

# Quality Criteria for Control of Epidemics.

Vikenty Mikheev  
vikentym@ksu.edu

April 2020

## Abstract

The total number of infected and the maximum number of infected are considered as criteria for control by delayed isolation of the SIR-type epidemic on a temporal Barabasi-Albert graph. Simulations are run to estimate optimal delays.

**Key words:** delayed isolation; SIR; control of epidemics; temporal Barabasi-Albert graph; total number of infected; maximum number of infected; temporal network

## 1 Introduction

How can we control the epidemics of type Susceptible  $\rightarrow$  Infected  $\rightarrow$  Recovered (SIR)?

In March 2020, many nations around the world reacted to the spread of COVID-19 through policies of social isolation to help stop the pandemic and minimize the damage. Isolation typically involves voluntary and non-voluntary restrictions on social interactions. Simply speaking, in-person contacts are limited to the closest circles. These measures have obvious economic repercussions. So, isolation has its own price. This economic criterion will certainly be studied extensively in the future. In this work, however, we analyze the effect of the isolation on the spread of the disease.

There have already been some examples of this kind of study. One good animation of the Brownian motion approach can be found in [Stevens, 2020] and more theoretical work is in [van der Hofstad et al., May 12, 2010]. The classic differential equations approach referred so much recently is explained well in [Smith and Moore, 2004].

We apply the network (graph) approach. A problem of SIS (Susceptible  $\rightarrow$  Infected  $\rightarrow$  Susceptible) epidemic control on graphs was considered in [Pereira and Young, 2015]. SIS model isn't applicable to COVID-19 and influenza. We consider SIR model on a temporal graph. This graph is built according to Barabasi-Albert preferential attachment mechanism [Albert and Barabási, 2002], which is known to be one of the best to reflect social connections. Then we run this SIR simulation with 2-,4-,6- week restriction on communication with

“popular” nodes (ones with the degree equal or greater than 5), starting with different delays from the beginning of the epidemic process. The result is presented as plots of the infected and recovered. The computer code can be found at [Mikheev, 2020].

## 2 Model and Parameters

### 2.1 Graph

Barabasi–Albert graph of size  $n$  is built by a consecutive augmenting process. After initiation of the graph  $G$  with  $m$  nodes and no edges, one connects a new node  $x$  to  $G$  with  $m$  edges in such a way that the probability that  $x$  is connected to  $a \in G$  is equal to the degree of  $a$ . So, “popular” nodes grow faster (see fig. 1). The number  $m$  is a parameter, in this work  $m = 3$ . The building process stops when the size of  $G$  reaches  $n$ . We transform  $G$  to a temporal graph by assigning the times of availability to the edges. We assume time to be discrete, measured in days by natural numbers.

If two nodes  $x, y \in G$  are connected by an edge with the assigned time moment  $t$ , we say  $x$  and  $y$  are *temporal neighbors* at  $t$ . The number of temporal neighbors of a node  $x \in G$  at  $t$  is called *temporal degree* of  $x$  at  $t$ .

### 2.2 Epidemic process

When the Barabasi–Albert graph of  $n$  nodes is built we can simulate an epidemic with the SIR process. All recovered assumed to be immune to secondary infection during the considered epidemic. Each node in  $G$  has a status: “susceptible”, “infected” or “recovered”. Nodes are contagious if and only if they have status “infected”. It may be asked why don’t we have a status “dead”. Despite how cruel it sounds, for the spread of the disease, dead and recovered play the same role because they can’t get infected no more. After  $G$  is initiated with  $i$  infected nodes at the day zero (see fig. 3, 5), every day each contagious node infects with  $p$  probability each of its temporal neighbors with the status “susceptible”, meaning changing the temporal neighbor’s status to “infected”. The duration of status “infected” is 30 days, after which the node becomes “recovered” (see fig. 4, 6). What should the value of parameter  $p$  be? The triple product of an average temporal degree, duration of “infected” status and  $p$  should be equal to the *reproductive number* of the infection, i.e. how many new people on average get infected from one infected. The reproductive number of COVID-19 is an open question. Different statistical methods give quite different results [Liu et al., 2020]. Here we assume that COVID-19 has reproductive number 3.6. The average temporal degree of 5000-node Barabasi-Graph is about 5.99. So, we set  $p = 0.02 \approx 3.6/(5.99 \cdot 30)$ . The Python code used for all simulations in this work is left in an open access [Mikheev, 2020]. Everyone is welcome to play with  $p$  and other parameters.

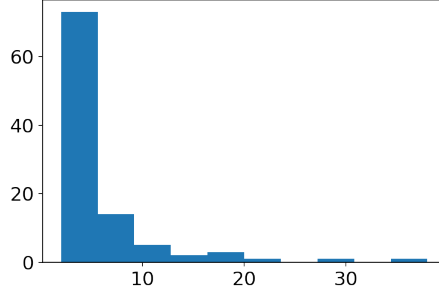


Figure 1: Histogram of the distribution of node degrees in Barabasi graph with  $n = 100$ ,  $m = 3$ .

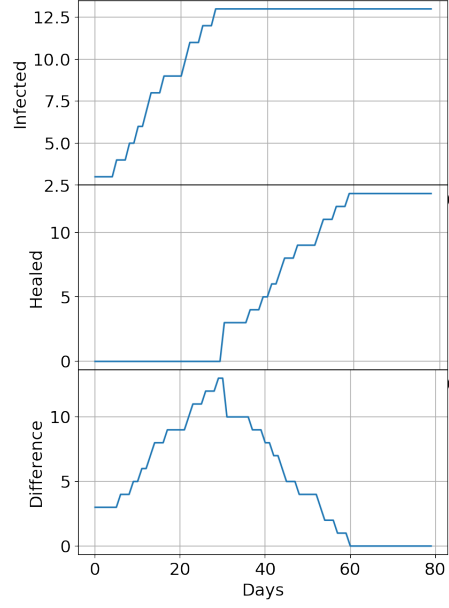


Figure 2: Number of infected, recovered nodes and their difference against time.

Since the random number generator is involved, the infection process may go differently even on the same graph and same initially infected nodes (see fig. 7, 8). We should take into account in the following experiments.

To observe the epidemic we plot the number of infected, recovered and difference of them against time (see fig. 2).

### 2.3 Isolation

One inconvenience of classic epidemic models based on differential equations or Brownian motion is that it is hard to conduct experiments with isolation. Temporal graphs allow us to do it relatively easy. First, for simplicity, we give one availability time moment per day to each edge of  $G$  during  $[t_1, t_T]$ . ( In the future works, someone can try to model more complex distributions of schedules on the graph, for example, make some edges available on weekdays and others only on the weekends.) Then we enforce “No gathering 5 or more people” rule from day  $t_s$  to day  $t_f$  as following: every node  $x \in G$  in the time interval  $[t_s, t_f] \subset [t_1, t_T]$  disconnects from its neighbor which has temporal degree 5 or more. Since we do it in an iterative manner for all nodes in  $G$  on each time step, at some point “popular” nodes stop being “popular” and keep some connections.

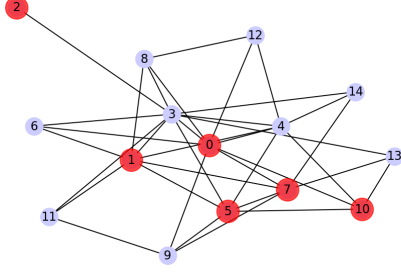


Figure 3: Barabasi 15-node graph with 6 infected nodes.

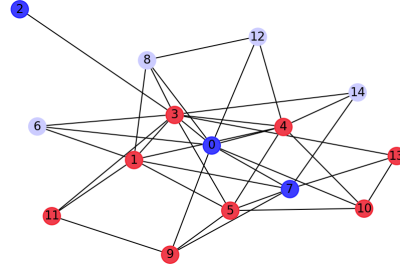


Figure 4: Infected 15-node graph after 32 days. Blue nodes are recovered or dead.

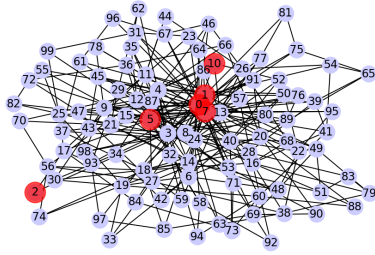


Figure 5: Barabasi 100-node graph with 6 initially infected.

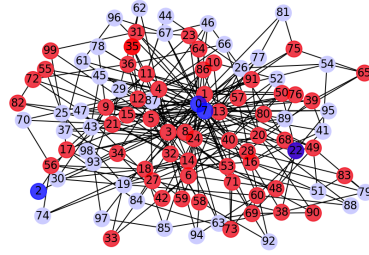


Figure 6: Barabasi 100-node graph after 32 days.

The time of the beginning of isolation  $t_s$  we call *isolation delay*.

Someone could say that isolating “popular” nodes isn’t the same that “no gathering”. Well, we think it is equivalent because gathering in a group of  $k$  people at day  $t$  means that the temporal degree of each person in this group greater or equal  $k$ . You may see a university lecture as a gathering of  $k$  students or you may say that the professor has a temporal degree  $k$  or higher. If students don’t come in their lecture because of the “no gathering”, they lower their temporal degree and the temporal degree of the professor.

### 3 Quality criteria

Epidemics bring complex damage to society. Individuals going through the illness work less and, in severe cases, may die. So, it makes sense to consider measures against the infection and study the effectiveness of them.

The first quality criterion of the actions against the epidemic processes we choose to be  $F_{tot}$ , the total number of people getting sick during the epidemic.

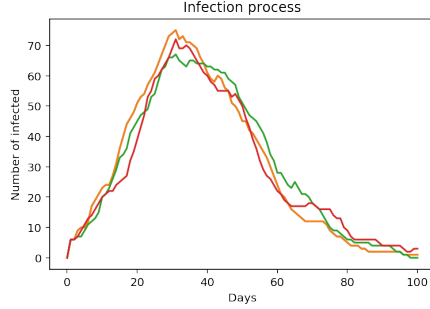


Figure 7: Variability of the epidemic process on 100-node graph showed by three different random seeds.

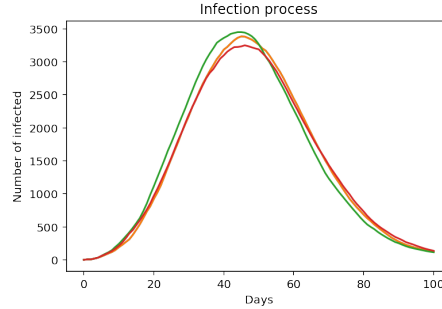


Figure 8: Variability of the epidemic process on 5000-node graph showed by three different random seeds.

Now let's think about fatal cases. Their number depends on the total number of infected but also depends on the available spots in hospitals. If too many ill people in severe conditions simultaneously come for medical service, there may not be enough resources to help them. Thus, another approach would be to minimize the maximum number of infected during the epidemic period (so-called "flattening"). So, the second quality criterion is  $F_{max} = \max_t F^t$ , where  $F^t$  is the number of currently infected at a day  $t$  of the epidemic.

We consider isolation as a single tool to reduce both  $F_{tot}$  and  $F_{max}$ . The isolation has two control parameters: its duration  $l$  and the starting time of it  $t_s$ . It is reasonable to assume that the cost of isolation does not depend on  $t_s$ . This allows to perform optimization of  $F_{tot}$  and  $F_{max}$  over  $t_s$  ignoring the cost of the isolation for a fixed  $l$ .

However, the cost of isolation does depend on  $l$ . This dependency is not studied in this work, where we make experiments with a few fixed values of  $l$ . So, we remove  $l$  from arguments of objective functions to their upper indices and limit ourselves in considering two following mathematical

**Problem 1:** Minimize the objective function  $F_{tot}^l$  over  $t_s$ .

**Problem 2:** Minimize the objective function  $F_{max}^l$  over  $t_s$ .

The experiments conducted in the next section are meant to address the two problems above but, of course, may give some other food for thought.

## 4 Experiments and Observations

We conducted experiments on 5000-node temporal Barabasi-Albert graph with no isolation and the isolation of lengths  $l = 14, 28, 42$  days starting at  $t =$

14, 21, 28 day from the beginning of the epidemic. Each simulation is run with the same random seed. For the convenience of comparison the graphs of simulations for the same length of the delay are combined in one plot. Thus, we obtained three plots 11, 13 and 14 for  $l = 14, 28, 42$ , respectively.

The total numbers of infected over the epidemic time people are in table 1 and table 2 contains the maximum numbers of infected during the epidemic periods. Observe that min-max values of infected obtained with delay equal 28, 28 and 21, for 2-,4-,6-week isolation, respectively. The minimum total numbers of infected are obtained on a 35-day delay of all 2-,4-,6-week isolation. Notice that the results may vary from one random seed to another (see fig. 12).

Giving a better estimate for the optimal delay meets some difficulties. First, we would have to average it over random seeds. Second, we don't observe the unimodality of the objective function  $F_{tot}$ . For example, its values for 2-week isolation from 24- to 35-day delay can be seen on fig. 9. The function  $F_{max}$  seem to behave nicer 10. The unimodality of  $F_{max}$  is a direction for a future theoretical research.

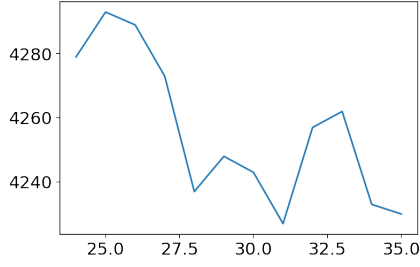


Figure 9: Non-convex behavior of  $F_{tot}^{14}$ .

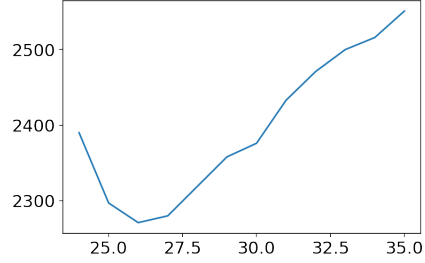


Figure 10: Nicer behavior of  $F_{max}^{14}$ .

For the same random seed, which was used to compute figures 11, 13, 14 and tables 1, 2, we compute the optimal delay for  $F_{max}$ . It is 26, 25, 23 for, respectively, for 2-,4-,6-week isolation.

Table 1: Total numbers of ever infected.

|         | No isol. | 14-th | 21-th | 28-th | 35-th | 42-th |
|---------|----------|-------|-------|-------|-------|-------|
| 2 weeks | 4450     | 4405  | 4342  | 4236  | 4230  | 4270  |
| 4 weeks | 4450     | 4328  | 4181  | 4023  | 4006  | 4148  |
| 6 weeks | 4450     | 4288  | 4070  | 3900  | 3888  | 4105  |

Table 2: Maximum numbers of infected.

|         | No isol. | 14-th | 21-th | 28-th | 35-th | 42-th |
|---------|----------|-------|-------|-------|-------|-------|
| 2 weeks | 3208     | 3013  | 2700  | 2319  | 2551  | 2985  |
| 4 weeks | 3208     | 2629  | 1925  | 1897  | 2547  | 2985  |
| 6 weeks | 3208     | 2447  | 1497  | 1897  | 2547  | 2985  |

On the figure 15 one may see that results are similar on 10,000 nodes. For 6-week isolation, the total numbers of infected are 8881, 8446, 8004, 7812, 7997, 8342 and the maximum numbers are 6462, 4332, 2866, 4407, 5545, 6252 for, respectively, no isolation, 14-, 21-, 28-, 35-, 42-day delays.

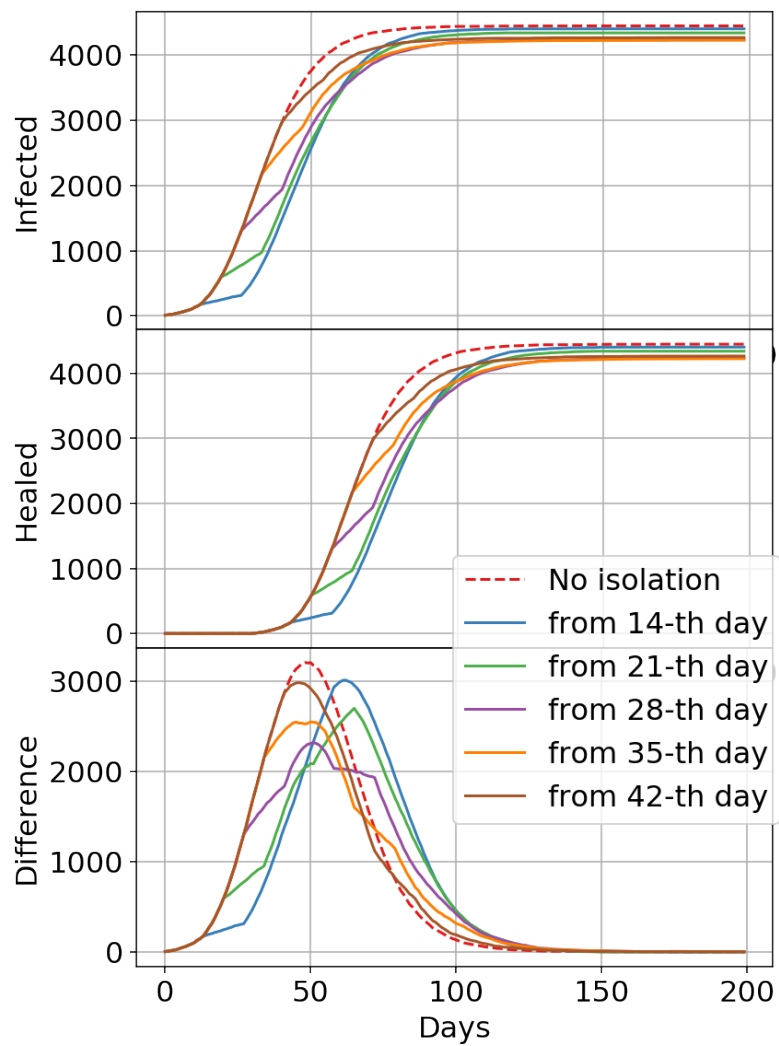


Figure 11: 2-week isolation.



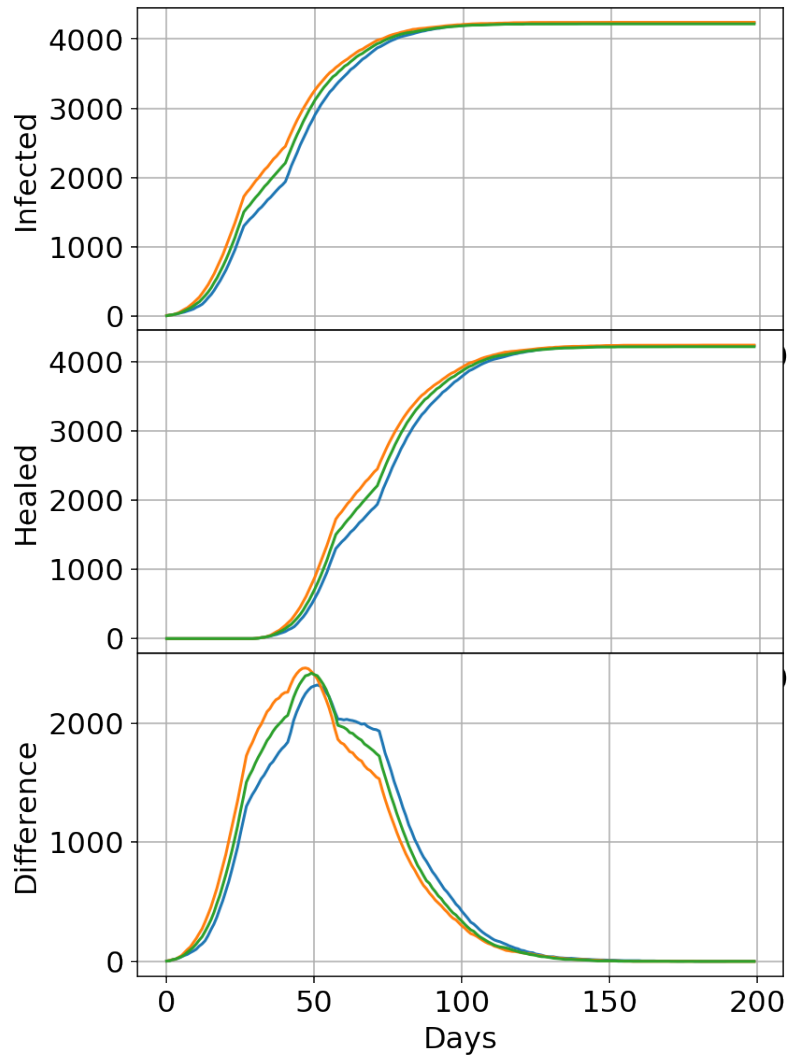


Figure 12: The variability of the infection process on 2-week isolation with 28-day delay showed by running the simulation with three different random seeds.

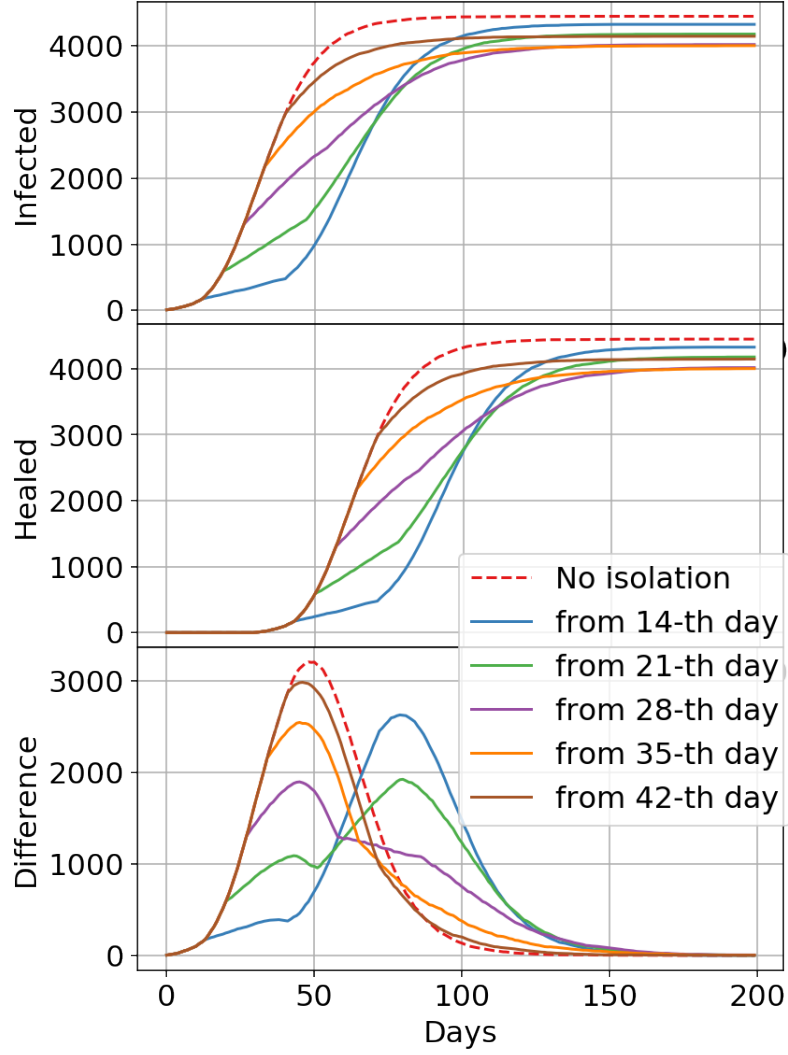


Figure 13: 4-week isolation.

## 5 Conclusion and Recommendations

The experiments we conducted, show that the objectives to decrease total number of infected and to decrease maximum number of infected (problems 1 and 2 stated in 3) are not antagonistic to each other. Firstly, the values of both criteria improve when the delay increases starting from zero. Secondly, they get worse when the delay becomes significantly large. Thirdly, a good compromise can be found because the optimal values are achieved at closely spaced delays

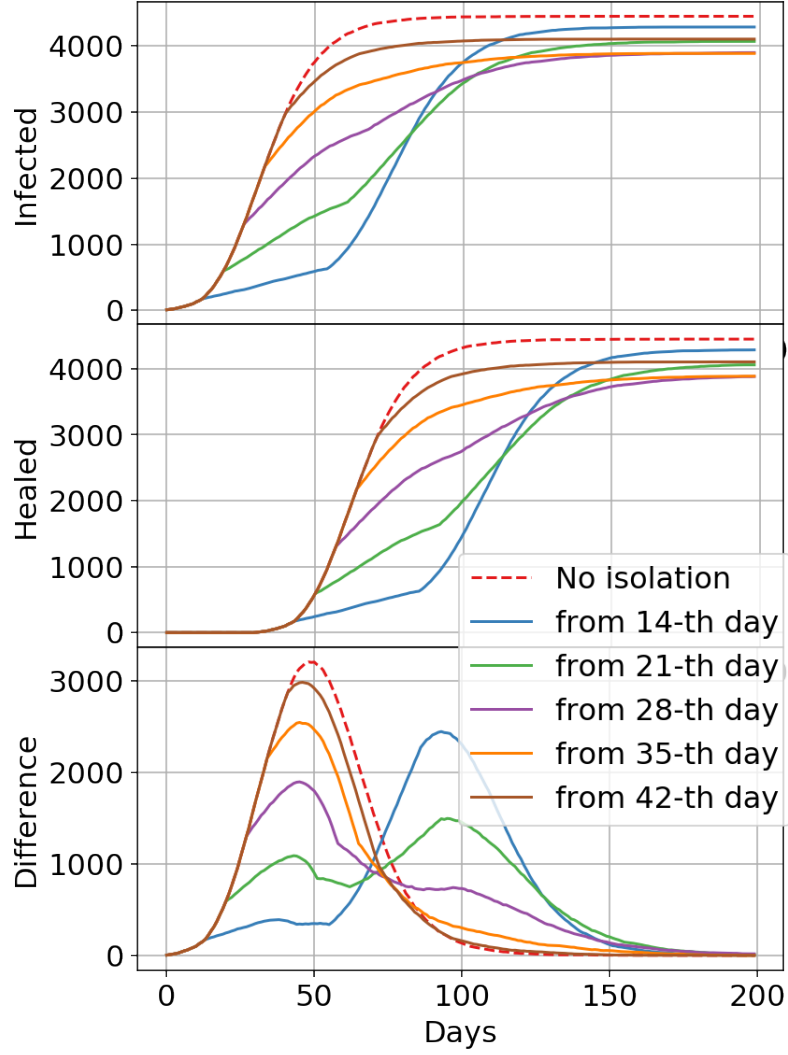


Figure 14: 6-week isolation.

delays. For the isolation with length from 2 to 6 weeks, that compromise is the delay between 21- and 35-day delay.

Alas, to compute the recommendation when to start an isolation at a certain day of epidemic, besides the optimal delay  $t_{opt}$ , we need to know the day of the beginning of epidemic, which is usually unknown. Fortunately, we know  $n$ , the size of population. Also,  $I$ , the amount of infected at the current day  $t_{today}$  can be estimated. Thus, we can run "no isolation" simulation on  $n$ -node graph to the day  $t_1$ , when the number of infected in the simulation is equal to  $I$ . Then

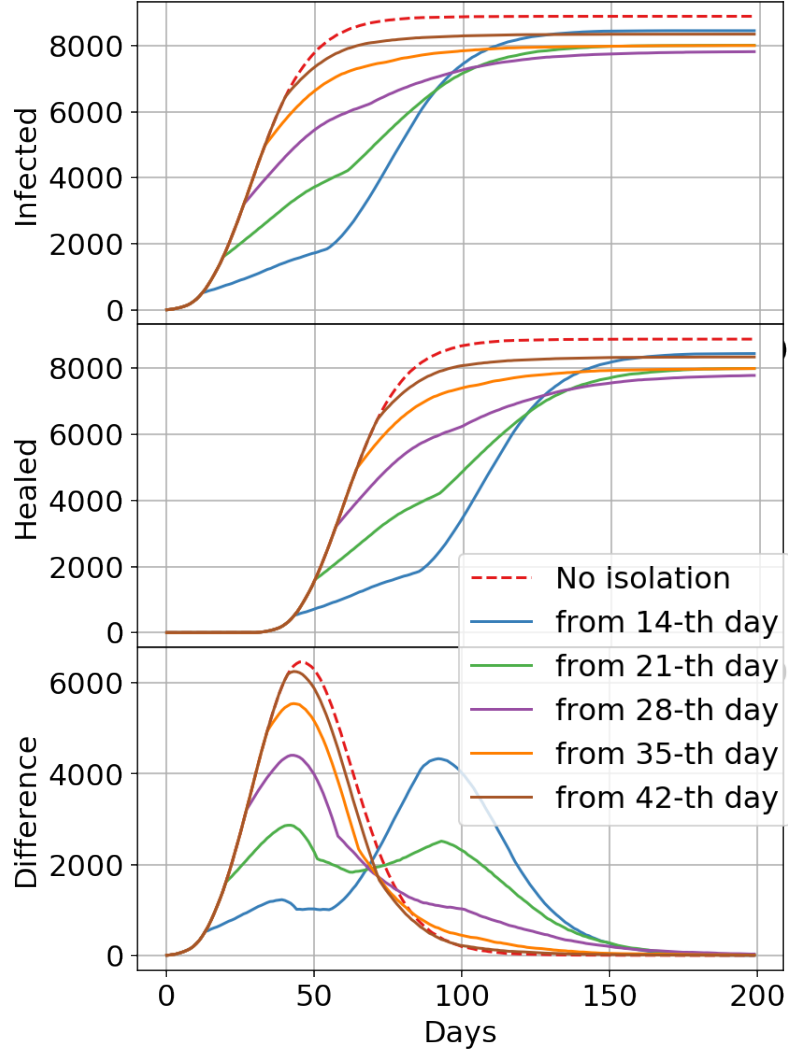


Figure 15: 10,000 nodes. 6-week isolation.

$t_{today} - t_1 + t_{opt}$  is the time to start the isolation.

For example, a town with population 5,000 people registers first infected. The town knows that they can afford 4 week isolation. They could randomly select a group of citizens and run the virus test on them. After running the "no isolation" simulation suggested in this paper, the town computes the day when the epidemic began. Table 2 tells them that to minimize the maximum number of infected they have to start the isolation around 28-th day of the epidemic, which is approximately when 27% of population is infected. If there is a daily

updates on estimate  $I/n$ , the town can use them instead of time scale, meaning simply start the isolation when  $I/n = 0.27$ .

## References

- R. Albert and A.-L. Barabási. Statistical mechanics of complex networks. *Rev. Mod. Phys.*, 74:47–97, Jan 2002. doi: 10.1103/RevModPhys.74.47. URL <https://link.aps.org/doi/10.1103/RevModPhys.74.47>.
- Y. Liu, A. A. Gayle, A. Wilder-Smith, and J. Rocklöv. The reproductive number of covid-19 is higher compared to sars coronavirus. *Journal of Travel Medicine*, 27(2), 02 2020. ISSN 1195-1982. doi: 10.1093/jtm/taaa021. URL <https://doi.org/10.1093/jtm/taaa021>.
- V. Mikheev. Isolationsironbarabasi. <https://github.com/keshmish/IsolationSIRonBarabasi.git>, 2020.
- T. Pereira and L.-S. Young. Control of epidemics on complex networks: Effectiveness of delayed isolation. *Phys. Rev. E*, 92:022822, Aug 2015. doi: 10.1103/PhysRevE.92.022822. URL <https://link.aps.org/doi/10.1103/PhysRevE.92.022822>.
- D. Smith and L. Moore. The sir model for spread of disease - the differential equation model. *MAA publications*, December 2004. URL <https://www.maa.org/press/periodicals/loci/joma/the-sir-model-for-spread-of-disease-the-differential-equation-model>.
- H. Stevens. Why outbreaks like coronavirus spread exponentially, and how to “flatten the curve”. *The Washington Post*, March 2020. URL <https://www.washingtonpost.com/graphics/2020/world/corona-simulator/>.
- R. van der Hofstad, A. Janssen, and J. S. van Leeuwen. Critical epidemics, random graphs and brownian motion with a parabolic drift. *Advances in Applied Probability*, (4):1187–1206, May 12, 2010. doi: 10.1239/aap/1293113157. URL [https://www.win.tue.nl/~rhofstad/critical\\_epidemic\\_fin.pdf](https://www.win.tue.nl/~rhofstad/critical_epidemic_fin.pdf).