

An Analysis of Forwards vs Backwards Contact Tracing in a Pandemic

Introduction:

The world was recently introduced to the challenges brought by a pandemic by the Covid-19 virus. To ensure that health care workers were not overwhelmed, most states and counties stressed the importance of ‘Flattening the Curve’ by urging citizens to reduce their exposure to others as much as possible. However, not all citizens will respect the guidelines set out by local governments until they themselves are put in danger. This means that short of a state mandated lockdown, it is difficult to stop a spread of infection from peer-to-peer interaction. There is also the issue of ‘super-spreaders’ and ‘super-spreader’ events where, due to one or few people, a large amount of people are exposed to the virus at one time.

One of the tools that governments have taken advantage of is contact tracing. This is relatively easy for governments to implement, as it usually can be done entirely over the phone with citizens. Contact tracing can be conducted in a few ways. The first of which is Forward Contact Tracing, which involves identifying people who have been exposed to the virus by contacting a currently infected individual. The second way is Backwards Contact Tracing, which attempts to retrace an infected individual’s steps to locate the source of their infection. The last method would be a combination of the previous two methods.

Forward Contact Tracing requires the least effort for governments to implement because the citizen is only required to give the names of the few people they have been in contact with, and then those people are warned that they may have been exposed. Both Boulder County and Colorado Public Health focused on forward contact tracing during the Covid-19 pandemic (3)(4). The issue with Forward Contact Tracing is that many cases go undetected. A study done in Austria in April 2020 suggested that only 13% of Covid-19 infections were being detected (2). This suggests that forward contact tracing may not be an ideal solution, as it does not help identify new cases, put only hopes to prevent further damage.

On the other hand, Backwards Contact Tracing is effective at finding undetected cases of a disease. A paper published by *Nature Physics* (2021) proved that Backwards Contact Tracing is effective at locating ‘Super Spreaders’ due to a phenomenon called the ‘Friendship Paradox’ (1). In short, this paradox relies on the fact that any node in a network is more likely to be connected to a high degree node than a low degree one. This means that through Backwards Contact Tracing, you are more likely to find nodes that are spreading the disease more than the currently identified node.

The last form of contact tracing would be a combination of Forward and Backward Contact Tracing. This would involve warning exposed citizens, as well as tracking the source of the infection.

The purpose of this paper is twofold: Using the SIR model and a simulated network, how effective is contact tracing in general, and which form of contact tracing is most effective? We have seen that most local governments in the Colorado area are relying on forward contact tracing, but is that the best course of action? Should they instead focus their efforts on Backwards Contact Tracing?

Methods:

SIR Model:

To simulate a pandemic, I used the SIR model (Susceptible, Infected, Recovered). The modification I made was to add a ‘Quarantined’ parameter. A node in the network could be ‘Quarantined’ when they were infected or susceptible, but once ‘Quarantined’ the node could not spread the disease or contract it. The parameters for my SIR model were loosely based on Covid-19.

- $\text{Beta} = 0.3$
- $\text{Gamma} = 0.071 = 1/14$
- $R_0 = 4.225$

A study done based on China’s pandemic numbers suggested an R_0 of 5.7 for Covid-19, but other studies mentioned in the same paper have reported much lower values for R_0 (5). My Gamma was chosen

to reflect a 14 day (average) recovery period. My Beta was chosen somewhat arbitrarily, but it was not changed to a different value because it generates a reasonable R_0 value.

Barabási-Albert Graph:

To simulate a population, I used the Barabási-Albert Graph. This was used in the paper published in *Nature Physics* (2021) to simulate a population with high degree nodes (1). To create the model, I used the pre-made NetworkX package that follows the BA Graph model. The network is based on growth and preferential attachment, where every node has a specified minimum number of edges, and every new node is more likely to be connected to a higher degree node than to a lower degree node. The following CCDF shows the degree distribution for the network I created.

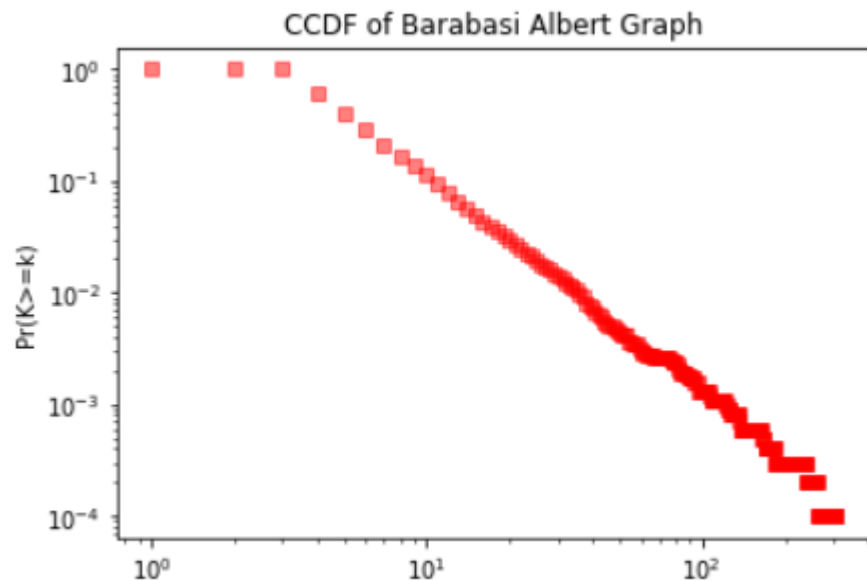


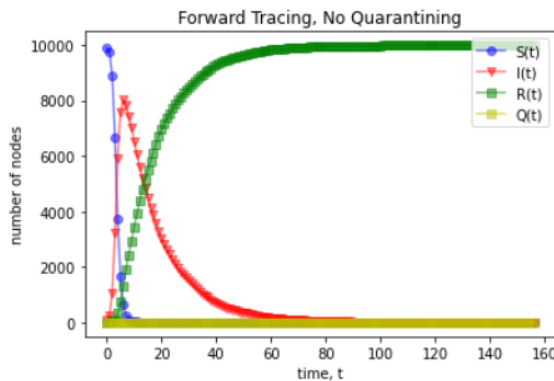
Figure 1: CCDF of 10,000 node Barabási-Albert Graph used for the simulations in this report. The mean degree of this network is $k = 5.9982$. The number of edges is $m = 29991$. The distribution shows that most of the nodes have a degree of 10 or less, while a select few have a high degree of over 100.

Results:

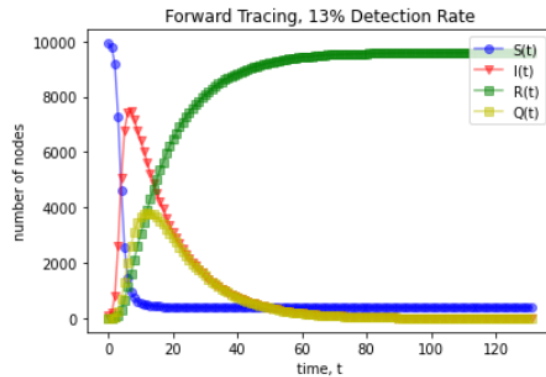
The first step in conducting the simulation was modifying the SIR model to reflect quarantining from the different types of contact tracing. The first type, Forward Contact Tracing was done by first

running a typical SIR iteration, which would be updating to find newly infected individuals, and updating from infected to recovered. To modify this, I added a third step which involved identifying infected individuals with probability 0.13. This was based off the study I mentioned earlier that was conducted in Austria (2). Next, we quarantine that individual, which means they can no longer infect other susceptible nodes, and we identify all their neighbors who aren't yet infected with probability 0.2. This is done to reflect how Forward Contact Tracing looks for exposed individuals, but not for the source of the infection. The neighbors identified are also quarantined and can not be infected until they 'recover', which is done by the same process as an infected individual would recover (so an average of 14 days in quarantine). The following charts show a simulation of 10,000 nodes that undergo this SIR simulation using Forward Contact Tracing with a detection rate of 0%, 13% (the relevant case), and for comparison's sake I also did 30%.

Safe: 13, Infected: 0, Recovered 9987, Quarantined: 0
Mean Degree of Quarantined Nodes: 0



Safe: 399, Infected: 0, Recovered 9601, Quarantined: 0
Mean Degree of Quarantined Nodes: 5.991809494918853



Safe: 1269, Infected: 0, Recovered 8731, Quarantined: 1
Mean Degree of Quarantined Nodes: 6.214240942710292

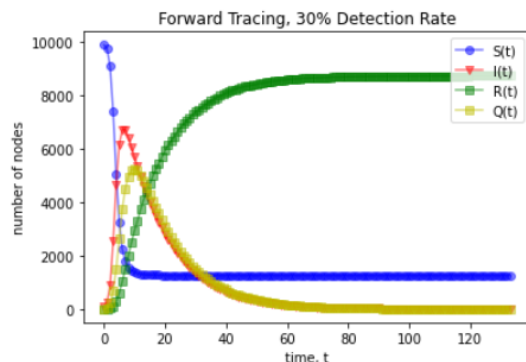
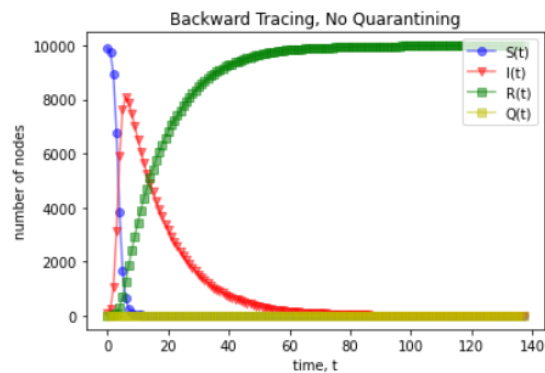


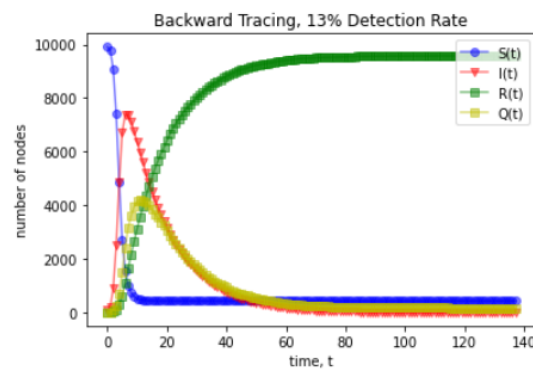
Figure 2: These 3 graphs show the results from 1 round of SIR simulation using Forward Contact Tracing. Note that as the detection rate for infection rose, the number of Safe (or uninfected, or susceptible) nodes that remained at the end also rose. Also note that the mean degree for quarantined nodes was between 5.9 and 6.2.

Next was the simulation using Backwards Contact Tracing. This was done in a similar way to Forward Contact Tracing, except instead of quarantining neighboring nodes that were exposed, we quarantined neighboring nodes that were infected. The following charts show a simulation of 10,000 nodes that undergo this SIR simulation using Backwards Contact Tracing with a detection rate of 0%, 13% (the relevant case), and for comparison's sake I also did 30%.

Safe: 11, Infected: 0, Recovered 9989, Quarantined: 0
Mean Degree of Quarantined Nodes: nan



Safe: 424, Infected: 0, Recovered 9576, Quarantined: 170
Mean Degree of Quarantined Nodes: 6.437800787516407



Safe: 1368, Infected: 0, Recovered 8632, Quarantined: 259
Mean Degree of Quarantined Nodes: 6.538329820864067

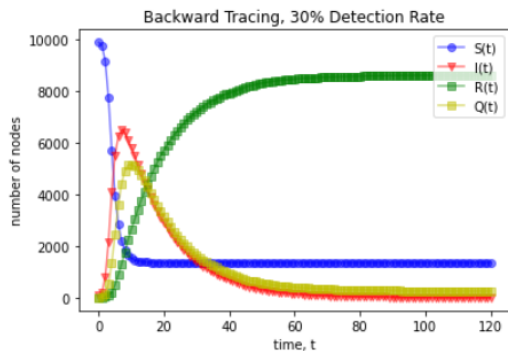
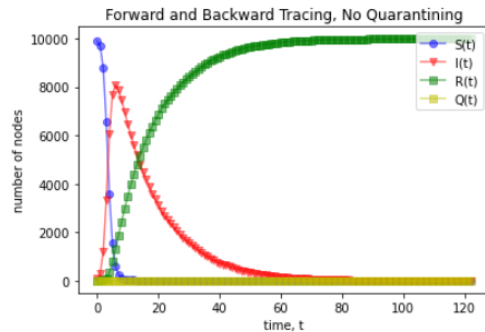


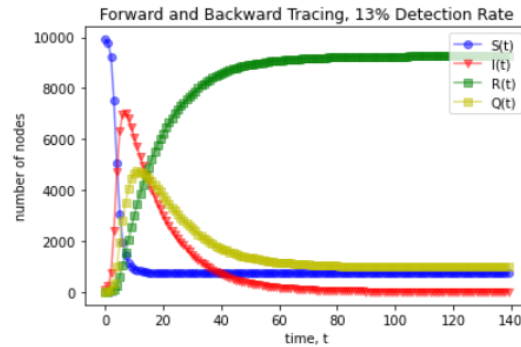
Figure 3: These 3 graphs show the results from 1 round of SIR simulation using Backward Contact Tracing. Note that as the detection rate for infection rose, the number of Safe (or uninfected, or susceptible) nodes that remained at the end also rose. Also note that the mean degree for quarantined nodes was between 6.4 and 6.5.

Lastly, I conducted the same simulation but with a combination of Forward and Backward Contact Tracing. This involved quarantining all neighboring nodes of the identified infected node with probability 0.2. The following charts show a simulation of 10,000 nodes that undergo this SIR simulation using Forward and Backward Contact Tracing with a detection rate of 0%, 13% (the relevant case), and for comparison's sake I also did 30%.

Safe: 14, Infected: 0, Recovered 9986, Quarantined: 0
Mean Degree of Quarantined Nodes: nan



Safe: 745, Infected: 0, Recovered 9255, Quarantined: 990
Mean Degree of Quarantined Nodes: 7.073882293508461



Safe: 1923, Infected: 0, Recovered 8077, Quarantined: 833
Mean Degree of Quarantined Nodes: 6.9472884525479195

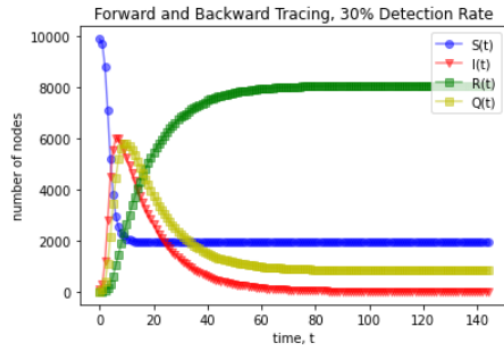


Figure 4: These 3 graphs show the results from 1 round of SIR simulation using Backward and Forward Contact Tracing. Note that as the detection rate for infection rose, the number of Safe (or uninfected, or susceptible) nodes that remained at the end also rose. Also note that the mean degree for quarantined nodes was between 6.9 and 7.1.

Lastly, for a more concrete comparison, I ran all three simulations using the 0.13 detection rate to find the number of uninfected individuals remaining at the end of each simulation. The following chart reflects the results from this.

| Type of Contact Tracing | Forward Contact Tracing | Backward Contact Tracing | Forward and Backward Contact Tracing |
|--------------------------|-------------------------|--------------------------|--------------------------------------|
| Average Uninfected Nodes | 394.9 | 466.5 | 710.7 |

Figure 5: This chart shows the average number of nodes that remained uninfected at the end of the SIR simulations. This is the average of 10 simulations run using each individual model. The results show that a combination of Forward and Backward Contact Tracing was most effective, followed by Backward Contact Tracing. Forward Contact Tracing was the least effective.

Discussion:

From my results, I can conclude that in general, Contact Tracing is an effective way of reducing the number of infections in a pandemic. This is based on the reduced peak infections seen in the SIR graphs, but also in the average number of uninfected individuals remaining at the end of the simulations. The more important takeaway is the effectiveness of the different types of Contact Tracing. Based on my results, a combination of Forward and Backward Contact Tracing is the most effective form of Contact Tracing. But there may be some bias in these results. For one, this method will always quarantine more individuals at once. In Figures 1-3, the number of quarantined nodes is shown in yellow, and in Figure 3 the number of quarantined individuals is shown to be highest during peak infection, and at the end of the simulation. This should be considered when considering how to conduct contact tracing.

When looking at Forward vs Backward Contact Tracing, the results show that, on average, Backwards Contact Tracing is more effective in preventing infections. While on average 4.665% of the simulated population remained uninfected at the end of the SIR simulation using Backwards Contact Tracing, only 3.949% of the population remained uninfected at the end of the Forwards Contact Tracing Simulation. It was shown in Figures 3 and 2 that the Backwards Contact Tracing was quarantining nodes of higher degree. This shows that Backwards Contact Tracing is more likely to quarantine a 'super spreader' or high degree node than Forwards Contact Tracing. This reflects the results of the *Nature Physics (2021)* Backwards Contact Tracing effectiveness study (1).

I believe that in the future, local and state governments should consider implementing a combination of Forwards and Backwards Contact Tracing to prevent infections and 'flatten the curve'. Because they already implement Forward Contact Tracing, that is the least effective form of the sort and they could further reduce infection by attempting to locate and quarantine the source of the infection.

References:

- 1) Kojaku, S., Hébert-Dufresne, L., Mones, E. *et al.* The effectiveness of backward contact tracing in networks. *Nat. Phys.* (2021). <https://doi.org/10.1038/s41567-021-01187-2>
- 2) Rippinger, C., Bicher, M., Urach, C. *et al.* Evaluation of undetected cases during the COVID-19 epidemic in Austria. *BMC Infect Dis* **21**, 70 (2021). <https://doi.org/10.1186/s12879-020-05737-6>
- 3) “Contact Tracing Boulder County.” *Boulder County*, 3 Dec. 2020, <https://www.bouldercounty.org/families/disease/covid-19/contact-tracing/>
- 4) “Contact Tracing Colorado Public Health.” *Colorado COVID-19 Updates*, Colorado Public Health & Environment, 21 Jan. 2021, <https://www.covid19.colorado.gov/contact-tracing>
- 5) Sanche, S., Lin, Y., Xu, C., Romero-Severson, E., Hengartner, N., & Ke, R. (2020). High Contagiousness and Rapid Spread of Severe Acute Respiratory Syndrome Coronavirus 2. *Emerging Infectious Diseases*, 26(7), 1470-1477. <https://doi.org/10.3201/eid2607.200282>

Appendix A:

- 1) <https://github.com/austinritz/Biological-Networks-Final-Project/blob/main/Networks.ipynb>
 - a. The above link is to the code I used to produce my results. It is public and viewable by anyone.