INFO6205

Program Structures & Algorithms

Spring 2021

Virus transmission simulation

Team Members

Anjali Sajeevan - NUID 001563277
Manasa Bhat - NUID 001032278
Moumita Halder - NUID: 001566147

Table of Contents

1. Introduction

Coronaviruses are a wide group of viruses that typically causes mild to moderate upper respiratory tract infections. Over the past two decades three new coronaviruses have emerged from the animal reservoirs causing serious widespread illness and death.

- SARS-CoV (SARS- severe acute respiratory syndrome)
- MERS-CoV (MERS- Middle East respiratory syndrome)
- SARS-CoV-2 (COVID-19- coronavirus disease 2019)

The COVID-19 is a communicable disease and was first reported in Wuhan, China in December 2019 and since then it has spread around the world. It has been declared a global pandemic by World Health Organization.

- **R- naught**

For communicable diseases in a population with no immunity, the basic reproduction parameter (R-naught or R_0) is the estimated average number of infections produced by a single infectious person. R_0 values indicate how rapidly transmission will occur and if a disease will spread or decline within a community.

If the value of R_0 is >1 an outbreak will continue and will come to an end for R_0 <1. The R_0 values are established using mathematical models and are dependent on the parameters and model hence it varies hugely depending on country, culture, calculation, stages of the outbreak. Any mitigation or containment strategy will aid in reducing the reproduction number, either by decreasing the transmission rate or the time before infectious individuals are isolated.

- **Dispersion parameter(k)**

The dispersion parameter (k) plays an important role in explaining the overdispersion and super-spreading of the SARS-CoV-2 virus, it is the measure of the virus's dispersion. As every infected person does not pass on the virus to the same number of people, this parameter helps in determining if the virus spreads by one person infecting many (big bursts) people or in a steady manner.

If the value of k is small, it means that one infected person can trigger many new cases in a very short time i.e., lower values of k corresponding to a broader distribution. Several studies suggest that the dispersion parameter k for SARS-CoV-2 is on the order of 0.1 [5], approximately 10% of infected people will result in 80% of new infections.

SARS-CoV
- k for SARS-CoV is 0.16 with the negative binomial model (Singapore)[8]
- R_0 is of the range 0.43–2.41 [9]

SARS-CoV-2
- k for SARS-CoV-2 is on the order of 0.1 [5]
- R_0 is of the range 2-3 [6][7]

2. Aim of the project

The aim of the project is to simulate a SARS-CoV-2 virus and analyze the effectiveness of various preventive measures. The following constraints/factors are taken into consideration for the simulation:

- The R-naught and k factor of the disease
- The usage and effectiveness of masks and social distancing
- The prevalence of testing and contact tracing
- The availability and efficacy of the vaccine
- Remote work and schools
- Quarantine
- Group event

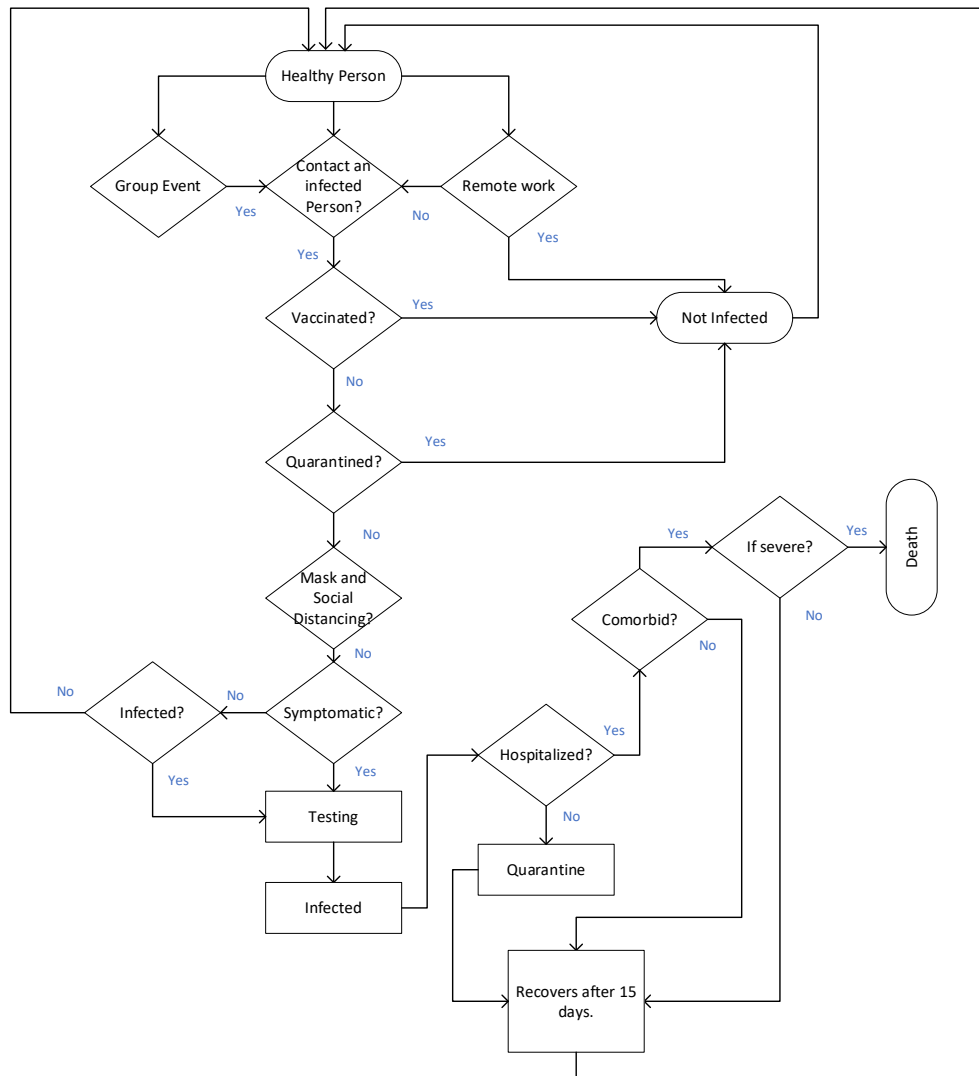## 3. Implementation

3.1. Algorithm Flow Chart

Figure 1: Flow Chart

3.2. Data Structure

The data structures used for the project are:

1. Array - Population: We store the entire population as an array, we loop through each person and based on their position and interaction with other nearby population we update their health conditions accordingly.

2. Directed graph: In this project directed graph is implemented to check contact tracing. We can get information on the people infected by a person. The directional graph is represented using adjacency list. The adjacency list complexity is O(N+M) where N denoted the number of nodes and M the number of edges. The unique id of people is represented as vertices and once they collide, in the direction of collision the edges are stored. A HashSet is used to store the vertices and the edges of each vertices are saved in the LinkedList.

3. Hash map: All the constraints and factors required for the run are stored and retrieved from the HashMap.

3.3. Simulation

The project is developed using Java and the simulation is implemented using Java swing graphical user interface (GUI) and Graphics 2D. The simulation is implemented on a desktop display resolution of 1920x1080 with scaling of 150%. The simulation is modeled for both SARS-CoV-2 and SARS-CoV. The screenshot of the dashboard is shown in Figure 2.

The simulation is based on various assumptions and recent study reports and does not reflect the real-world scenario of SARS-CoV-2 or SARS-CoV, as it would include several other factors. This simulation helps us in getting a better understanding of the importance of the factors that we have considered and the effect they have on the transmission of the disease.

In the simulation each person is represented as a ball, the user can choose the total number of populations from the slider from a range of 50 to 1000. When we start the simulation, initially each person will be given a position, the position of the people keeps changing and based on their collision with other people we update the health status of the person.

The health conditions of the population are represented by various colors as shown in the legend. The constraints can be selected and the equivalent simulation incorporating the constraint will be generated.

The value of R-naught can be defined by the user, the simulation will take it into consideration and during simulation each person can only infect that many numbers of people. In case if the R-naught value is not specified the simulation transmits on a general scenario where a person can infect random number of people.

The real time display of the number of healthy people, infected, severe, recovered, and dead people are provided for both the viruses. A plot of the number of infections over time is displayed.

Once the simulation is stopped, a plot can be generated by clicking on the plot button. The plot displays the number of healthy, infected, recovered, hospitalized and dead over the number of days the simulation was run as shown in Figure 13. The log pane displays the values obtained for each runs of the application – Figure 14. The contact tracing details are displayed on the console of the IDE -Figure 12.
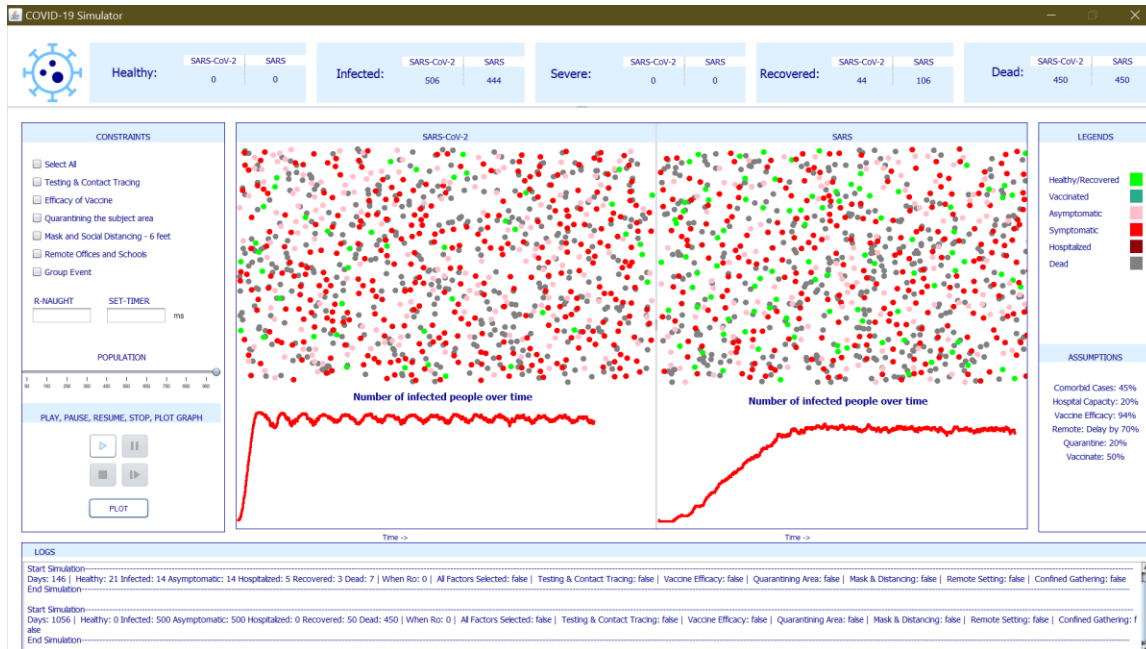

Figure 2: Dashboard

3.4. SARS-CoV (Compared Virus)

When simulating the SARS-CoV (SARS) virus, we take into account the incubation period of 3-5 days that is mapped to 3ms before it starts infecting healthy people. Due to the incubation period and recovery time for the virus, the transmission is comparably lower than the SARS-CoV-2 which results in lower dispersion parameter. When the vaccination factor is selected, it will not be considered for SARS-CoV simulation as the vaccine is still not developed in the real world for this virus.

4. Project Details

4.1. Assumptions

4.1.1. Population types

1. Healthy (Light Green): Susceptible to the virus. They include people with comorbid condition. This also includes recovered individuals.
2. Vaccinated (Dark Green): 94% of the population are protected from the virus.
3. Infected – Symptomatic (Bright Red): Primary transmitter of the virus until tested and quarantined.
4. Infected – Asymptomatic (Pink): Undetected transmitter until tested and quarantined.

5. Hospitalized/Severe (Dark Red): If hospitalized they are immobile, and comorbid condition is checked.
6. Dead (Grey): Immobile and will not further transmit the virus.

4.1.2. Factors considered for simulation

1. R-naught value:
   a. The R-naught parameter is user defined, the transmission changes depending on the R-naught value.
2. Vaccine:
   a. 50 percent of the population is vaccinated.
   b. The vaccine efficiency is considered to be 94% [11].
   c. Individuals with comorbid condition will be vaccinated first.
   d. There is no vaccine currently available for SARS-CoV, hence it is not considered for the comparison.
3. Testing and Contact Tracing:
   a. Testing will detect all the infected population including symptomatic and asymptomatic population. So once testing is considered, the entire simulation will reveal all infected individuals.
   b. Quarantine constraint is also considered if infected.
   c. For contact tracing, the part of the population that comes in contact with the infected person is traced using directed graph.
4. Quarantining the subject area:
   a. 20% of the population will be quarantined.
   b. This will include healthy people who have recently travelled, people with health issues, people practicing self-isolation and infected people.
   c. The quarantined healthy population will not be infected by the virus.
   d. Infected people who quarantine will not be transmitting the virus to nearby healthy people.
5. Mask and Social Distancing – 6ft:
   a. Population displacement will reduce when social distancing is practiced.
   b. In this case the collision(interaction) becomes lesser resulting in less transmission of the virus.
6. Remote Offices and Schools:
   a. Population movement is delayed by 70%.
   b. Due to the delay the spreading reduces, as the recovery time elapses.
7. Group Events:
   a. A closed confined massive gathering event is attended by a part of the population.
   b. A single infected person can infect the entire attending population.

4.1.3. General

1. The recovery time:
   a. Considering the number of population relative to the size of the panel, we have factored the recovery time of 15 days to 5ms.

b. After recovery, they are considered as a healthy population and become susceptible to the virus.
2. Death Rate:
    a. Both infected people with and without comorbidities get hospitalized based on hospital bed availability.
3. Asymptomatic Population:
    a. We have considered one out of every 3 persons infected to be asymptomatic.
4. Hospitalizing severe population:
    a. Severe patients would be hospitalized up to the hospital bed capacity.
    b. The hospital capacity is considered to accommodate 20% of the population.
5. SARS-CoV incubation period:
    a. We have factored the incubation period of 5 to 7 days to 3ms.
    b. After the incubation period it gets transmitted to others.

4.2. Config.ini

The parameters that can be modified in the Config.ini file is given below:
1. Infection Parameters:
    a. infectionRate: This is the percentage of population initially infected. We have taken it as 0.1, which means 10% of population would be infected at the start of simulation.
    b. asymptoticFraction: Fraction of infected people that are asymptomatic. We have taken 3, i.e, 1 out of 3 infected people is asymptomatic.
    c. comorbidPercentage: Percentage of population that are comorbid. Currently considered as 45%.
    d. infectedQuarantinePercentage: When quarantine constraint is selected, this is the percentage of infected population that actually quarantine. Taken as 20%.
2. Vaccine Parameters:
    a. vaccinatedPercentage: This is the percentage of population that are vaccinated when we consider efficacy of vaccine constraint. Currently considered as 50%.
    b. vaccineEffectiveness: Percentage of vaccinate population that do not get infected. Effectiveness is considered as 94%.
3. Population Ball Parameters:
    a. populationBallHeight: Height of the 2D Graphics ball shape components that we have considered for displaying population.
    b. populationBallWidth: Width of the 2D Graphics ball shape components that we have considered for displaying population.
4. Other Conditions:
    a. hospitalCapacity: Capacity of hospital is considered based on percentage of population taken. i.e, if hospitalCapacity is 0.2 then the hospital capacity is 20% of the population.

5. Outputs

The following are the results observed for the simulation for various constraints/factors.

1. **Comparison plot for social distancing and mask condition**

Time(ms)	With Social Distancing and Mask	Without Social Distancing and Mask
1	20	33
2	34	43
3	69	87
4	101	104
5	189	322
6	254	273
7	229	325
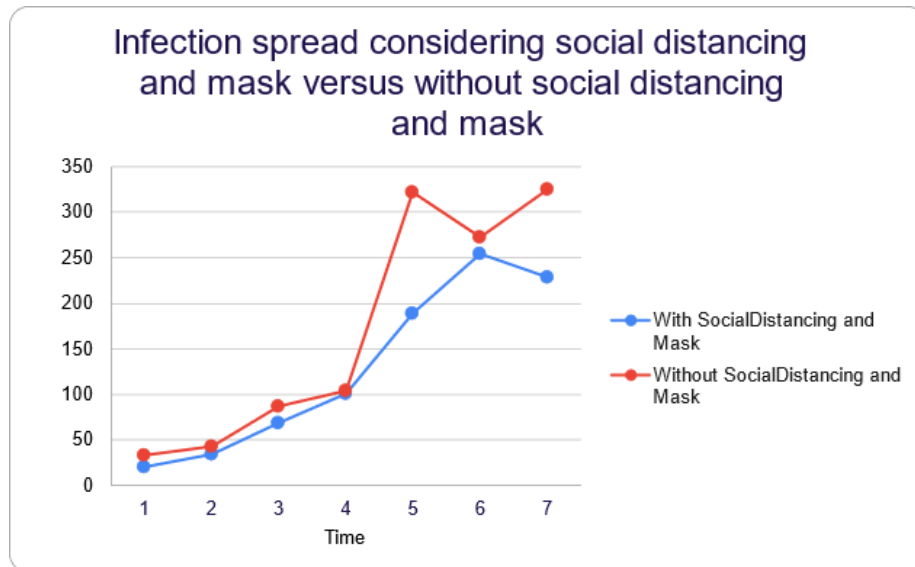
Figure 3:Comparison plot for social distancing and mask condition

We observed that when quarantine is not considered the mean infected count is 208 and when quarantine is considered it is 55. There is 73% reduction of infection if we follow quarantine.

2. **Comparison plot for quarantine condition**

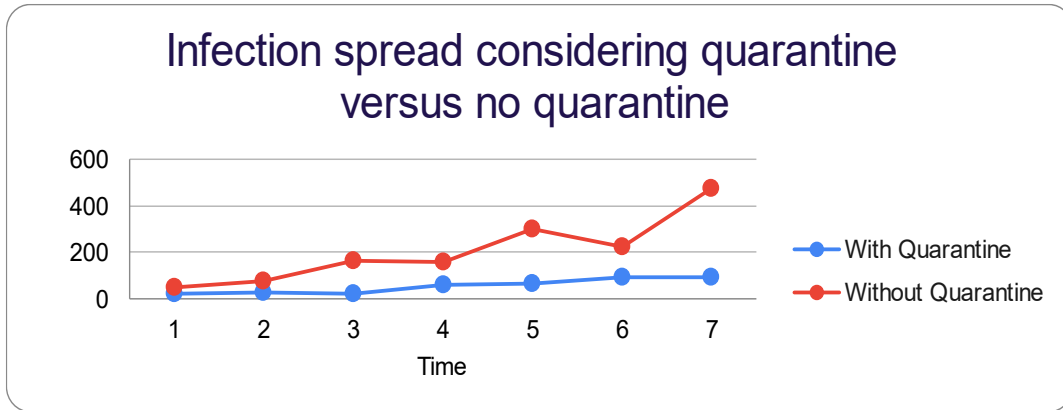| Time(ms)	With Quarantine	Without Quarantine
1	22	49
2	29	79
3	20	166
4	62	161
5	65	302
6	94	226
7	94	474

Figure 4: Comparison plot for quarantine condition

We observed that when social distancing is not followed the mean infected count is 169 and when social distancing is considered the mean infection count is 128. There is 24% reduction of infection if we follow social distancing and use masks.

3. Comparison Plot for Remote Work/School condition

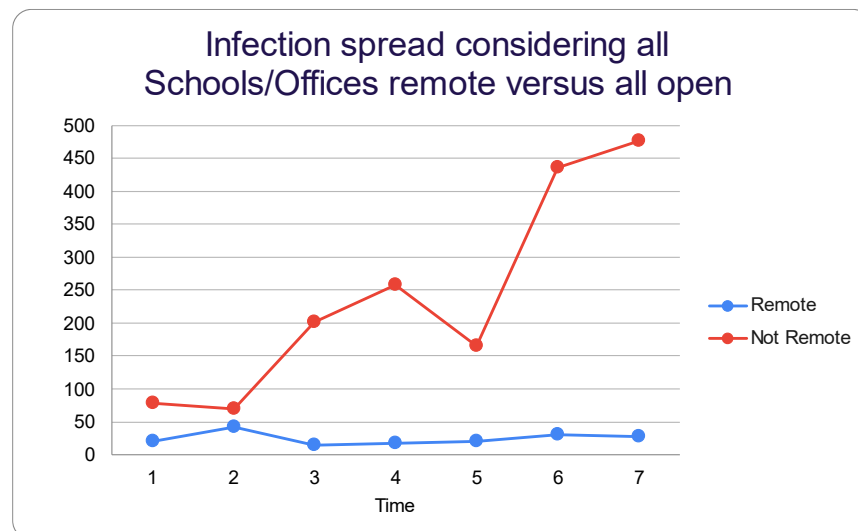| Time(ms)	Remote	Not Remote
1	20	78
2	42	69
3	14	202
4	18	258
5	21	165
6	30	436
7	27	476

Figure 5: Comparison Plot for Remote Work/School condition

We observed that when all Offices and schools are open the mean infected count is 240 whereas the mean count is 24 when offices and schools go remote. There is 90% reduction of infection.

4. Comparison plot for vaccination

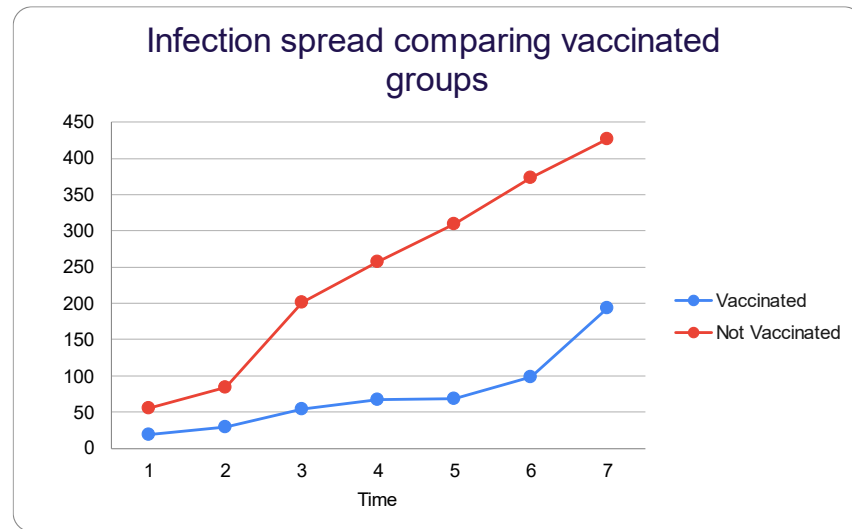| Time (ms)	Vaccinated	Not Vaccinated
1	19	56
2	29	84
3	54	202
4	67	257
5	69	310
6	98	373
7	194	427

Figure 6:Comparison plot for vaccination

We observed that when in group without any vaccination the mean infected count is 244 and when group with vaccination is considered, the mean infection count is 75. There is 69% reduction of infection.

All the above graphs depicted the impact of each factor on the SARS-CoV-2 virus transmission.

5. Comparison by considering all factors and no factors

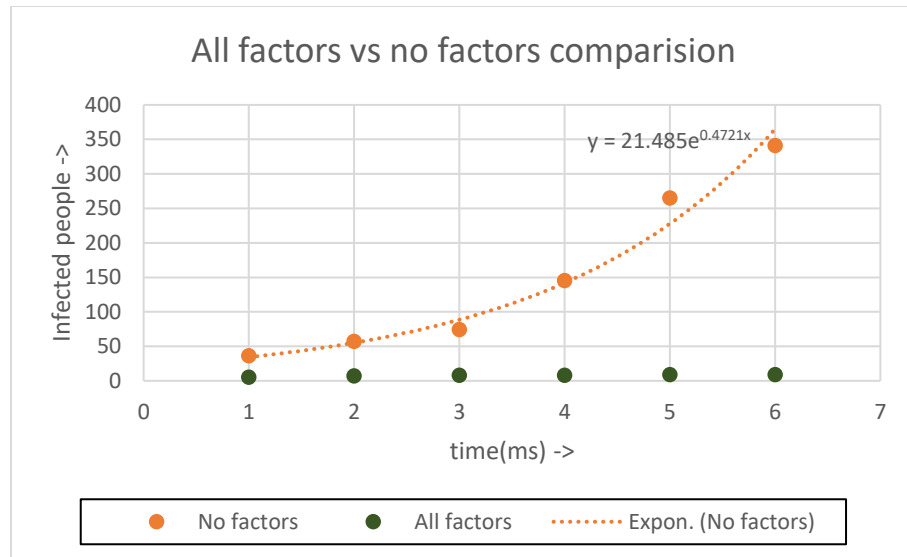| Time(ms)	All Factors	No Factors
1	36	5
2	57	7
3	74	8
4	145	8
5	265	9
6	341	9

Figure 7:Comparison by considering all factors and no factors.

From this graph it can be observed when no factors are considered the number of infected people increase exponentially over time whereas when all factors are considered the growth is logarithmic.

6. **Comparison between different R-naught values**

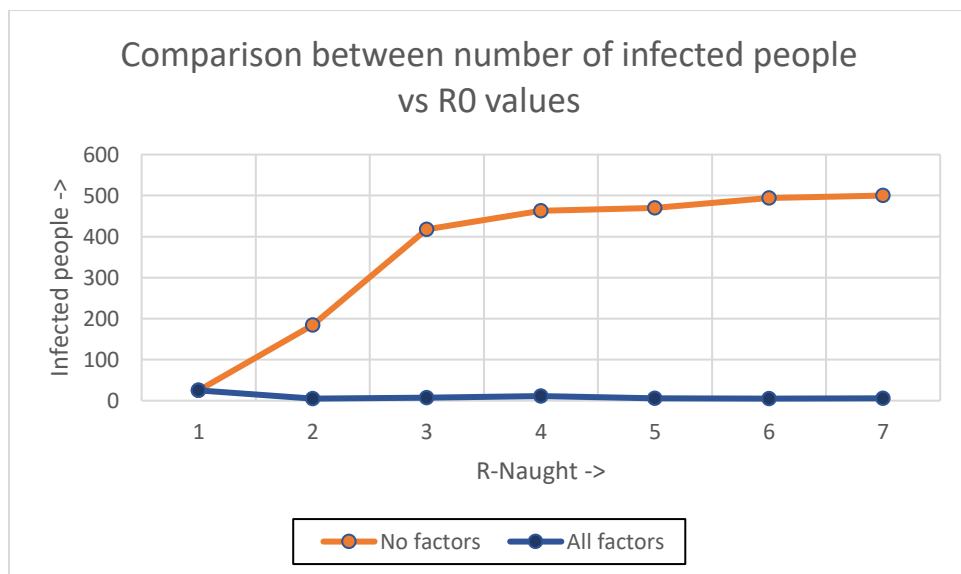| R0	No factors	All factors
1	25	25
2	184	5
3	418	7
4	463	11
5	470	6
6	494	5
7	500	6

Figure 8:Comparison between different R-naught values.

12

This graph depicts the relation of the basic reproduction parameter with virus transmission for a constant time(5ms). The population was taken as 500 for this simulation and we can observe that by R0 = 7 all people were infected.

7. Comparison between SARS-CoV-2 and SARS-CoV without any constraints.

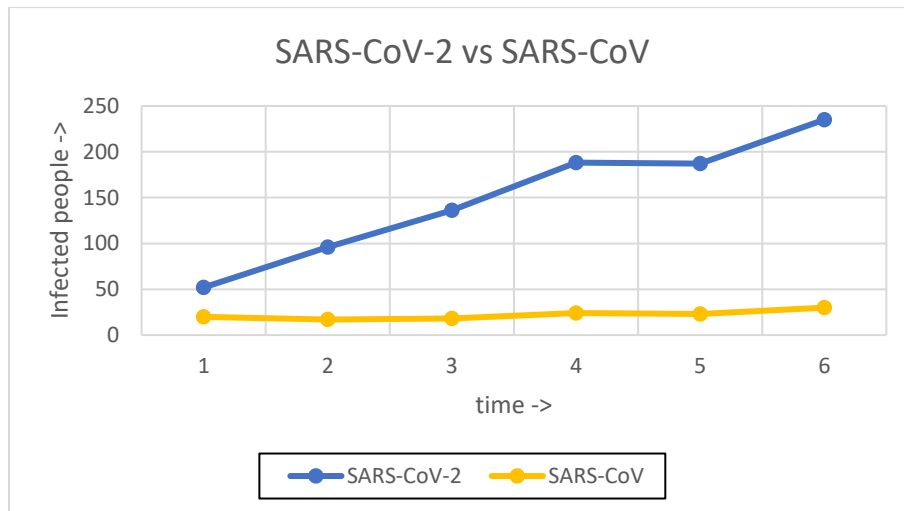| Time(ms)	SARS-CoV-2	SARS-CoV
1	52	20
2	96	17
3	136	18
4	188	24
5	187	23
6	235	30

Figure 9:Comparison between SARS-CoV-2 and SARS-CoV

As per the graph it can be analyzed that the number of infected people increases linearly over time for SARS-CoV-2 whereas for SARS-CoV it increases only by a small fraction.

8. **Comparison between R0=0 value vs R0=1 and R0=2.5 over time.**

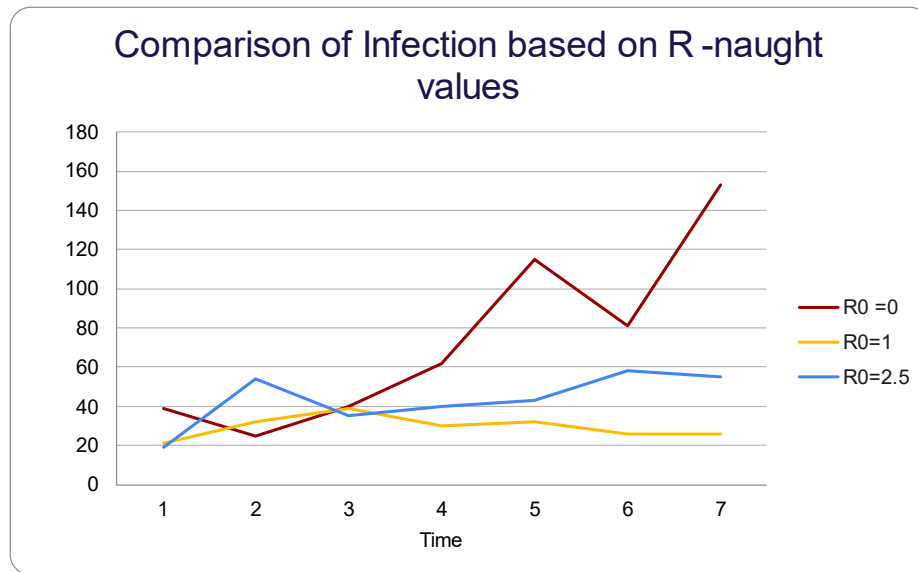| Time(ms)	R0 =0	R0=1	R0=2.5
1	39	21	19
2	25	32	54
3	40	39	35
4	62	30	40
5	115	32	43
6	81	26	58
7	153	26	55

Figure 10:Comparison between no R0 value vs R0=1 and R0=2.5

When R0 is not considered the spreading is very high compared to R0=1 and R0=2.5. This implies that R0 alone is not efficient enough.

9. Comparison plot for infection based on group event over time with and without factors.

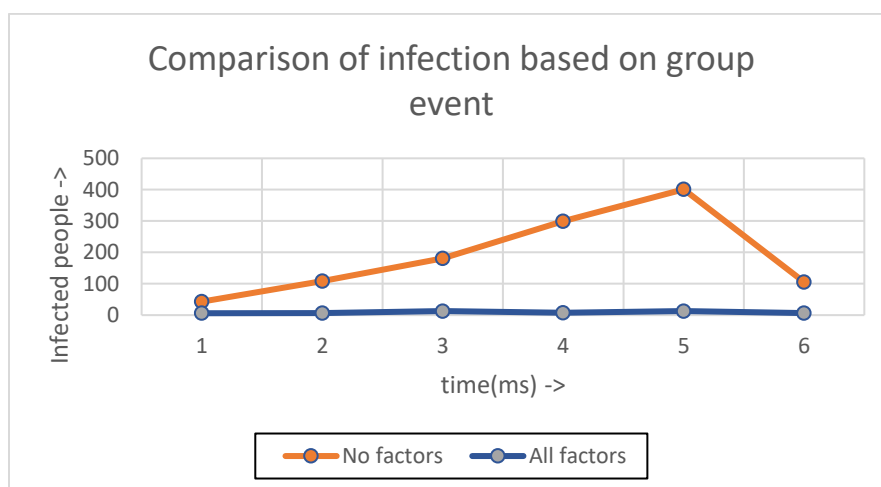| Time (ms)	No factors	All factors
1	43	6
2	109	7
3	181	13
4	299	8
5	401	13
6	105	7

Figure 11 : Comparison plot for group event

14

It can be observed from the graph that if at least one infected person is present inside the group it can infect everyone, and result is a drastic spike if no factors/constraints are considered. This brings the dispersion constant k into the picture. Based on our analysis of the R0 comparison plots figure 9 and 11, we can infer that the basic reproductive parameter alone is not efficient enough to determine the transmission. In cases like the group event, a single person can transmit the virus to multiple people whereas in other scenarios a person may not even transmit to another person.

10. Contact Tracing

```
--------------------Contact Tracing----------------------------
Person: A472053 -> infected: A309125 |
Person: A704361 ->
Person: A357641 -> infected: A704361 |
Person: A309125 -> infected: A357641 |
```

Figure 112: Contact tracing screenshot with all constraints.

On the IDE console after each run the contact tracing for each person is displayed. It shows that person A234795 has infected person A309125. The contract tracing is checked only when an infected person comes in contact with a healthy person. The number of infected people when all constrains are selected is low.
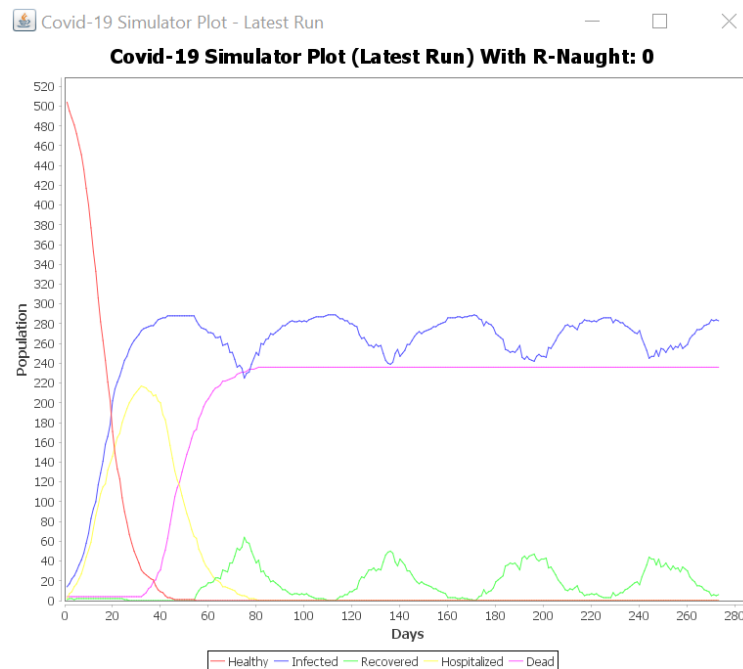
11. Result Plot

Figure 123: Contact tracing screenshot with all constraints.

12. Log pane

Figure 134: Log pane

6. Mathematical Analysis/Evidence

6.1. Big O

It can be observed from Figure 7 that when considering the scenario when no factors/ constraints are selected or followed by people, the transmission increases for infected people and the growth is exponential – O (2^n) **.** By incorporating the algorithm that including all factors or constraints the growth was improved to logarithmic time **-** O(log(n)).

6.2. Entropy

The entropy of the project is in the beginning of the simulation the available data is minimum this is when the entropy is at the highest. With time we analyze the effect of factors/constraints on the transmission of the virus, this results in entropy gets lower.

6.3. Invariant

1. Array:
   a. Whenever a person collides with another person the position will change, and the status will be updated accordingly.

2. Diagraph Graph:
   a. When an infected person collides with a healthy person, the direction of edge is only from infected to healthy person.
   b. There will not be any duplicated vertices.

7. Unit tests result:

The unit test coverage for the project is 72.9%. The snapshot of unit test results are given below:

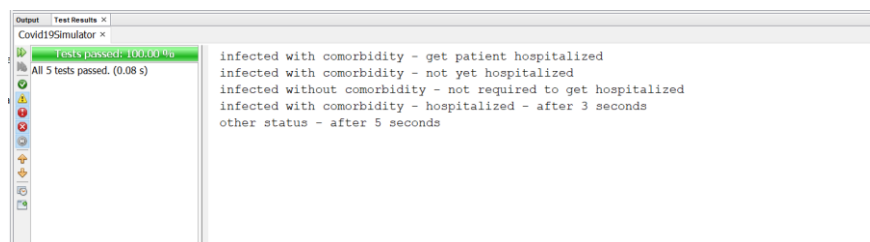1. Unit test for Population package:

   a. CountStatusTest.java

Figure 145: Unit test for CountStatusTest.java

16

b. PopulationTest.java

Figure 15:Unit test for PopulationTest.java

2. Unit test for Simulation package:

a. PopulationPaintPanelTest.java

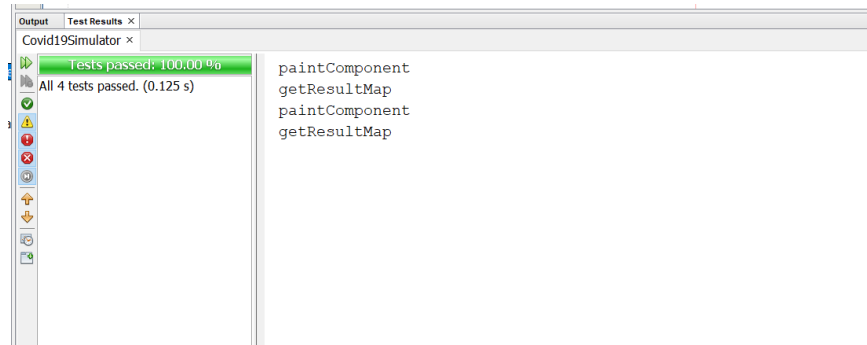*Figure 167:Unit test for PopulationPaintPanelTest.java*

3. Unit test for UI package:

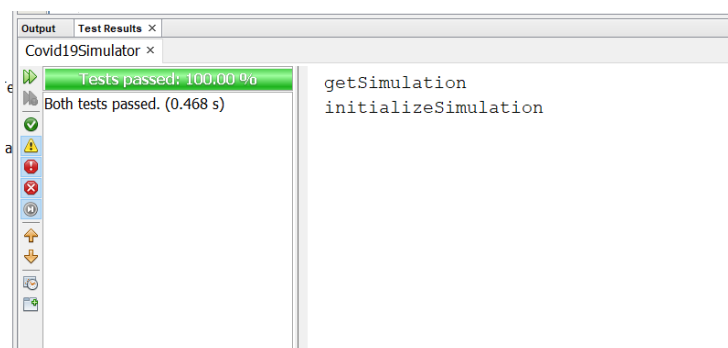a. DashboardTest.java

Figure 178:Unit test for DashboardTest.java
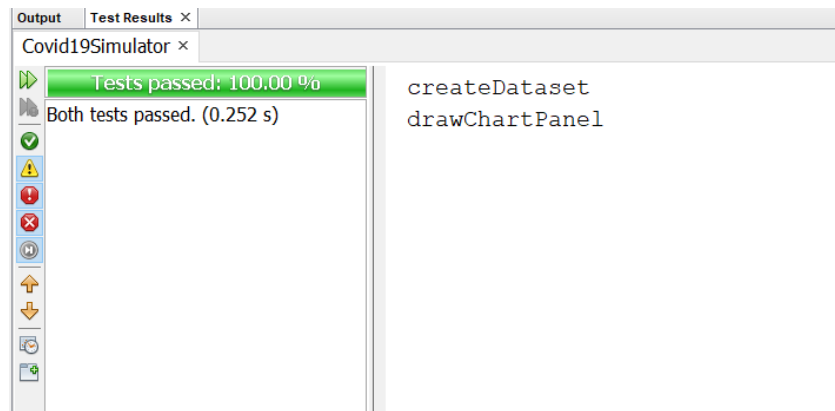
b. PlotChartJFrameTest.java

Figure 19: Unit test for PlotChartJFrame

8. Conclusion

The spread of SARS-CoV-2 and SARS-CoV virus was simulated and analyzed. When two virus - SARS-CoV-2 and SARS-CoV of different k/R factors was compared, the spreading was relatively smaller for the SARS-CoV virus.

The R-naught value that estimates the average number of transmissions by a single infectious person does not give us the actual virus transmission rate when compared to real life scenarios where one person can transmit the infection to random number people based on the dispersion parameter(k) in case of events or gatherings.

It can be concluded from the results that when no precautions(factors/constraints) are considered the transmission growth rate is exponential, this transmission growth rate was improved to be logarithmic time when all precautions (factors/constraints) were incorporated.

9. Reference:

1. *Healthcare Workers*. (2020, February 11). Centers for Disease Control and Prevention. https://www.cdc.gov/coronavirus/2019-ncov/hcp/planning-scenarios.html
2. *Coronaviruses*. (2021, March 26). NIH: National Institute of Allergy and Infectious Diseases. https://www.niaid.nih.gov/diseases-conditions/coronaviruses

3. Linka, K., Peirlinck, M., & Kuhl, E. (2020). The reproduction number of COVID-19 and its correlation with public health interventions. *Computational Mechanics*, *66*(4), 1035–1050. https://doi.org/10.1007/s00466-020-01880-8

4. Kleczkowski, A. (2020, June 16). *Is the K number the new R number? What you need to know*. The Conversation. https://theconversation.com/is-the-k-number-the-new-r-number-what-you-need-to-know-140286

5. Sneppen, K., Nielsen, B. F., Taylor, R. J., & Simonsen, L. (2021). Overdispersion in COVID-19 increases the effectiveness of limiting nonrepetitive contacts for transmission control.

Proceedings of the National Academy of Sciences, *118*(14), e2016623118.
https://doi.org/10.1073/pnas.2016623118

6. Zhang, S., Diao, M., Yu, W., Pei, L., Lin, Z., & Chen, D. (2020). Estimation of the reproductive number of novel coronavirus (COVID-19) and the probable outbreak size on the Diamond Princess cruise ship: A data-driven analysis. *International Journal of Infectious Diseases*, *93*, 201–204. https://doi.org/10.1016/j.ijid.2020.02.033

7. Endo, A., Abbott, S., Kucharski, A. J., & Funk, S. (2020). Estimating the overdispersion in COVID-19 transmission using outbreak sizes outside China. *Wellcome Open Research*, *5*, 67. https://doi.org/10.12688/wellcomeopenres.15842.1

8. Chun, B. C. (2016). Understanding and Modeling the Super-spreading Events of the Middle East Respiratory Syndrome Outbreak in Korea. *Infection & Chemotherapy*, *48*(2), 147. https://doi.org/10.3947/ic.2016.48.2.147

9. Chowell, G., Castillo-Chavez, C., Fenimore, P. W., Kribs-Zaleta, C. M., Arriola, L., & Hyman, J. M. (2004). Model Parameters and Outbreak Control for SARS. *Emerging Infectious Diseases*, *10*(7), 1258–1263. https://doi.org/10.3201/eid1007.030647

10. He, D. (2020, October 16). *Low dispersion in the infectiousness of COVID-19 cases implies difficulty in control*. BMC Public Health. https://bmcpublichealth.biomedcentral.com/articles/10.1186/s12889-020-09624-2

11. STAT. (2021, April 14). *Comparing three Covid-19 vaccines: Pfizer, Moderna, J&J*. https://www.statnews.com/2021/02/02/comparing-the-covid-19-vaccines-developed-by-pfizer-moderna-and-johnson-johnson/

12. Howard, J. (2021, January 26). *An evidence review of face masks against COVID-19*. PNAS. https://www.pnas.org/content/118/4/e2014564118