

Modeling Disease Spread Report

Group Members: Abby Carr, Alex Schad, and Sam Zlota

I. Project Motivation and Goals

A shared interest and common topic of this group's work has been Network Science with other interests in data science and epidemiology. Through our work looking into the Network Science research at Northeastern university and the interview with Samuel Scarpino, a joint interest in the effects of networks and the current topic of COVID-19 led us to be interested in the way disease is handled today.

Needing to find a more scalable approach to disease in network science, we decided to focus on the modeling of disease. The main model we looked at was the SIR model; this model was important as it was one of the basic and core models in epidemiology. The SIR model provided our group with a more specialized scope for our broader interest in modeling disease.

Our group's goals were to find some insight into how models were affected by different factors of disease. After work looking into research papers, explanations on the math of the SIR model, and discussing together, the project transformed from work on understanding the SIR model to trying to model COVID-19 ourselves with a version of the SIR model.

Outside of our problem statement and research, we had personal learning goals for this project as well. One of the most important parts of our project was planned to be understanding the model and math we found; our shared goal included understanding how to start our own research. We wanted to get the experience of figuring out where to start and how to continue working with brand new material.

II. Background

When looking at the current work surrounding epidemiology, we found a few different papers of SIR model-related research spanning from simple model explanation[1] to vaccination models[2] to using branching in modeling[3]. Modeling disease is not a small subject and, while our research could only skim the surface, our group feels that the papers and sources discussed provide a solid base for the work we were able to complete.

Looking into some of the more advanced research on modeling disease- and more specifically COVID-19- was the backing for our project. Research into vaccination modeling of influenza[2], although not focused on COVID-19, provided insight into the

Modeling Disease Spread Report

Group Members: Abby Carr, Alex Schad, and Sam Zlota

potential of models. The work showed how more specialized and complex equations could be created to model the spread of disease and the influence of time and various vaccination effects influenced the future work possibilities for our group. Although there are large quantities of COVID-19 research flooding in, this paper[3] on modeling the disease caught our group's attention; we appreciated their discussion of the SIR model and were interested to see a discussion on the state of COVID-19 compared to their modeling.

One particular source caught our group's attention while working on our research topic. Henri Froese's articles found on Medium's Towards Data Science website covered the SIR model[4], work expanding that model[5], and some coronavirus data fitting to a SIR model[6]. Although we didn't plan to meet the same levels as Froese, his work and resources allowed our group to gather the necessary information to create the project.

III. Results

The objective of our project was to determine what parameters best fit real COVID case data. Through the course of our project, we learned that this was a much more niche and complex topic than we had imagined. Assuming we were using valid and reliable sources, we attempted to apply known values to our model. However, we needed to tweak these parameters to best fit real data outcomes.

Another layer of complexity we needed to account for was how many compartments we were going to use in our model. At the simplest level, a given population can be described with as little as two compartments, susceptible and infected, but there can be many more, including, exposed, deceased, recovered, carrier, and maternally derived immunity. In addition, the compartments can be ordered in many ways, describing different disease evolutions.

Although the focus of this model is on the rates of change of the compartments, the realistic picture involves dynamic rates; in other words, a more comprehensive model would take into account the second derivative of the compartments to capture how rates of change change. For example, we were able to capture a logistic rate of change for our R_0 value such that over time it would slowly decrease to reflect lockdown measures. This affected the transmission rate and thus slowed the trajectory of the infected state.

All the more, COVID projections and data are not universal, meaning that any analysis needs to be confined to a certain population usually defined geographically and needs to be confined to a certain time point or time frame. It was very hard to capture truly

Modeling Disease Spread Report

Group Members: Abby Carr, Alex Schad, and Sam Zlota

accurate results. However, we were able to achieve results where the population of the infected and the population of the dead compartment were reflective of real COVID data. The parameters that gave these results were $\alpha=0.02$, $\gamma=\frac{1}{4}$, $\delta = \frac{1}{4}$, and a beta value that ranged from 0.833 to 0.158 over time. We selected an N of 350,000,000 to reflect the United States population. These values produced accurate results for US COVID data.

IV. Conclusions and Future Work

While we were able to understand all of the complexities and intricacies of the SIR model and were then able to model the current state of the Covid-19 pandemic through tweaking parameters in the model, with more time our group could have advanced our project even further. For example, while we did choose our parameters specifically so that the resulting model would reflect the current state of the pandemic in the United States, more accurate parameters could have been calculated if we used data fitting through machine learning. We, unfortunately, didn't have the time or expertise needed to explore this further but it could have improved our results if we had the resources to do so.

As mentioned previously, exploring the rate of change of the rate of change, also known as the 2nd derivative could have provided a more realistic measure for our resulting graph by having inflection points for the lines. This could have in turn reflected more accurate results compared to the current data. There are plenty of other parameters that we could have included that are present in other models such as natural births and deaths. Both of these variables would have been easy to implement if we had more time as they would require only simple addition and subtraction.

Finally, we could have changed the model to include a new state of vaccinated individuals which would represent those that couldn't be infected by Covid. This would have explored the potential effects the current Covid vaccine would have had on the United States' population.

In conclusion, this project has taught us a lot about epidemiology and its relation to network science. Learning about how scientists and researchers model the spread of a disease such as Covid with the SIR model was an exciting opportunity to research a topic we wouldn't have been able to learn in any of our other CS and DS classes. Through this project, we were introduced to the possibilities and the implications of taking part in Computer Science research. By diving deep into this topic of the SIR model, we read

Modeling Disease Spread Report

Group Members: Abby Carr, Alex Schad, and Sam Zlota

scientific literature and learned what working in scientific research would be like. All in all, we had a great time learning about the SIR model and feel that this was a great opportunity to explore research.

V. References:

- [1] Abou-Ismaïl A. (2020). Compartmental Models of the COVID-19 Pandemic for Physicians and Physician-Scientists. SN comprehensive clinical medicine, 1–7. Advance online publication. <https://doi.org/10.1007/s42399-020-00330-z>
- [2] Feng, Z., Towers, S., & Yang, Y. (2011). Modeling the effects of vaccination and treatment on pandemic influenza. *The AAPS journal*, 13(3), 427–437. <https://doi.org/10.1208/s12248-011-9284-7>
- [3] Bertozzi, A. L. (2020, July 21). The challenges of modeling and forecasting the spread of COVID-19. PNAS. <https://www.pnas.org/content/117/29/16732>
- [4] Froese, H. (2020, May 03). Infectious Disease Modelling: Understanding the models that are used to model Coronavirus. Retrieved December 09, 2020, from <https://towardsdatascience.com/infectious-disease-modelling-part-i-understanding-sir-28d60e29fd6c>
- [5] Froese, H. (2020, April 22). Modelling Coronavirus: Beyond the SIR Model | Towards Data Science. Medium. <https://towardsdatascience.com/infectious-disease-modelling-beyond-the-basic-sir-model-216369c584c4>
- [6] Froese, H. (2020, April 22). Infectious Disease Modelling: Fit Your Model to Coronavirus Data. Retrieved December 09, 2020, from <https://towardsdatascience.com/infectious-disease-modelling-fit-your-model-to-coronavirus-data-2568e672dbc7>