Epidemic Simulation Using Percolation

Qijian Liu

**Abstract**

This study presents a simple model for COVID-19, which was used to simulate the effects of the epidemic when no one is taking protective measures against the virus. It is found that in this scenario, the virus can spread quickly, and infect 91.31% of the total population on an average of 136.4 days. It is also found, by running the simulation until the active case counts became 0, that the threshold for herd immunity is 91.20% for the model in being studied.

**Introduction**

One of the most prominent current issues, if not the most prominent, is the COVID-19 outbreak that continues to ravage the world. COVID-19 is a highly infectious disease that was first discovered in late 2019, and the outbreak was declared a pandemic by the World Health Organization (WHO) on March 11th, 2020 (WHO, 2020). To combat against such a dangerous disease, modelling how the disease spread is necessary, as it provides important information like how fast it spreads and how deadly it is. A method that can simulate the spread is by using a model of percolation (Moore & Newman, 2000). A simple model for COVID-19 based on percolation will be the focus of this study.

Percolation can be described as the movement of fluid through porous materials. The model can be constructed using a *n*-by-*n* grid of “sites”. Each site is either open or blocked, and the system is said to percolate if there is a full site, which is an open site that can be connected to the top row through a chain of neighboring (left, right, up, or down) open sites (Sedgewick & Wayne, 2008). For the purpose of simulating an epidemic, the study will modify this model described.

The goal of this study is to determine the following: the effects of not practicing social distancing and taking proper measures against the spread of the disease, and the threshold for herd immunity.

**Methods**

**I. Setting Up the Grid**

To construct the model, the grid of sites will need to be constructed first. To generate this grid, one can first generate a *n*-by-*n* grid full of open sites. Then, randomly select a site in the grid and close it. Repeat this process until number of open sites reach the desire amount. In this study, an open site will be defined as a human, the agent of transmission of the disease. Therefore, the final number of open sites is also the population in the simulation. This study will use a 100-by-100 grid with a total population of 6500.

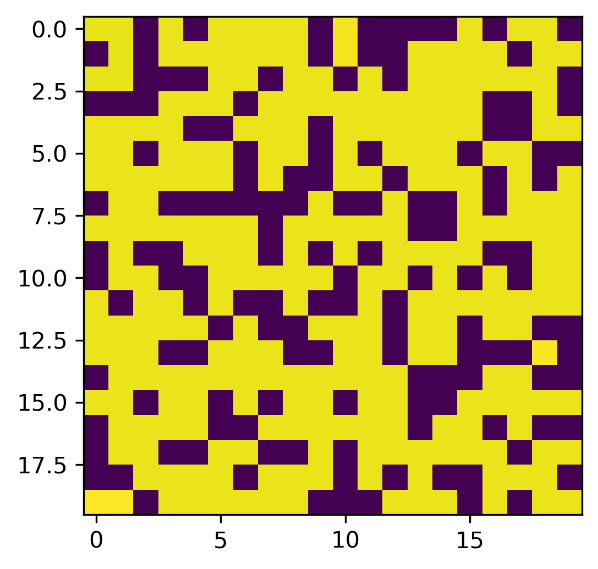


Figure 1: Sample 20 by 20 grid. The yellow tiles are the open sites, and the other colour indicates the blocked sites. In the program, open sites are labeled 1, closed sites are labeled 0.

After generating the grid, the population need to be redistributed into 3 different age groups: youth, adult, and elderly. This is because the transmissibility of COVID-19 is different depending on the age. A study finds that individuals under the age of 20 is about half as susceptible as those over the age of 20, while individuals over 70 years old is significantly more vulnerable than those aged between 20 and 70 (Davies et al., 2020). Based on the population estimate by Statistic Canada, 21% of the Canadian falls into the 0-19 years age group, 66% of the Canadian falls into the 20-69 years age group, and 13% of the Canadian falls into the 70+ years age group (Government of Canada, 2021). These 3 age groups will correspond to the youth, adult, and elderly groups defined thereof. The population will be randomly distributed into one of these 3 groups based on the ratio provided by the population estimate. Additionally, a dictionary containing the status of the person will be generated. The status is a list which includes the current health condition of the person (infected or healthy), the age group the person belongs to, and the number of days the person has contracted the virus. After assigning the age groups, the model will randomly give a number a people the status of infection. These are the initial patients in the epidemic modelled. In this study, the probability of being the initial patients is set to be 1%.

**II. Transmission of COVID-19**

A healthy person can be infected if they are the neighbouring site (left, right, up, or down) of an already infected individual. This model will assume that once infected, the patient will stay ill for 14 days (14 iterations of the loop), as this is the usual period of isolation for anyone who may have been in close contact with an infected individual or is a traveller from abroad **(**Public Health Agency of Canada, 2021). The patient will not be able to infect its neighbouring sites at day 0 and day 14. At day 14, the patient will either recover or they will die from the virus. Based on the latest statistics, the ratio between total deaths from COVID-19 to the total cases of infection in Canada is roughly 1.6% **(**Public Health Agency of Canada, 2021). This will be used as the probability of death for a patient after 14 days.

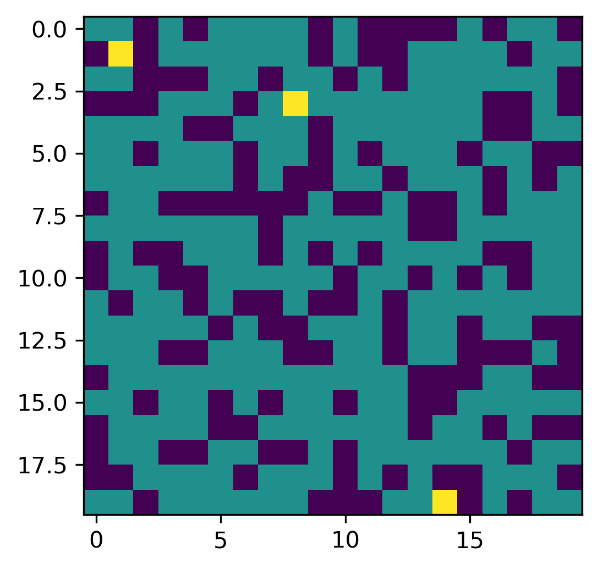


Figure 2: Sample 20 by 20 grid with initial patients coloured in yellow, healthy individuals are coloured in dark green. In the program, healthy individuals are labeled 1, infected individuals are labeled 2.

Whether a healthy individual will be infected by a patient is determined randomly. Since this study’s goal is to investigate the effects of not practicing social distancing and taking the necessary measure against the spread, it is assumed that each of the each of the individuals will come into close contact with their neighbour everyday without wearing a mask. A study shows that the probability of not wearing a mask and coming into close contact (0.5 m) with an infected patient is about 50% (Agrawal & Bhardwaj, 2021). Therefore, 50% is used as the probability of infection for adults and assume that 25% and 75% are the probability of infection for youth and elderly respectively. It must be noted that Agrawal and Bhardwaj’s study was done on the original COVID-19 variant. As of the writing of this report, there are 5 more notable variants of COVID-19 (*Alpha*, *Beta*, *Gamma*, *Delta*, and *Omicron*), each having different rate of transmission. For simplicity, the original’s probability of transmission is used.

**III. Running the Simulation**

The simulation will run until the number of active cases reaches 0 to see how the virus will spread. While recovering from the virus will grant the individual immunity, the new *Omicron* variant seems to evade this immunity (Massey, 2018). Therefore, two scenarios will be simulated: a patient recovering will be granted immunity that protects them from being infected again, and the patient can be infected again as soon as they recover. Since the simulations rely on random numbers, for robustness of the results, 10 simulations with different seeds were done.

**Results**

**Scenario 1: Immunity in Recovered Patients**

On average, the simulation lasted 136.4 days before the total active cases within the population reached 0. In this period, the total number of cases averaged to 5935.4, and the total deaths is averaged 98.5. For detailed result in each of the simulations, see Table 1 in Appendix.

Figure 3 illustrates the trend of active cases and total deaths for one of the simulations. The trends observed in the other instances are similar to the one exemplified.

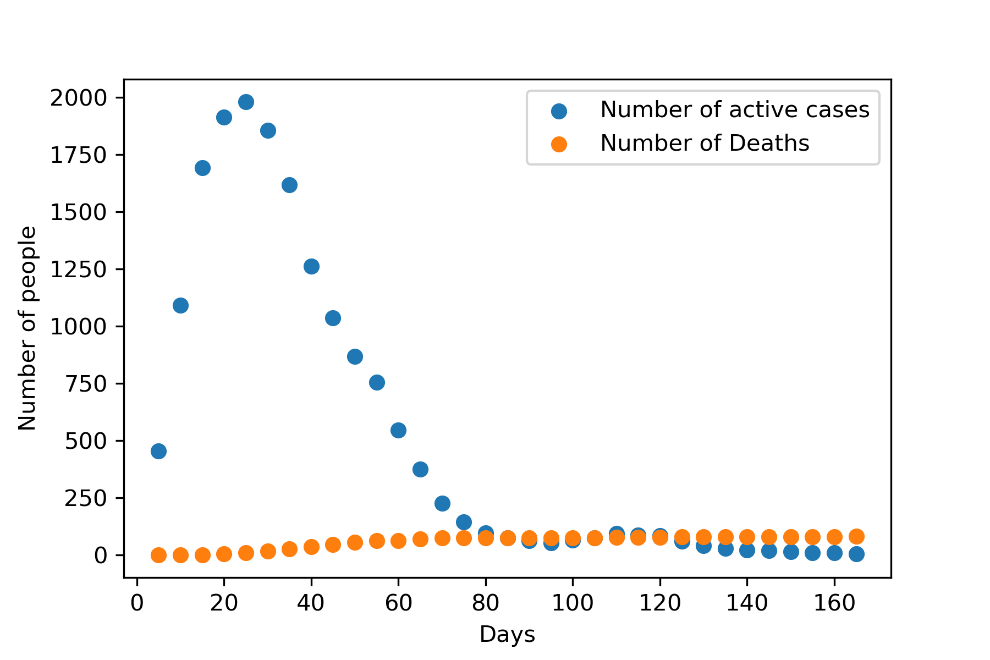
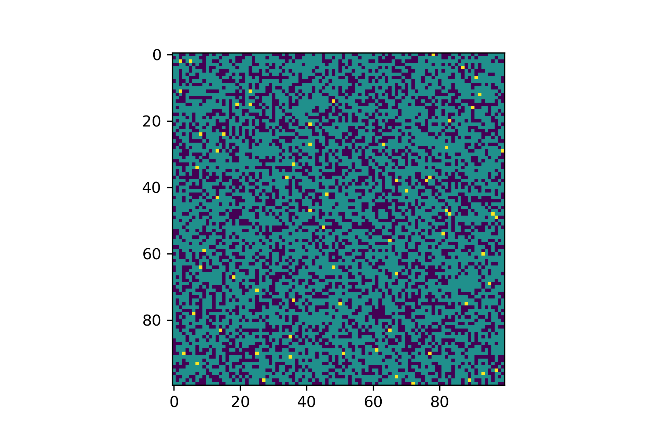


Figure 3: Number of active cases and total deaths vs. time for scenario 1. Seed=1.

Figure 4 shows the initial condition of the epidemic before the simulation, and the final condition after the simulation stopped.

 Chart, qr code

Description automatically generated

Figure 4: Left: Before the epidemic simulation. Initial patients are coloured in yellow, and healthy individuals are coloured in dark green. Right: After the simulation. Healthy patients are coloured in yellow, with the deceased becoming one of the closed sites. Seed=1.

**Scenario 2: No Immunity for Recoveries**

The simulations ran 2741.2 days on average, with an average of 205109.2 cases of infections in total, and an average of 3340.6 total deaths. In this scenario, the ratio of active case counts to total population will cross the 50% threshold in 29.7 days on average. Table 2 and 3 provides a more detail summary of the results in Appendix.

Figure 5 illustrates the trend of active cases and total deaths for one of the simulations. The trends observed in the other instances are similar to the one exemplified.

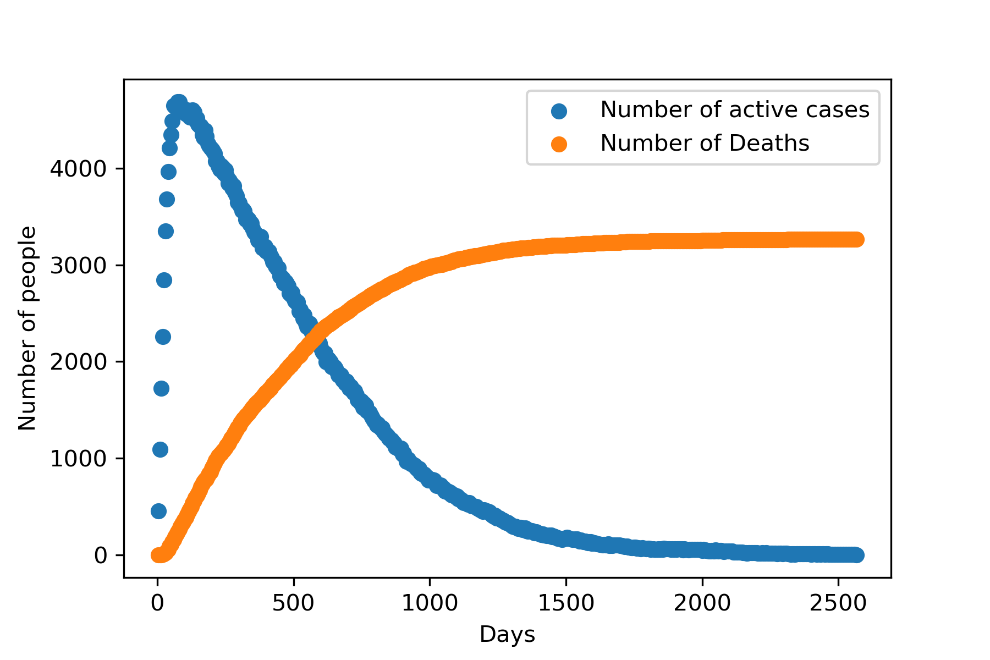
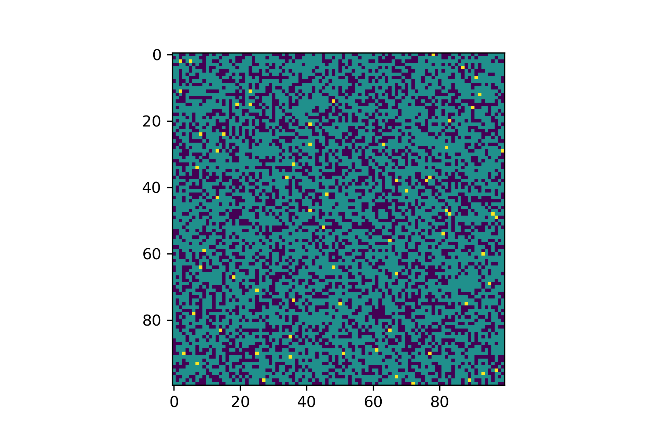


Figure 5: Number of active cases and total deaths vs. time for scenario 2. Seed=1.

Figure 6 shows the initial condition of the epidemic before the simulation, and the final condition after the simulation stopped.

Chart

Description automatically generated

Figure 6: Left: Before the epidemic simulation. Initial patients are coloured in yellow, and healthy individuals are coloured in dark green. Right: After the simulation. Healthy patients are coloured in yellow, with the deceased becoming one of the closed sites. Seed=1.

**Discussion**

For both scenarios, the trends for the total number of active cases are similar. In the start, the number of active cases climb very rapidly. Then, after reaching a peak, the active cases start to gradually decrease down to 0. However, the peak for scenario 2 is more than twice as much as scenario 1. On average, the active case count will cross the 50% threshold of the total population in 29.7 days for scenario 2. Finally, the average length of the simulation is significantly longer for scenario 2, and the average total death count is over 50% of the total population. Furthermore, from the right image of Figure 6, it can be seen that most of the connections between open sites have been severed due to death of patients. This means that in scenario 2, the virus only stopped spreading because of the lack of connections between open sites, which also indicates that being socially distanced can help to stop the spread of the virus.

However, scenario 2 is too extreme to become a reality. The results of scenario 1 are more meaningful. They illustrate how fast the virus can spread when no proper protective measures are taken against it. The total infections averaged to 5935.4. Therefore, in these simulations, it took the virus an average of 136.4 days to infect 91.31% of the total population. In addition to the rapidness of the spread, this also indicate the threshold for herd immunity. At the end of the simulations, the ratio of the people who are immune to the total surviving population is . Because the epidemic stopped when this percentage of people is immune, this should also be the herd immunity threshold for COVID-19. In the early days, the herd immunity threshold was estimated to be in the region of 60-70% (Aschwanden, 2021), but scientists have to rethink this threshold as the pandemic progressed. In an interview with The New York Times, Dr. Anthony Fauci, the current chief medical advisor to the U.S. president, states that their recent studies suggest that it may take close to 90% immunity to stop the virus (McNeil, 2020), which agrees with the findings of this study. This interview was published at the time before any of the notable COVID-19 variants became dominant. Therefore, their result is comparable to the results in this study.

**Future Studies**

There are improvements that can be made over the current model. First, the population is modelled in a 2-D lattice. This means that each individual will only ever come in close contact with up to 4 neighbouring sites, which will also remain the same throughout the entire simulation. This is unrealistic, as social groups in reality are significantly more complicated and dynamic. A more robust model should have a better model for the social network of the population. Second, assumptions were made about the death rate and immunity to keep the model simple. The death rate was assumed to be constant, but in reality, it should depend on the occupancy of the health care system and the age group the patient belongs to. The death rate should be higher when the system is more occupied or when the patient is older. While gaining immunity was assumed to have 100% protection against the virus. More parameters can be added to remove these assumptions. Finally, the efficiency of the program is low when the size of the grid becomes large. Currently, the program has to loop through every element in the *n*-by-*n* grid. An attempt at improvement was made by only looping through the elements that is not a blocked site. This attempt used the function numpy.nonzero. However, runtime analysis indicates that the original program is faster. Therefore, the newer version was abandoned. Nevertheless, improvement on the efficiency should be possible, and should be investigated so that the model can handle a larger population.

**Conclusion**

This study found that the spread of COVID-19 is very rapid when no protective measures are taken, as simulations show that it can infect 91.31% of the total population in 136.4 days on average. The results of the simulations also indicate that the threshold for herd immunity is 91.20%. It is evident, from the results of this study, that it is crucial to have proper protections. These protections could be staying socially distanced or getting immunized.

References

Agrawal, A., & Bhardwaj, R. (2021). Probability of covid-19 infection by cough of a normal person and a super-spreader. *Physics of Fluids*, *33*(3), 031704. <https://doi.org/10.1063/5.0041596>

Aschwanden, C. (2021). Five reasons why covid herd immunity is probably impossible. *Nature*, *591*(7851), 520–522. <https://doi.org/10.1038/d41586-021-00728-2>

Davies, N. G., Klepac, P., Liu, Y., Prem, K., Jit, M., Pearson, C. A., Quilty, B. J., Kucharski, A. J., Gibbs, H., Clifford, S., Gimma, A., van Zandvoort, K., Munday, J. D., Diamond, C., Edmunds, W. J., Houben, R. M., Hellewell, J., Russell, T. W., Abbott, S., … Eggo, R. M. (2020). Age-dependent effects in the transmission and control of COVID-19 epidemics. *Nature Medicine*, *26*(8), 1205–1211. <https://doi.org/10.1038/s41591-020-0962-9>

Government of Canada, Statistics Canada. (2021, September 29). Population estimates on July 1st, by age and sex. Retrieved December 17, 2021, from <https://doi.org/10.25318/1710000501-eng>

Massey, N. (2021, December 18). *Omicron 'largely evades immunity from past infection or two vaccine doses'*. The Independent. Retrieved December 18, 2021, from <https://www.independent.co.uk/news/uk/delta-omicron-pfizer-biontech-imperial-college-london-b1978106.html>

Moore, C., & Newman, M. E. J. (2000). Epidemics and percolation in small-world networks. *Physical Review E*, *61*(5), 5678–5682. <https://doi.org/10.1103/physreve.61.5678>

Mcneil, D. G. (2020, December 24). *How much herd immunity is enough?* The New York Times. Retrieved December 17, 2021, from <https://www.nytimes.com/2020/12/24/health/herd-immunity-covid-coronavirus.html>

Public Health Agency of Canada (2021, December 17). *Covid-19 Daily Epidemiology update*. Canada.ca. Retrieved December 17, 2021, from <https://health-infobase.canada.ca/covid-19/epidemiological-summary-covid-19-cases.html>

Public Health Agency of Canada. (2021, November 30). *For unvaccinated travellers without symptoms of COVID-19 arriving in Canada by land, air or water*. Canada.ca. Retrieved December 10, 2021, from <https://www.canada.ca/en/public-health/services/publications/diseases-conditions/2019-novel-coronavirus-information-sheet.html>

Sedgewick, B., & Wayne, K. (2008). *Programming assignment 1: Percolation*. Princeton University. Retrieved December 6, 2021, from <https://coursera.cs.princeton.edu/algs4/assignments/percolation/specification.php>

World Health Organization. (2020, March 11). *Who director-general's opening remarks at the media briefing on COVID-19 - 11 March 2020*. World Health Organization. Retrieved December 15, 2021, from <https://www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020>

Appendix

Table 1: Results of simulating until 0 active cases with immunity

|  |  |  |  |
| --- | --- | --- | --- |
| Seed | Total days | Total infections | Total deaths |
| 1 | 169 | 5892 | 81 |
| 2 | 103 | 6014 | 91 |
| 3 | 172 | 5920 | 103 |
| 4 | 139 | 6058 | 105 |
| 5 | 147 | 5835 | 128 |
| 6 | 147 | 6026 | 103 |
| 7 | 113 | 5918 | 97 |
| 8 | 115 | 5886 | 94 |
| 9 | 145 | 5856 | 93 |
| 10 | 114 | 5949 | 90 |
| Average | 136.4 | 5935.4 | 98.5 |

Table 2: Results of simulating until 0 active cases without immunity

|  |  |  |  |
| --- | --- | --- | --- |
| Seed | Total days | Total infections | Total deaths |
| 1 | 2565 | 203126 | 3266 |
| 2 | 2940 | 199721 | 3367 |
| 3 | 2191 | 204212 | 3301 |
| 4 | 2782 | 216391 | 3433 |
| 5 | 2600 | 207582 | 3277 |
| 6 | 3147 | 210561 | 3359 |
| 7 | 2730 | 198772 | 3321 |
| 8 | 2595 | 202031 | 3317 |
| 9 | 2702 | 201474 | 3341 |
| 10 | 3160 | 207222 | 3424 |
| Average | 2741.2 | 205109.2 | 3340.6 |

Table 3: Results at when the active case count is equal to at least 50% of the population for scenario 2

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Seed | Total days | Total active | Total infections | Total deaths |
| 1 | 29 | 3266 | 1810 | 18 |
| 2 | 31 | 3319 | 2124 | 29 |
| 3 | 34 | 3257 | 2611 | 40 |
| 4 | 25 | 3371 | 1316 | 23 |
| 5 | 34 | 3260 | 2510 | 40 |
| 6 | 27 | 3329 | 1578 | 27 |
| 7 | 29 | 3335 | 1901 | 33 |
| 8 | 28 | 3309 | 1552 | 30 |
| 9 | 36 | 3291 | 2724 | 38 |
| 10 | 24 | 3282 | 1425 | 23 |
| Average | 29.7 | 3301.9 | 1955.1 | 30.1 |