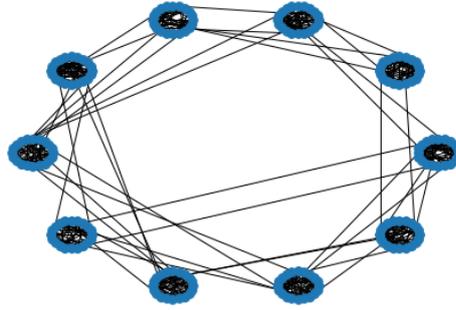


Figure 2.3: Community Graph, each community contains 20 nodes

### 2.1.2 Edge Weights

Since we want to use the edge weights to calculate infection probabilities between neighbouring nodes (see Section 2.2.2) two methods are implemented to generate the edge weights.

- **Poisson**

Each edge weight is independently selected from a Poisson distribution with parameter  $\lambda$ . This method obviously yields edge weights that can lie outside the interval  $[0, 1]$  which can be problematic for further probability calculations. To circumvent this issue normalization can be applied optionally in a way that yields incoming weights that sum to 1 per node.

- **Gauss-Sigmoid**

Each edge weight is independently selected from a Gauss distribution before being fed through a sigmoid function to get values in  $[0, 1]$ . For most simulations we used a normal distribution ( $\sigma = 1, \mu = 0$ ), but other settings are possible.

## 2.2 SIR Model

The SIR Model divides the total population of  $N$  individuals into three distinct groups, Susceptible, Infected and Recovered. Those groups are referenced by their first letters,  $S$ ,  $I$  and  $R$  hence also the model name. The size of these groups, i.e. the number of individuals belonging to them, is represented by a non-negative integer which we will reference as  $|S|$ ,  $|I|$  and  $|R|$  respectively.

A healthy individual starts in  $S$ . Upon infection it is moved to  $I$  and after recovering moved to  $R$ . It is assumed that once an individual recovers it is immune and can not be infected a second or third time. This is a reasonable assumption for many viral infections like chickenpox [18].

In contrast to the usual model, we look at SIR in discrete time-steps where one time-step roughly corresponds to one day. Additionally the graph based approach allows to distinguish individuals from each other, meaning we know in which of the three groups each individual is at any time-step. However, the basic continuous SIR-Model is essentially equivalent to the one we use with a fully connected graph.

The infection and recovery processes are described in detail in the following subsections.

### 2.2.1 Virus Initialization

Before the first time-step starts we randomly infect a number of individuals to kick-start the whole infection process. By default this number is set to 10. This loosely corresponds to some infected individuals entering a population, maybe from a foreign country. However, in our model we do not consider any other external factors like additional immigration or any additional external infections once the simulation has started.

### 2.2.2 Infection

The transition from being susceptible to being infected is calculated by the so called *SI-Method* (*S* to *I* Method) as depicted by the left arrow in Figure 2.4. We developed five different SI-Methods which are explained here. For any given simulation only one of these methods is used. In Section 4.1 we compare the different methods.

- **Threshold Stay-Ok**

For this and the second method we assume that a viral infection crosses an edge from an infected individual  $j$  to a susceptible individual  $i$  with probability  $w_{ji}$ . Here  $w_{ji}$  is the edge weight from node  $j$  to node  $i$ .

We can calculate the probability of individual  $i$  *not* getting infected, also known as staying ok, hence the method name, in the following way:

$$p_i = \prod_{\text{infected neighbour } j} b \cdot (1 - w_{ji}) \quad (2.1)$$

where  $b$  is a fixed base infection probability in  $(0, 1]$ . Individual  $i$  gets infected in this time-step when  $p_i$  is smaller than a fixed threshold.

- **Probabilistic Stay-Ok**

Here  $p_i$  is calculated as in 2.1. Individual  $i$  is then infected with probability  $1 - p_i$ .

- **Neighbour Count**

In this method we count the infected neighbours of a susceptible individual. If this number is larger than a set threshold the individual gets infected in this time-step, otherwise it stays susceptible. This implies that edge weights are ignored in this method.

- **Edge Sum**

This method is similar to Neighbour Count but instead of counting the infected neighbours we sum over the edge weights. A susceptible individual  $i$  is infected if the sum over its incoming edge weights  $w_{ji}$  is larger than a given threshold.

- **Sigmoid Edge Sum**

For a susceptible individual  $i$  we compute the same sum over its edge weights as for the Edge Sum method. This sum is then mapped to the interval  $[0, 1]$  by the function

$$p_i = 1 - \text{sigmoid}(s_i - c) \quad (2.2)$$

Where  $s_i$  is the edge weight sum of individual  $i$  and  $c$  is a correction constant which is by default set to 5. Individual  $i$  gets infected with probability  $p_i$ .

### 2.2.3 Recovery

The transition from  $I$  to  $R$  is calculated by the *IR-Method* as shown by the right arrow in Figure 2.4. Three different IR-Methods were implemented of which we mostly used the last one (Poisson). As with the infection methods, only one of these was used in any simulation.

- **Fixed Threshold**

In this model every infected individual is recovered after being infected for a fixed number of time-steps. This threshold is the same for all individuals.

- **Probabilistic**

Each infected individual will recover with probability  $p_r$  at any time-step.

- **Poisson**

As in the Fixed Threshold model individuals recover based on the number of days of being infected. However, here the threshold is independently selected from a Poisson distribution for each individual.

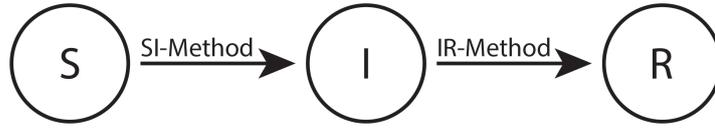


Figure 2.4: Group membership and transition functions

### 2.2.4 Super Spreader

We implemented so called *Super Spreaders*, individuals that have a significantly higher infectiousness. In a real world scenario these Super Spreaders can be individuals that refuse to follow health regulations like wearing face masks or individuals that for some reason are just by their biology more infectious.

We select super spreading individuals uniformly at random from the total population based on a constant ratio. For example with a ratio of 0.2 and a population of  $N = 100$  we would select 20 individuals as Super Spreaders. The desired effect of increasing the infectiousness is achieved by changing the outgoing edge weights of Super Spreaders depending on the selected edge weight method, both of which are described in Section 2.1.2.

- If the Poisson Method is used:  
Outgoing edge weights of Super Spreaders are *multiplied* by a constant factor.
- If the Gauss-Sigmoid Method is used:  
Outgoing edge weights of Super Spreader are calculated by *adding* a constant term  $f$  to the values from the Gaussian distribution before applying the sigmoid function as seen in equation 2.3:

$$w = \text{sigmoid}(G + f) \quad (2.3)$$

where  $G$  is a Gaussian random variable and  $w$  is the edge weight. This is equivalent to selecting edge weights for Super Spreaders from  $\mathcal{N}(\mu + f, \sigma^2)$  instead of  $\mathcal{N}(\mu, \sigma^2)$  before applying the sigmoid function.

### 2.2.5 Very Vulnerable individuals (VVI)

As a model for individuals with a high vulnerability to the infection we implemented the group of *very vulnerable individuals*. We select  $|VVI|$  individuals from the total population based on a constant ratio. Individuals are ranked by their degree (lowest to highest), the top  $|VVI|$  individuals are selected to be the very vulnerable individuals. This resembles our assumption that elderly people are often VVIs and also often have less social contacts than an average individual.

This also gets reinforced since most VVIs probably know that they are vulnerable and try to minimize social contacts to protect themselves.

VVIs behave just like any other individual, the only difference is that we value an infection of a VVI as being worse than an infection of a normal individual. This plays a central role in the evaluation of different vaccination strategies.

### 2.2.6 Vaccination

Vaccination was implemented as an optional setting. We modeled vaccination by moving some percentage of individuals from  $S$  to  $R$  before starting the first time-step. The selection of the individuals to vaccinate is done by four different strategies. The total number of individuals to vaccinate is denoted by  $|V|$ , which is always strictly smaller than  $N$ . This represents a situation where only a limited quantity of vaccinations is available. Independent of the chosen strategy we assume that the vaccine provides perfect and lifelong protection from the virus.

As we will further discuss in Sections 4.7 and 4.8 all vaccination strategies are compared in terms of how well they can reduce the total number of VVIs that get infected. Furthermore, strategies will be measured in their ability to protect the general public as well. The first and second strategy serve as baselines for those comparisons.

The goal of comparing these strategies is to see whether finding crucial points in the network is better at protecting VVIs than vaccinating them directly.

- **Uniformly Random**

Individuals are selected uniformly at random from the total population.

- **VVI - Baseline**

Individuals are selected uniformly at random from the group of VVIs. As mentioned above, this strategy only serves as a baseline to be able to compare the other strategies against. Obviously if the ratio of vaccinated individuals is higher or equal to the ratio of VVIs this strategy yields perfect results in the sense that no VVI will be infected.

- **Highest Edge-Sum**

The individuals are ranked by their *outgoing* edge weight sum (highest to lowest), the top  $|V|$  individuals are selected for vaccination. This has the goal to vaccinate individuals that are very infectious, such as Super Spreaders, and/or have many social contacts. Realizing this in a real world setting would require good knowledge of the infectiousness and social contacts of every individual.

- **Highest Degree**

In this vaccination strategy the individuals are ranked by their degree (highest to lowest), the top  $|V|$  individuals are selected for vaccination. As the Highest Edge-Sum strategy this one also requires good knowledge about social contacts of individuals. With this strategy we aim to cut as many infection paths as possible.

# Experimental Setup

---

In this chapter we explain how a simulation is carried out.

## 3.1 Simulation

Every simulation consists of multiple runs, by default we did a total of 100 runs, in groups of 10 runs that happen on the exact same topology. During one simulation all the values of the different settings discussed in the model Chapter 2 are fixed. We chose this approach in order to obtain a better statistical representation of the infection dynamics compared to a single run.

Each run is terminated when  $|I|$  becomes 0 or after a fixed number of time-steps.  $|I| = 0$  corresponds to the infection being exterminated in the population.

Unless stated otherwise, we use the following default settings:  $N = 300$  individuals, 10 communities (when a Community Graph is used), Gauss-Sigmoid weights, Poisson recovery and Probabilistic Stay-Ok as infection method. The number of initial infections is 10 and by default VVIs, super spreaders and vaccinations are not used in our simulations.

A repository with the python code that was used as well as the figures can be found in [19].

## 3.2 Figure Interpretation

Our main analysis tool of the simulations is the SIR-Plot. Figure 3.1 shows a generic example of such a plot. The horizontal axis show time-steps ranging from 0 to about 130 in this example while the vertical axis represents individual counts with values from 0 to 300, which is our default total population  $N$ . The exact numbers however, are rarely important as we will analyze the plots qualitatively.

Yellow curves show the number of susceptible  $|S|$ , red the number of infected  $|I|$  and green the number of recovered  $|R|$  individuals. Each curve is the result from one single run. The blue dotted curves represent the respective average values of  $|S|$ ,  $|I|$  and  $|R|$  over all runs.

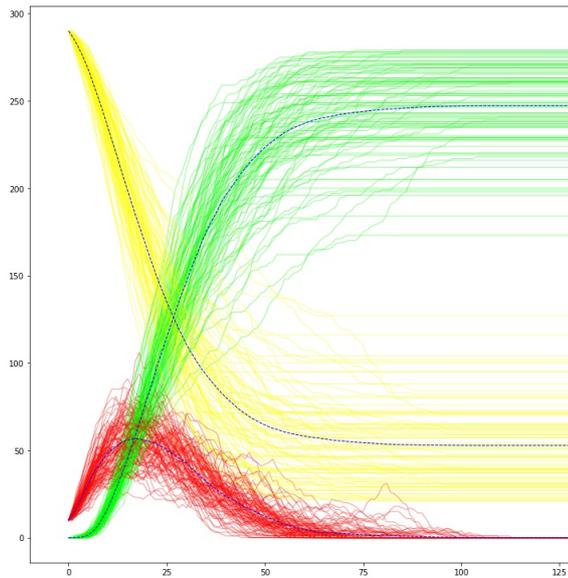


Figure 3.1: A generic SIR-Plot

## 4.1 SI-Method Comparison

As a first result we discuss how the different SI-Methods presented in Section 2.2.2 compare to each other under otherwise equal conditions. For this setting we use a community graph with 300 individuals in 10 communities.

Figures 4.1 through 4.5 show a progression through different values for the given methods. All plots are at the same scale, showing about 90 time-steps.

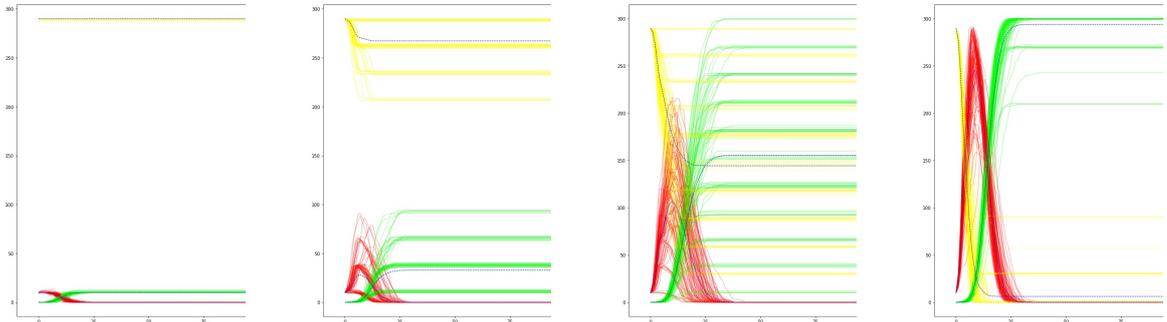


Figure 4.1: SI-Method Threshold Stay-OK with base infection probability 0.3, 0.4, 0.5 and 0.6 respectively

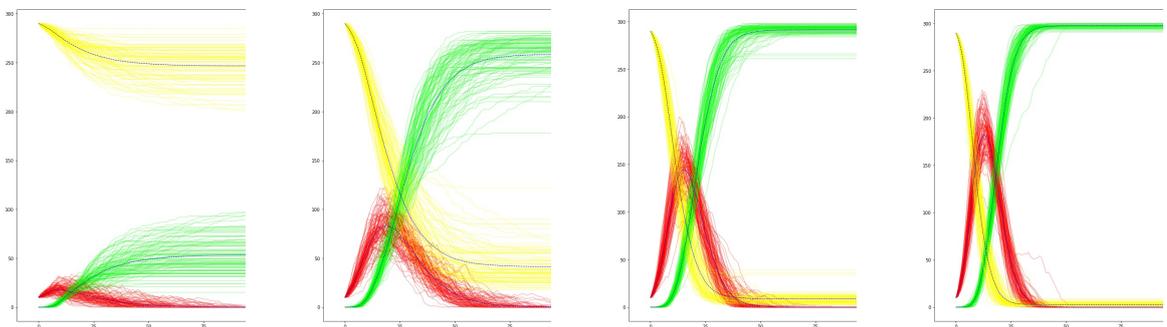


Figure 4.2: SI-Method Probabilistic Stay-Ok with base infection probability 0.02, 0.05, 0.08 and 0.11 respectively

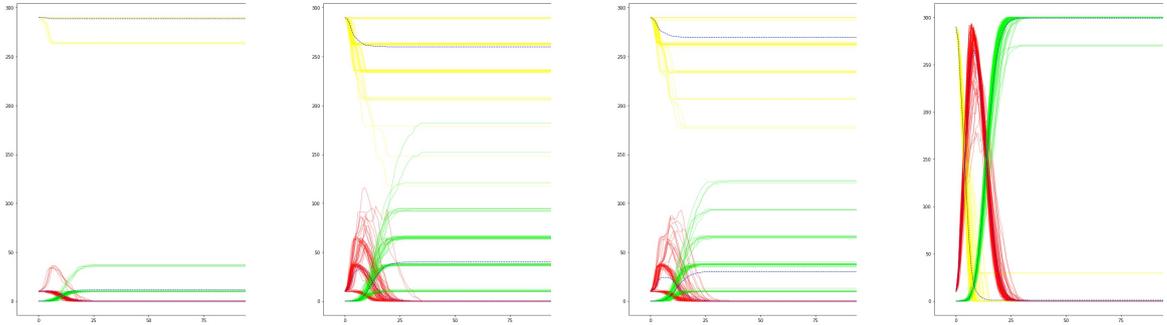


Figure 4.3: SI-Method Neighbour Count with threshold 3.5, 3.0, 2.5 and 2.0 respectively

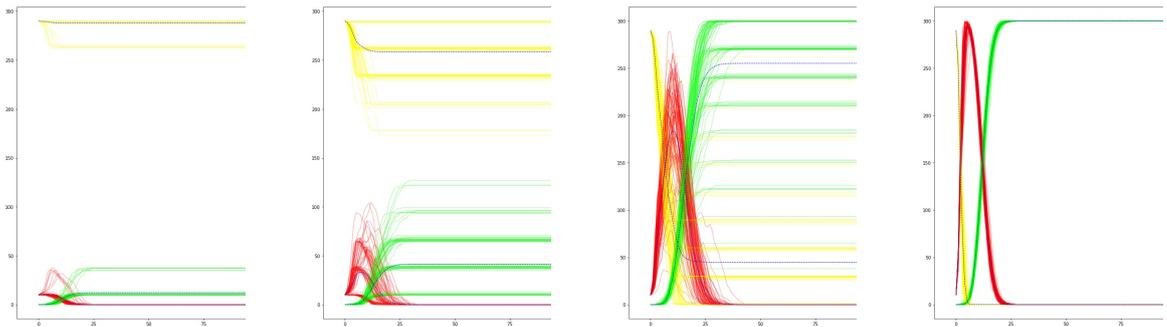


Figure 4.4: SI-Method Edge-Sum with threshold 1.8, 1.4, 1.0 and 0.6 respectively

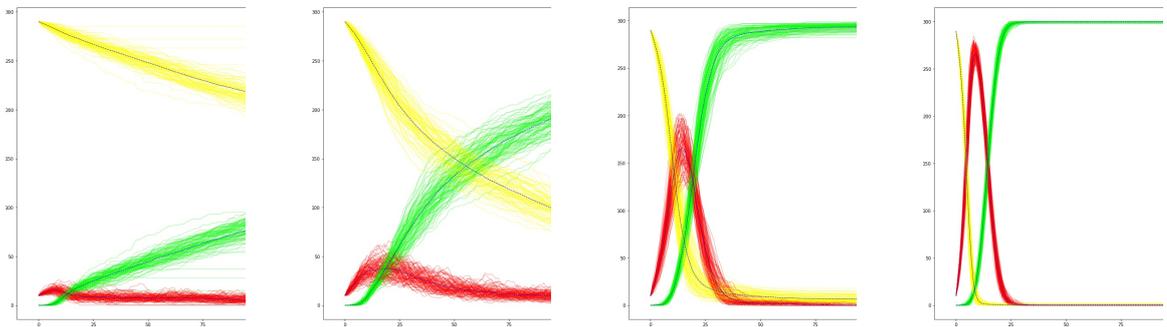


Figure 4.5: SI-Method Sigmoid Edge Sum with correction term 6.0, 5.0, 4.0 and 3.0 respectively

All methods show similar behaviours for extreme (very low or very high) values. This can be observed by comparing the right most and left most plots from each Figure, which do not show a lot of differences. However, in the respective medium ranges the dynamics are quite different. The three methods which use a threshold, Threshold Stay-OK, Neighbour Count and Edge Sum, show clear discrete levels in the  $S$  and  $R$  curves while the other methods create continuous distributions.

Interestingly the number of discrete levels in Figures 4.1, 4.3 and 4.4 seem to be similar and about 10. In order to verify whether this is related to the number of communities being 10, we did some further simulations. Figure 4.6 shows the effect that changing the number of communities has on the number of discrete levels for the Threshold Stay-Ok case. We observe a strong correspondence between those two numbers since there are 4 levels in the left graph and 19 in the right one. In Figure 4.7 we can see that in Erdős–Rényi Graphs there are only two discrete levels, one where everyone gets infected and one where almost no one gets infected during a run.

We therefore conclude that the discrete levels are caused by the thresholding SI-Methods which are discrete functions by design. However, the number of levels is determined by the number of communities and each level corresponds to a certain number of communities being infected more or less completely while other communities are left largely unharmed.

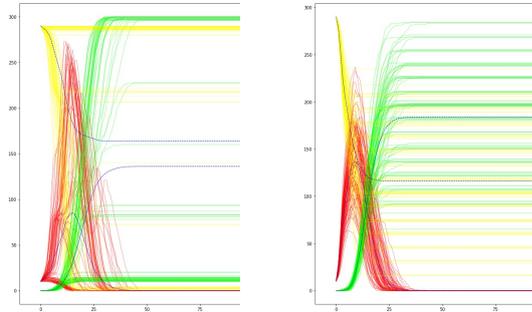


Figure 4.6: SI-Method Threshold Stay-Ok with 4 respectively 20 communities

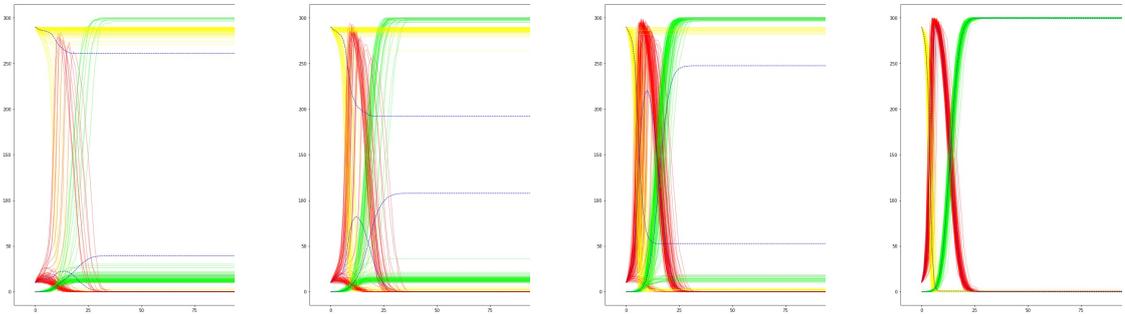


Figure 4.7: SI-Method Threshold Stay-Ok on Erdős–Rényi Graphs with base infection probabilities 0.41, 0.43, 0.48 and 0.55

## 4.2 IR-Method Comparison

In this subsection we compare the three IR-Methods described in Section 2.2.3. Figure 4.8 shows a progression through three different threshold values of the Fixed Threshold method. Very similar plots are created by the Poisson method in Figure 4.9 for three different values of  $\lambda$ . The main difference is only visible at the first 20 time-steps. Figure 4.11 displays a closeup on this region of the plots with a threshold of 12 and a  $\lambda$  value of 12 of the corresponding methods. The Fixed Threshold method has a discontinuity in the  $|R|$  curve at the 12<sup>th</sup> time-step because all of the initially infected individuals recover at this exact moment. This also leads to a sharp drop in the  $|I|$  curve. The Poisson method provides smooth curves, which is our main reason to use this method by default.

The SIR-plots corresponding to the Probabilistic method are displayed in Figure 4.10 for four different values of  $p$ . This method behaves similar to the Poisson method when set to the correct range of values.

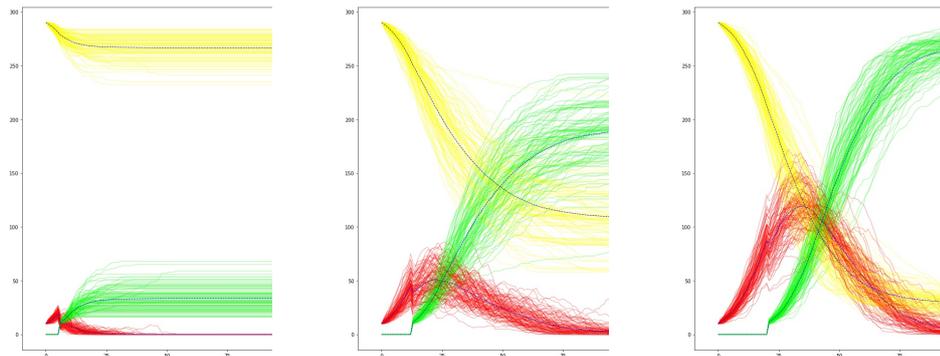


Figure 4.8: IR-Method Fixed Threshold with recovery time of 5, 12 and 20 time-steps

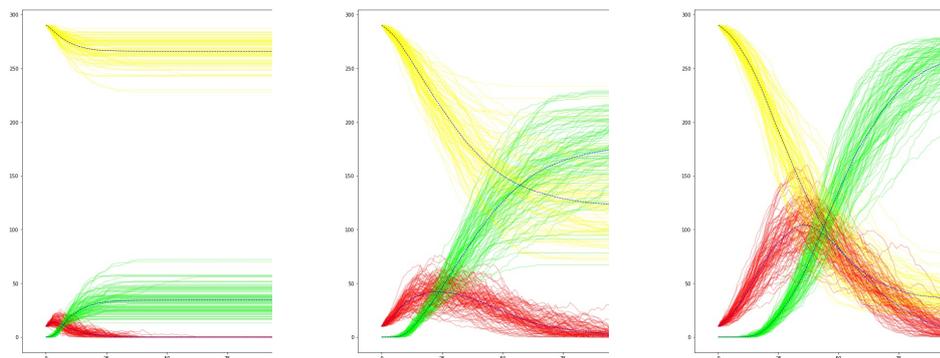


Figure 4.9: IR-Method Poisson with  $\lambda$  set to 5, 12 and 20 respectively

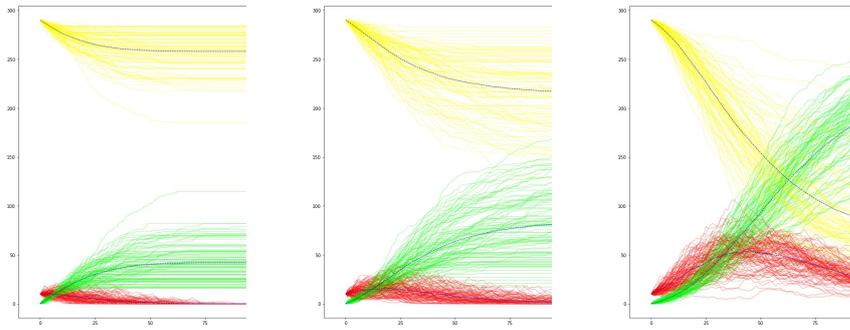


Figure 4.10: IR-Method Probabilistic with  $p$  set to 0.15, 0.1 and 0.05 respectively

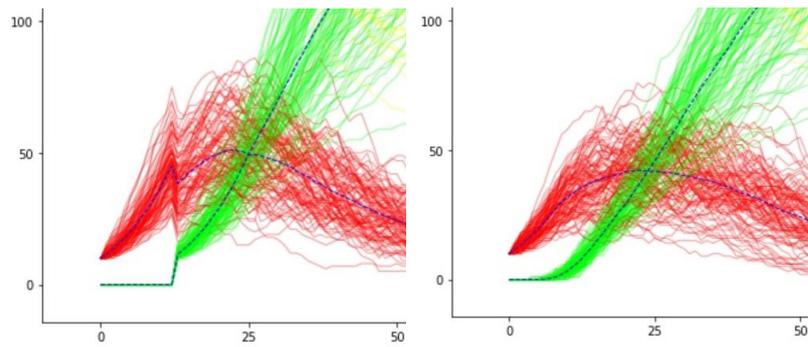


Figure 4.11: Closeup of the Fixed Threshold (left) and Poisson comparison plots

### 4.3 Graph Type Comparison

In this section we compare the effect of the graph type on the infection dynamics. To that end we create Erdős–Rényi, Community, and Small World Graphs with similar average degrees. Figure 4.12 shows the resulting plots. Plots in the same column correspond roughly to the same average degree. It is hard to spot large differences between the three graph types which is surprising since the topologies are quite different. The Erdős–Rényi and Small World plots however display a smaller variance in all three curves for higher degrees (right) than the corresponding Community Graph plots. This is explained by the fact that for higher degrees Erdős–Rényi as well as Small World becomes more and more a fully connected graph where statistical variance should decline due to homogeneity. Whereas the communities in the Community Graph are still only sparsely connected to each other, which can lead to more variance between simulation runs depending on how much communities are reached.

For the Small World Graph we only display plots with rewiring probability  $p = 0.2$  because changes in this factor did not yield large changes on the SIR curves.

From a practical standpoint this implies that forming small communities within the population does not slow down infection spreading as long as the average number of social contacts are kept constant. The reduction of this average degree however does have a strong impact on infection dynamics as we will see again in Section 4.5.

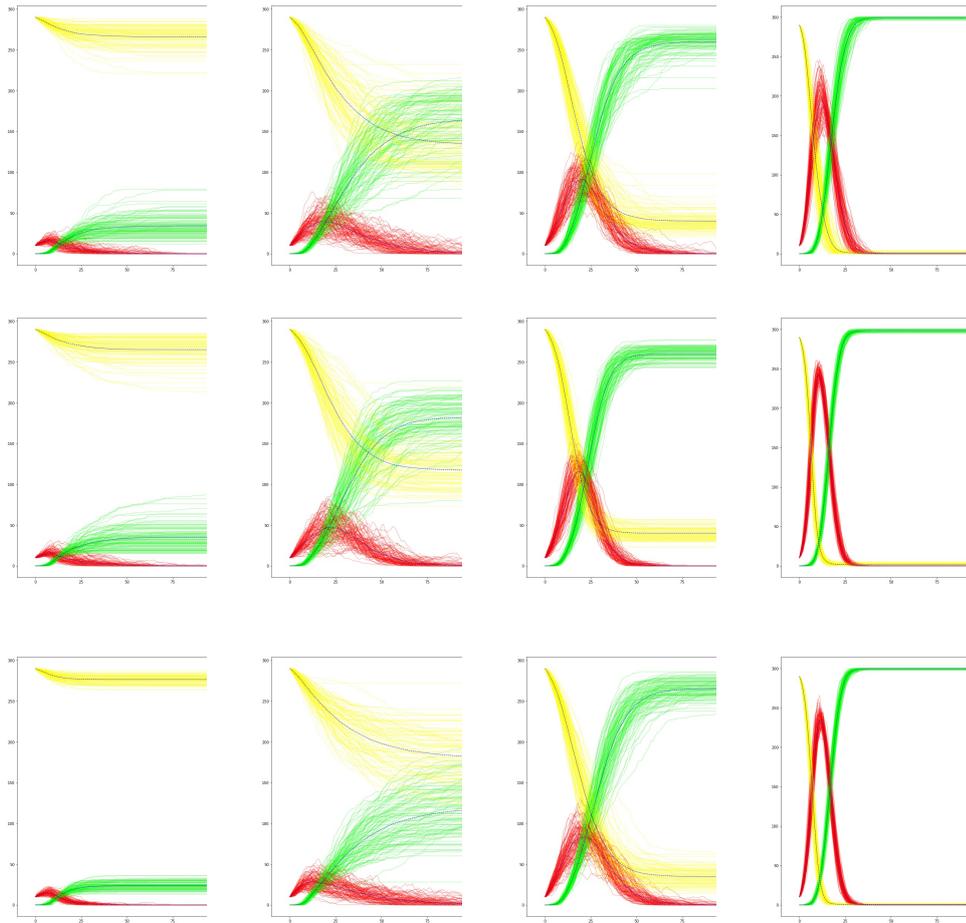


Figure 4.12: Community Graph plots in the first row, Erdős-Rényi Graph plots below and Small World Graph plots at the bottom, from left to right with average degree of about 2, 4, 6 and 14

## 4.4 Weight Methods

As our last basic setting comparison we explore the differences in the two weighting methods described in subsection 2.1.2. For the Gauss method we only focus on the impact of changing the standard deviation  $\sigma$ , since the effect of changing the mean is related to Super Spreaders. In Figure 4.13 we can see that changing  $\sigma$  does only have small effects on the SIR curves. A larger standard deviation leads to weights close to 1 and 0 while less weights are close to 0.5. However, the larger amount of weights close to 0 seem to be able to overpower the larger number of weights close to 1 and lead to an overall reduction of infectiousness, which can be observed by the slight rise of the yellow  $|S|$  curve from the left to the right image.

Figure 4.14 shows the effect of changing the value of the Poisson parameter  $\lambda$  from 1 in the left plot to 20 in the right one. The impact of this change is bigger than the one observed in the Gauss method above. It is interesting that there is such a big increase of infectiousness even though normalization is applied.

As both methods yield similar results we conclude that selecting just one of them (Gauss-Sigmoid) as default for the other simulations was a justified assumption.

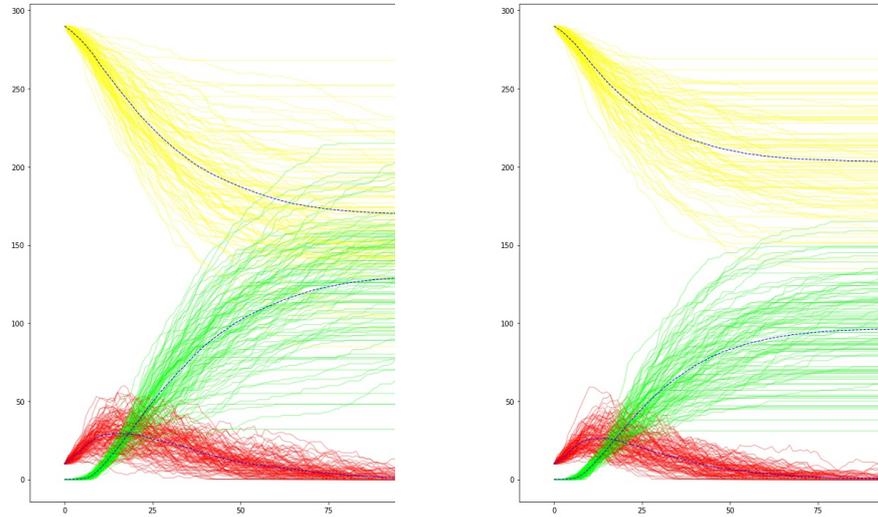


Figure 4.13: Gauss-Sigmoid weight method with standard deviation  $\sigma$  set to 1 (left) and 20

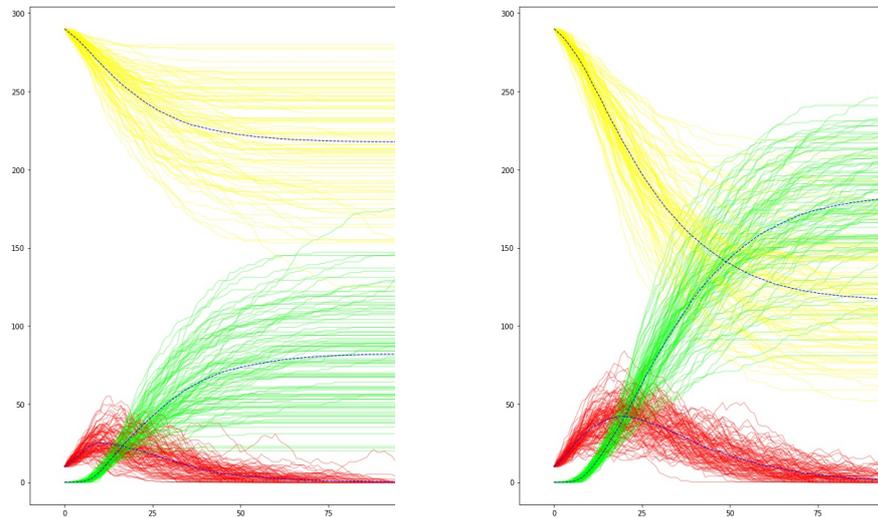


Figure 4.14: Poisson weight method with  $\lambda$  set to 1 (left) and 20

## 4.5 Flatten the Curve

In this section we explore whether we can replicate the concept of flattening the curve in our model and which settings have a great impact on this effect. The expression *flatten the curve* refers to the idea to reduce the maximum peak value in the  $I$  curve as much as possible. Any health care system has a limited capacity and applying flattening of the curve aims to not overwhelm this capacity. As the curve gets flatter the total number of individuals that got infected does not necessarily decrease but the number of infected individuals at any given time should get lower.

We explore two strategies with the goal to achieve a flattening of the curve. Reducing the base infection probability and reducing the average degree. These strategies correspond to making social meetings safer and reducing the number of different social contacts.

As Figure 4.15 shows, reducing the base infection probability ( $b$  in 2.1) can to some extent flatten the curve. However, the rising  $S$  curve in the two right most images show that lowering this value too much leads to some individuals not being infected at all instead of being infected at a slower time rate. Of course this is a good thing for all individuals but it is not the effect we were trying to reproduce. In a real world scenario this reduction of the base infection probability could be achieved by wearing face masks, frequent washing of hands or keeping physical distance during a meeting.

In the progression of plots in Figure 4.16 the average node degree is reduced from left to right. To some extent flattening of the curve can be observed, however the  $S$  curve starts to rise quite soon and stops the flattening process. Again, this is a good thing to happen for the population.

We have seen that we can flatten the curve to some extent but could not achieve the ideal flattening behaviour. Nevertheless we observed that both reducing the infection risk during a meeting and reducing the number of social contacts have great potential to reduce infection spreading through the population.

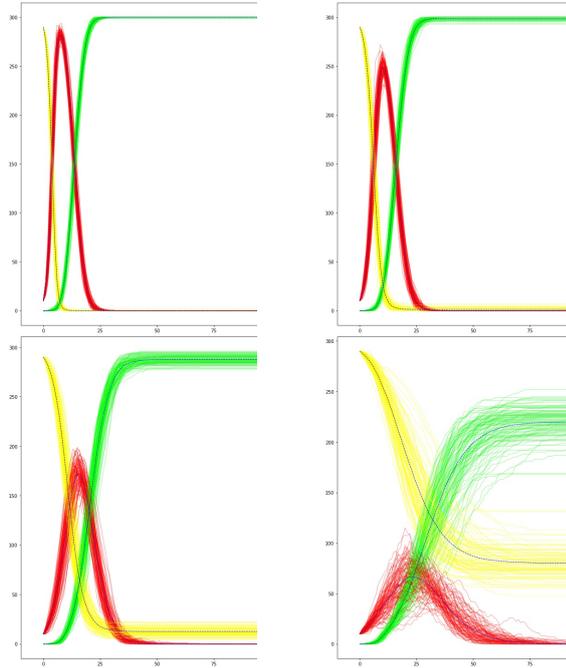


Figure 4.15: Flatten the Curve by reducing base infection probability from left to right: 0.18, 0.1, 0.05 and 0.025 respectively.

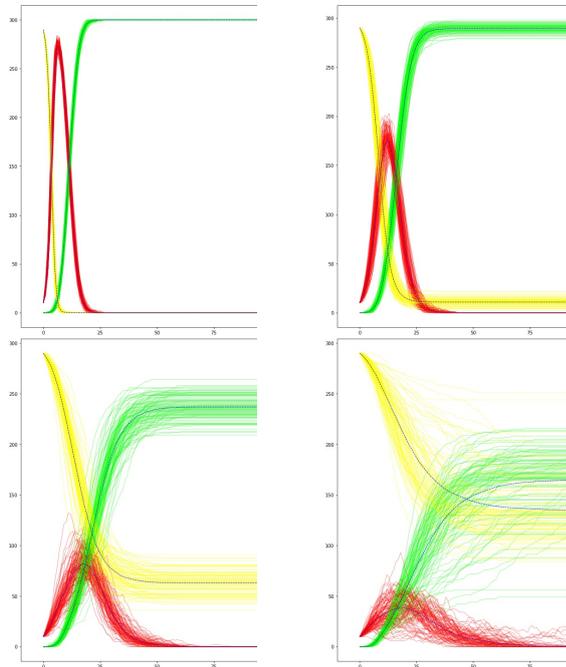


Figure 4.16: Flatten the Curve by reducing the average degree, from left to right: 30, 10, 6 and 4.

## 4.6 Super Spreaders in Community Graphs

The effect of Super Spreaders largely depends on two factors: The ratio of Super Spreaders (SSR) in the total population and the Super Spreader factor (SSF).

Figure 4.17 shows the effect of these two factors. From left to right the SSF is increased and takes values 0.0, 2.0 and 9.5. From top to bottom the SSR increases and takes values 0.2, 0.4 and 0.6.

We observe that the  $S$  and  $R$  curves approach a constant value for high SSF. Changes from 2.0 to 9.5 in the SSF result in diminishing changes of the respective curves. Our explanation for this observation is: At some SSR value every Super Spreader reaches an infectiousness such that he almost instantly infects every of its susceptible neighbours. An even higher SSR does not lead to an increase in this infectiousness.

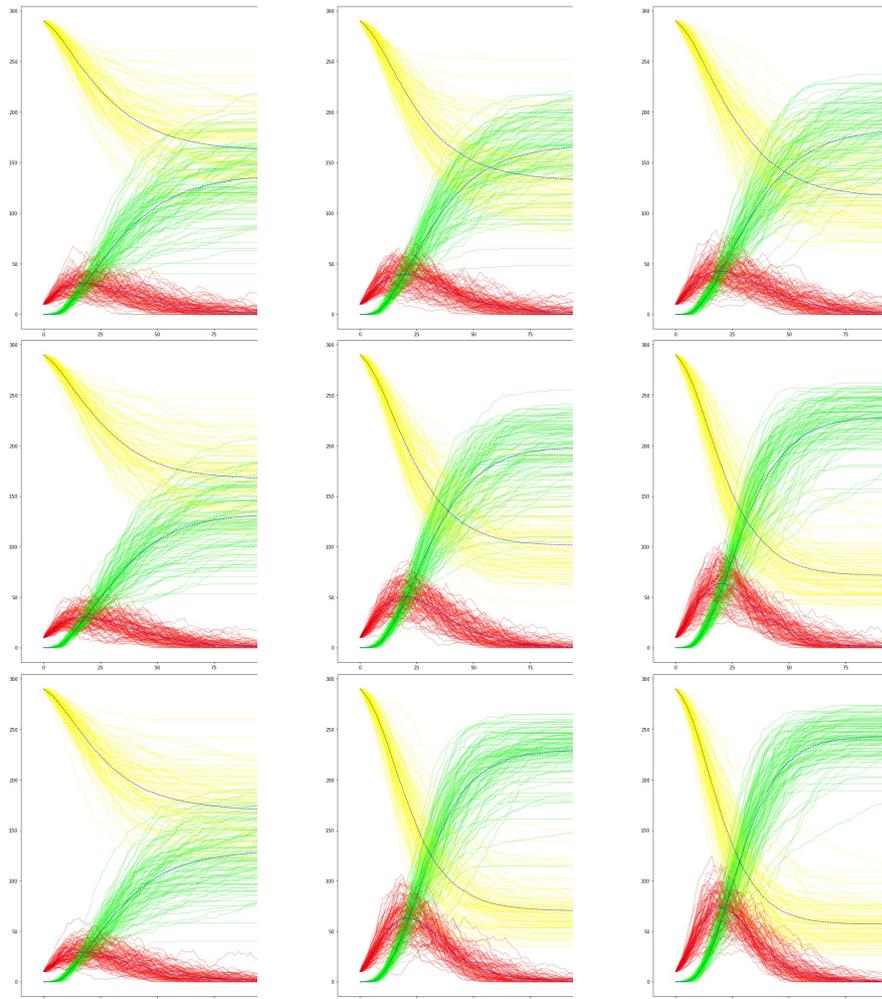


Figure 4.17: Effects of Super Spreaders in a Community Graph

## 4.7 Vaccination in Community Graphs

As described in Subsection 2.2.5 some individuals, which we refer to as VVIs, are more prone to suffer from an infection than others, maybe even with fatal consequences. In this section we compare our vaccination strategies based on the goal to protect as many VVIs as possible with limited vaccination resources.

Of course every strategy that is called good by any means should perform better than using the same number of vaccinations directly on the VVIs. This will therefore be our main baseline comparison. Random vaccination is also regarded largely as a baseline.

We did simulations on a Community Graph, 300 individuals in 10 communities with 40% VVIs and 10% Super Spreaders. Figure 4.18 shows the average infection ratio on the last time-step of the simulation of VVIs for each strategy over a spectrum of vaccination ratios. As we approach a vaccination ratio of 0.4 the baseline curve drops to 0 since almost every VVI is getting vaccinated. For vaccination ratios below 0.25 both the Edge Sum as well as the Degree strategies outperform the baseline. Random vaccination performs poorly over the whole spectrum.

Figure 4.19 displays the ratio of all individuals that never got infected, meaning they are either vaccinated or still susceptible at the end of the simulation. The Edge Sum curve overlaps quite well with the Degree curve so that it has low visibility on the plot.

While the baseline strategy outperforms all other strategies in protecting VVIs on vaccination ratios above 0.25 it performs poorly in protecting other individuals.

For both metrics, Edge Sum and Degree vaccination strategies are both highly effective.

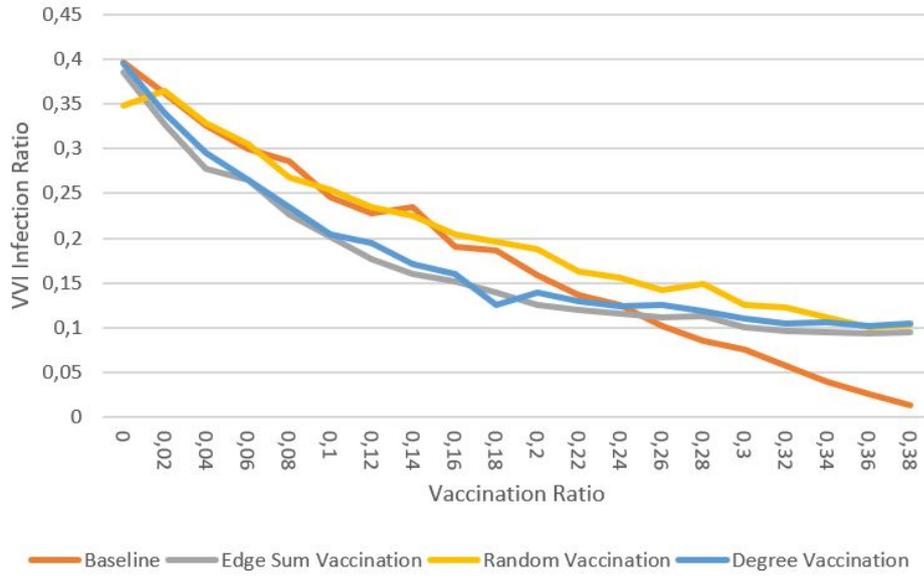


Figure 4.18: VVI infection ratio in Community Graphs

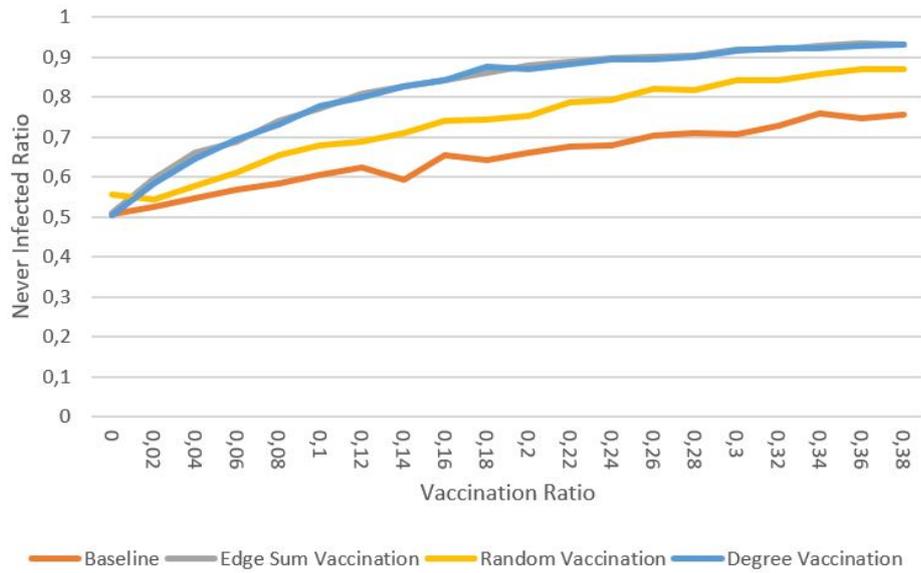


Figure 4.19: Ratio of never infected individuals in Community Graphs

## 4.8 Vaccination in Erdős–Rényi Graphs

The same simulation as in Section 4.7 was also run on a Erdős–Rényi Graph with 300 individuals and an average degree of about 12 which is almost twice as high the average degree in Section 4.7. All other settings are the same.

Figure 4.20 displays the ability of a strategy to protect VVIs. Surprisingly the baseline strategy which simply aims to directly vaccinate VVIs outperforms all other strategies. We hypothesize that the high degree is responsible for the poor performance of the strategies that proved to work well in this metric on the Community Graph. With a high degree there are simply too many paths over which an infection can spread. A look at Figure 4.21 shows that Edge Sum vaccination and Degree vaccination have a better performance at protecting the whole population than the baseline strategy.

In order to check our assumption we run the same simulation again with an average degree similar to the one in subsection 4.7, about 6.4. Figure 4.22 and Figure 4.23 display the corresponding results. As the good performance of the Edge Sum and Degree strategies are back again our hypothesis seems to hold.

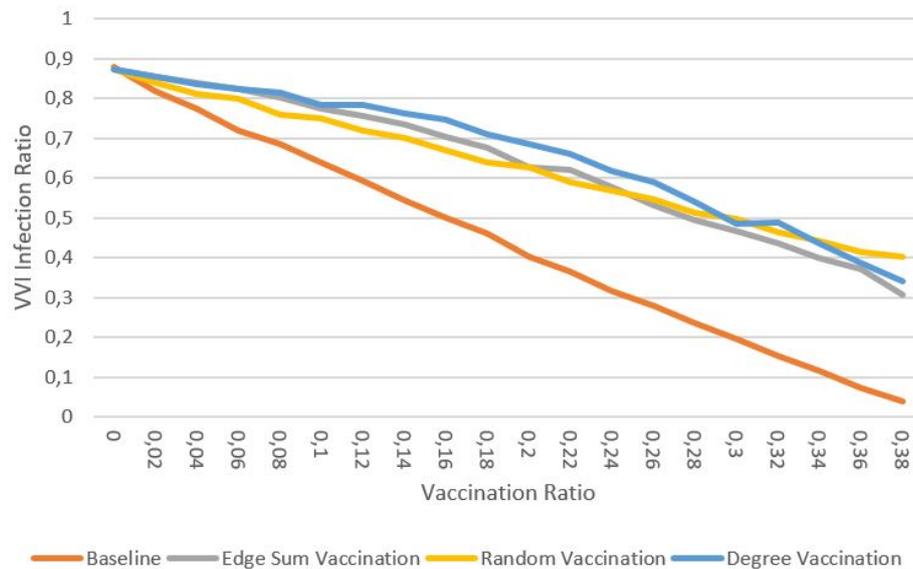


Figure 4.20: VVI infection ratio in Erdős–Rényi Graphs with high average degrees

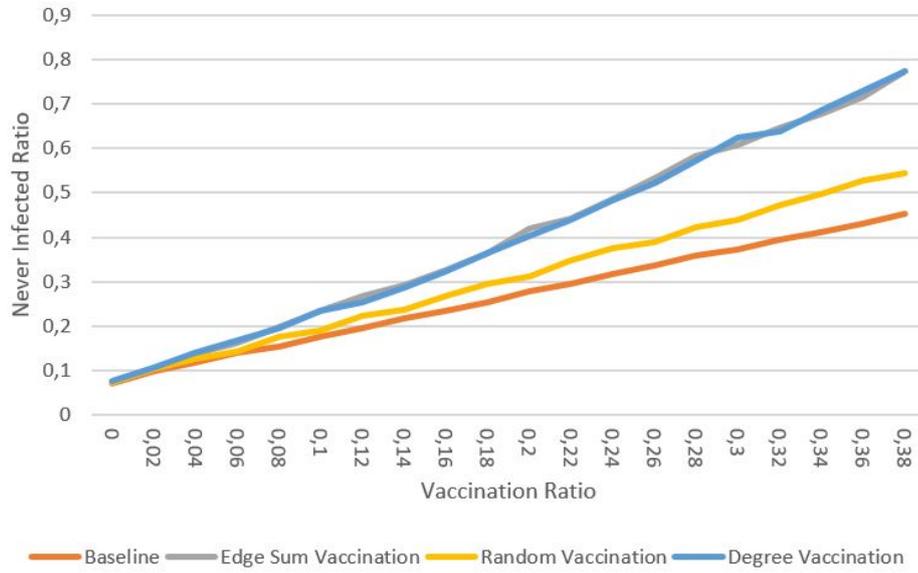


Figure 4.21: Ratio of never infected individuals in Erdős-Rényi Graphs with high average degrees

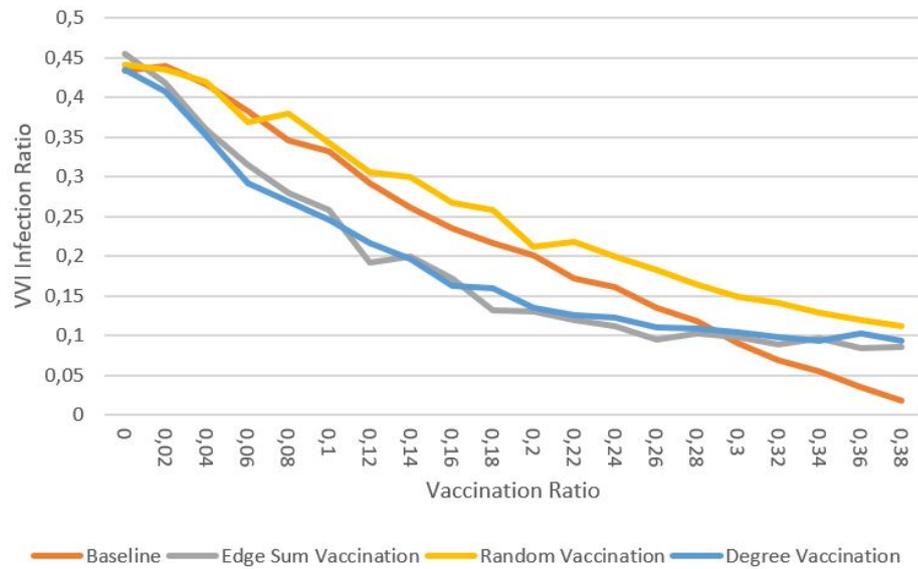


Figure 4.22: VVI infection ratio in Erdős-Rényi Graphs with low average degrees

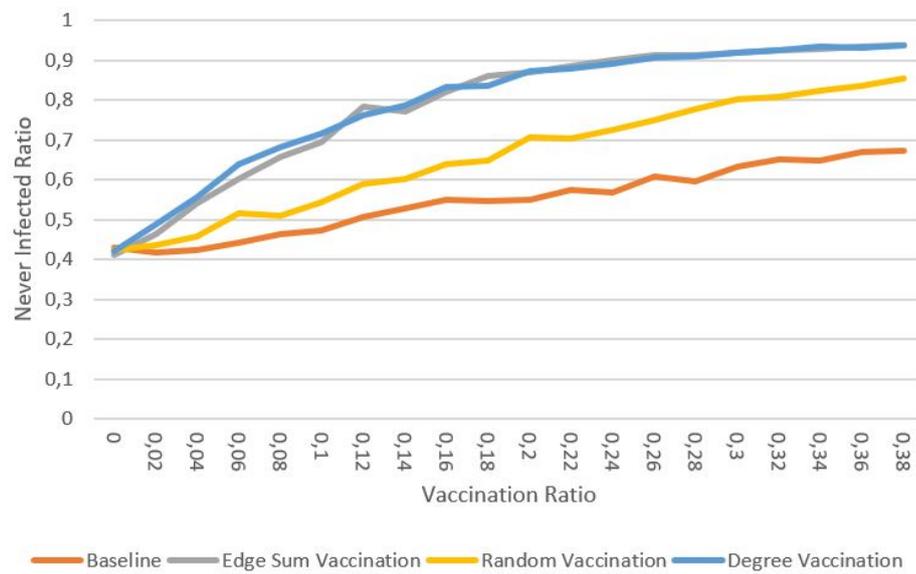


Figure 4.23: Ratio of never infected individuals in Erdős-Rényi Graphs with low average degrees

# Conclusion

---

We could replicate basic viral spreading behaviour in a realistic way with our model. However, as we have seen in 4.5, replicating flattening of the curve or even multiple infection waves like they occur in real viral pandemics such as Covid-19 is difficult and would probably require multiple different factors of influence to change simultaneously.

Arguably our most interesting findings are the results from 4.7 and 4.8 where we compare different vaccination strategies. As we have seen it can be more effective to *not* vaccinate VVIs directly but to vaccinate individuals that are likely to spread the virus through the population by having many social contacts or being more infectious. As long as vaccination resources are limited, say less than 25% of the total population can be vaccinated, these strategies not only perform better in protecting VVIs but also provide good protection for the general population. The Indonesian government decided to apply such strategies in their Covid-19 vaccination campaign [20]. On the other hand the Swiss government pursues the strategy to vaccinate VVIs first [21][only available in Swiss national languages].

As we have seen in 4.8, the average degree, which corresponds to the number of social contacts of an individual, has a big impact on the effectiveness of all considered strategies. Governments around the world seem to understand this concept in their struggle to fight the Covid-19 pandemic and still suggest to reduce social contacts as much as possible. For example the Swiss government still prohibits gatherings of more than five individuals [22] even though the vaccination campaign has already started. Of course there might also be other reasons for this.

## 5.1 Future Work

We can imagine multiple concepts being implemented as optional extensions to our model. Both the addition of new individuals to the population through immigration or birth as well as temporal changes in the graph topology could be

interesting changes to our work. The latter could for example represent short-lived events with large participant numbers like sport events, concerts or demonstrations. We propose that changes of this kind would allow to observe second and third infection waves.

The SIR-Model could be expanded to an SIRD-Model which considers the deceased individuals in a separate group. Also considering the concept that not every individual is immune after recovery, as done in [23], could allow for more complex infection dynamics.

A more realistic approach to vaccination could be implemented by reducing the effectiveness of its protection against the virus. For example [24] shows the effectiveness of a specific vaccine against Covid-19 being between 90% and 100%.

Tests for detecting an infection in an individual could be implemented, individuals that are tested positively could change their behaviour, for example reducing their social contacts, in order to prevent the virus from spreading.

# Bibliography

- [1] W. O. Kermack and A. G. McKendrick, “A contribution to the mathematical theory of epidemics,” Aug. 1927.
- [2] H. H. G. I. o. T. M. D. Weiss, “The sir model and the foundations of public health,” *Materials matemàtics*, pp. 1–17, 2013. [Online]. Available: <https://ddd.uab.cat/record/108432>
- [3] B. Shulgin, L. Stone, and Z. Agur, “Pulse vaccination strategy in the sir epidemic model,” *Bulletin of Mathematical Biology*, vol. 60, no. 6, pp. 1123–1148, Nov 1998. [Online]. Available: [https://doi.org/10.1006/S0092-8240\(98\)90005-2](https://doi.org/10.1006/S0092-8240(98)90005-2)
- [4] T. Harko, F. S. Lobo, and M. Mak, “Exact analytical solutions of the susceptible-infected-recovered (sir) epidemic model and of the sir model with equal death and birth rates,” *Applied Mathematics and Computation*, vol. 236, pp. 184 – 194, 2014. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S009630031400383X>
- [5] M. J. Keeling and K. T. Eames, “Networks and epidemic models,” *J. R. Soc. Interface.*, 2005.
- [6] D. Chakrabarti, Y. Wang, C. Wang, J. Leskovec, and C. Faloutsos, “Epidemic thresholds in real networks,” *ACM Trans. Inf. Syst. Secur.*, vol. 10, no. 4, Jan. 2008. [Online]. Available: <https://doi.org/10.1145/1284680.1284681>
- [7] Y. Jiang, R. Kassem, G. York, M. Junge, and R. Durrett, “Sir epidemics on evolving graphs,” 2019.
- [8] P. Montagnon, “A stochastic sir model on a graph with epidemiological and population dynamics occurring over the same time scale,” 2019.
- [9] K. Zhu and L. Ying, “Information source detection in the sir model: A sample-path-based approach,” *IEEE/ACM Transactions on Networking*, vol. 24, no. 1, pp. 408–421, 2016.
- [10] W. Zang, P. Zhang, C. Zhou, and L. Guo, “Locating multiple sources in social networks under the sir model: A divide-and-conquer approach,” *Journal of Computational Science*, vol. 10, pp. 278 – 287, 2015. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S1877750315000824>

- [11] Y.-C. Chen, P.-E. Lu, C.-S. Chang, and T.-H. Liu, “A time-dependent sir model for covid-19 with undetectable infected persons,” vol. 7, no. 4. Institute of Electrical and Electronics Engineers (IEEE), Oct 2020, p. 3279–3294. [Online]. Available: <http://dx.doi.org/10.1109/TNSE.2020.3024723>
- [12] G. C. Calafiore, C. Novara, and C. Possieri, “A modified sir model for the covid-19 contagion in italy,” 2020.
- [13] I. Cooper, A. Mondal, and C. G. Antonopoulos, “A sir model assumption for the spread of covid-19 in different communities,” vol. 139, 2020, p. 110057. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0960077920304549>
- [14] R. Durrett, “Erdős–rényi random graphs,” in *Random Graph Dynamics*. Cambridge: Cambridge University Press, 2006, pp. 27–69.
- [15] D. J. Watts and S. H. Strogatz, “Collective dynamics of small-world networks,” in *Nature*, 393, pp. 440–442, Jun. 1998.
- [16] “Watts–strogatz small-world graph.” NetworkX Developers, last accessed 28. December 2020. [Online]. Available: [https://networkx.org/documentation/stable/reference/generated/networkx.generators.random\\_graphs.watts\\_strogatz\\_graph.html?highlight=watts\\_strogatz\\_graph#networkx.generators.random\\_graphs.watts\\_strogatz\\_graph](https://networkx.org/documentation/stable/reference/generated/networkx.generators.random_graphs.watts_strogatz_graph.html?highlight=watts_strogatz_graph#networkx.generators.random_graphs.watts_strogatz_graph)
- [17] N. M. Girvan M, “Community structure in social and biological networks,” 2002, p. 7821–7826, last accessed 20. January 2021. [Online]. Available: <https://doi.org/10.1073/pnas.122653799>
- [18] G. N. H. S. W. S. U. M. Musharrafieh, I. A. Nuwayhid and A. R. N. Bizri, “Immunity to chickenpox among school adolescents in lebanon and options for vaccination,” Jan. 2003.
- [19] E. Jampen, “Gitlab repository for infection spreading in graphs.” [Online]. Available: [https://gitlab.ethz.ch/disco-students/hs20/ejampen\\_virus/](https://gitlab.ethz.ch/disco-students/hs20/ejampen_virus/)
- [20] “Indonesia coronavirus: The vaccination drive targeting younger people.” BBC, last accessed 20. January 2021. [Online]. Available: <https://www.bbc.com/news/world-asia-55620356>
- [21] “Covid-19-impfstrategie: Besonders gefährdete personen sollen zuerst geimpft werden.” Federal Office of Public Health FOPH, last accessed 20. January 2021. [Online]. Available: <https://www.bag.admin.ch/bag/de/home/das-bag/aktuell/medienmitteilungen.msg-id-81667.html>
- [22] “Coronavirus: Federal council extends and tightens measures.” Federal Office of Public Health FOPH, last accessed 20. January 2021.

[Online]. Available: <https://www.bag.admin.ch/bag/en/home/das-bag/aktuell/medienmitteilungen.msg-id-81967.html>

- [23] A. Osemwinyen and A. Diakhaby, “Mathematical modelling of the transmission dynamics of ebola virus,” vol. 4, 07 2015, pp. 313–320.
- [24] F. P. Polack, S. J. Thomas, N. Kitchin, J. Absalon, A. Gurtman, S. Lockhart, J. L. Perez, G. Pérez Marc, E. D. Moreira, C. Zerbini, R. Bailey, K. A. Swanson, S. Roychoudhury, K. Koury, P. Li, W. V. Kalina, D. Cooper, R. W. Frenck, L. L. Hammitt, Türeci, H. Nell, A. Schaefer, S. Ünal, D. B. Tresnan, S. Mather, P. R. Dormitzer, U. Şahin, K. U. Jansen, and W. C. Gruber, “Safety and efficacy of the bnt162b2 mrna covid-19 vaccine,” vol. 383, no. 27, 2020, pp. 2603–2615, pMID: 33301246. [Online]. Available: <https://doi.org/10.1056/NEJMoa2034577>