

Probability & Stochastic Processes II Final Report
Yet Another Covid Model
Zachary Swarm

Introduction

Since the recent outbreak of COVID-19 across the world, there have been numerous of models attempting to predict how good/bad things will get. One such model that caught my interest was by the Washington Post [1] which demonstrates the effectiveness of social distancing as a way of controlling the spread of disease. This model treats the socially distancing population as being stationary agents spread randomly around a two-dimensional area, vs the non-distancing population which will move randomly. An infected agent will infect others if they collide, and after being infected will become immune to the disease a short time later.

A limitation of the Washington Post model, along with this and other similar models [2] is that the distancing agents are treated as purely being stationary, which is a departure from the reality of occasional trips to locations like grocery stores, which are staffed by the essential workers of the local populace. This leads to an unrepresented concentration of people in specific locations, even from a socially distancing populace. As the sick population becomes symptomatic, their movement behavior changes, which is not realistically represented in these models. Finally, it has been shown that COVID-19 often has an incubation period, during which someone is contagious while being asymptomatic [3]. The model developed for this project is an attempt to address those additional levels of fidelity in studying the effects of social distancing on disease spread.

Model Description

This model, dubbed “Yet Another Covid Model,” is an agent-based modeling approach to represent the effects of population mobility to specific locations on disease spread. Instead of representing movement as unrestricted within a generic space, there is a discrete set of locations that agents can move to at each time step of the model. A single time step represents one hour of real time. Infected agents have a chance of spreading disease to other agents that are located within the same node at every time step. The code uses the Mesa [4] library to support the underlying simulation structure and is publicly available at <https://github.com/Zwami/yacm>.

The location nodes are split into three types, homes, essential services and hospitals. Each agent is randomly assigned a single home node, which may be shared with other agents. Each time step, an agent can move either to a randomly selected hospital or essential service, their home node, or remain at their current location.

The population is divided into three types of agents, doctors, essential workers and the control population. Each type of agent can be configured with a different movement behavior, which

allows the model to handle cases where the control population is spending more or less time at home, vs at services or hospitals, while also allowing doctors and essential workers to move about. The long-term proportion of time any type of agent spends at each type of location node is configurable before running the simulation.

Agent movement is probabilistically determined based on the configured proportion of time spent to be spent at each location type. In effect, this is similar to a Markov chain, but there is an additional simplification as shown in Equation 1 below. The result of this simplification is that the agents effectively have the same state transition probabilities regardless of what state they are in. This simplification was made to reduce the complexity in specifying the simulation parameters, as otherwise the user would be required to either fully define the state transition probabilities for each state and type of agent, or would need to determine appropriate eigenvalues to represent the desired behavior. The Mesa web interface used for simulation configuration does not readily allow for matrix-style user inputs, so it was decided that a simplified transition matrix would be the most user-friendly way to adapt the relevant movement parameters.

$$\begin{aligned}
 [p_{home}, p_{service}, p_{clinic}] &= [\pi_1, \pi_2, \pi_3] \quad p_{home} + p_{service} + p_{clinic} = 1 \\
 &\quad \lambda_1 = 1, \quad \lambda_2 = \lambda_3 = 0 \\
 P &= \begin{bmatrix} 1 & v^{(2)} & v^{(3)} \\ 1 & v^{(2)} & v^{(3)} \\ 1 & v^{(2)} & v^{(3)} \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} \begin{bmatrix} \pi_1 & \pi_2 & \pi_3 \\ \pi_1 & \pi_2 & \pi_3 \\ \pi_1 & \pi_2 & \pi_3 \end{bmatrix} = \begin{bmatrix} \pi_1 & \pi_2 & \pi_3 \\ \pi_1 & \pi_2 & \pi_3 \\ \pi_1 & \pi_2 & \pi_3 \end{bmatrix}
 \end{aligned}$$

Equation 1: Simplified state transition matrix, P , for agent movement resulting from having a single nonzero eigenvalue, with a value of 1.

One special case of movement that is not configurable is the movement of any agent that is infected and symptomatic. These agents have a specific transition matrix that encourages these agents to either remain at home or to move to a hospital. Once at a hospital, the agents will remain there until recovered, at which point they will resume their normal movement. The numbers chosen for this are primarily a result of tuning desirable behavior rather than drawing from literature on home recovery rates vs cases that require hospital visits. This is a point of future improvements that can be made to the model.

Disease progression for an infected agent is also modeled as a Markov Chain, with the states being defined as an augmented version of the Susceptible-Infected-Recovered model by Harko et al [5]. To account for the incubation period of COVID-19 where someone is both infected and asymptomatic, an additional "Asymptomatic" state is added, making the progression chain Susceptible-Asymptomatic-Infected (and symptomatic)-Recovered. The transition probabilities for disease progression of an infected agent are not directly exposed as user configurable parameters, and are set as shown in Table 1. These parameters assume once an agent has recovered from the disease that they are completely immune to reinfection and are no longer contagious to others. For the purposes of this model, deaths are not relevant and are therefore not directly modeled within the disease progression.

	Asymptomatic	Symptomatic	Recovered
Asymptomatic	0.998	0.00175	0.00025
Symptomatic	0.0	0.998	0.002
Recovered	0.0	0.0	1.0

Table 1: Disease progression state transmission matrix. Numbers derived from sources on COVID-19 estimated progression rates [3], [6], [7], [8] and model tuning based on publicly available data from Howard County, MD [9] [10].

Disease transmission, the transition from susceptible to asymptomatic, is primarily a function of the number of other contagious agents within the same node. A contagious agent is defined as an agent that is either in the Asymptomatic or Symptomatic states. Additionally, clinics and essential services can be configured to have infection rate modifiers to represent the effects of personal protective equipment, such as masks, and additional protective measures that are at a higher level of fidelity than this model can support. The transmissibility of the disease itself can also be modified by configuring the R_0 parameter, which represents the average number of other people a single person will infect, along with the Average number of days someone is contagious. This forms a simplified probability of infection per time step for contagious agents. The end result is shown in Equation 2 below.

$$p_{infect\ per\ dt} = \frac{R_0}{(Avg\ days\ contagious * 24)}$$

$$p_{agent\ avoids\ infection} = \left(1 - (p_{infect\ per\ dt} * modifier_{location})\right)^{n_{infected\ agents}}$$

Equation 2: Disease transmission equations

The model also allows configuration of the high-level population and scenario parameters. The total population size, number of doctors and number of essential workers can all be configured. All remaining population that is neither a doctor or essential worker is allocated as the control population. Additionally, the number of households, clinics and essential workplace locations can all be configured. Finally, to seed the start of the scenario, a user can configure the number of initially sick agents in the population. All initially sick agents begin as Asymptomatic, and are randomly dispersed amongst the total population.

The configuration for the model can be setup using the web display, which was created using the underlying Mesa [4] utilities. After configured as desired, the web display will show live results as the model runs. The model will run until it hits a preconfigured time step, or there are no more contagious agents present.

As with all models, there are a number of major assumptions. First being that the agent population is closed, so all cases are related to the sick population at the start of the simulation. This is obviously a departure from the reality of transmission spread. Secondly there is the assumption that recovered agents are no longer contagious or susceptible to future infection, which has not been yet fully verified. The disease spread numbers are based on available data,

but the reality is that many of the papers and resources on this have yet to be fully peer reviewed and validated due to how recent these events are, which may lead to discrepancies. Additionally, there is the built in assumption that the moment an agent becomes symptomatic, they are known to be contagious and will immediately begin socially isolating at home or moving to the hospital. Future study and development of the model could address some of these assumptions and bring the results closer in line with reality.

Example Results

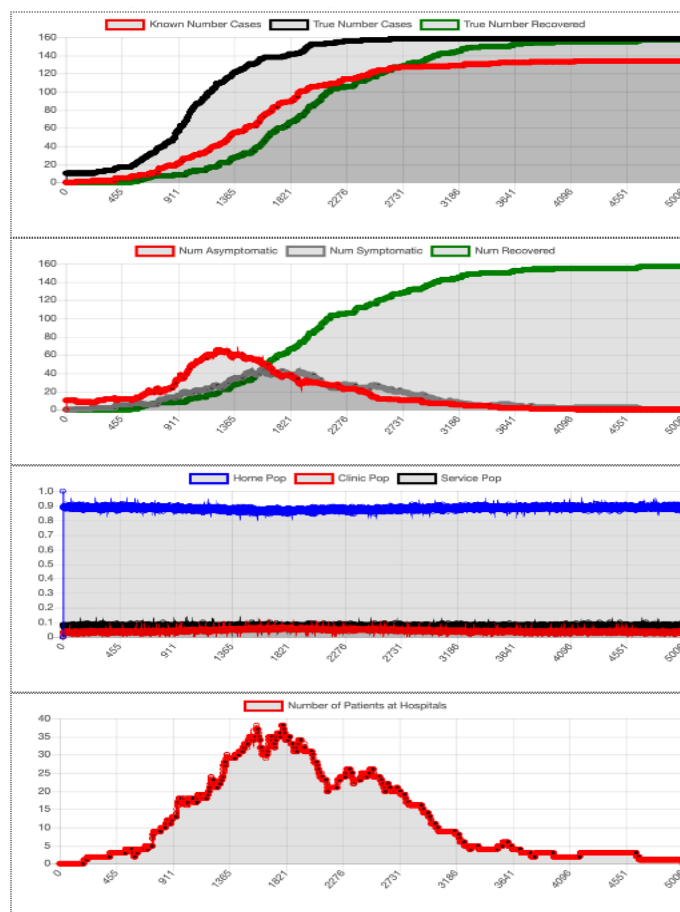


Figure 1: Example results from a YACM simulation of 1000 agents. Control agents spend 98% of their time at home, while essential workers were only home 70% of the time and doctors were home 48% of the time.

Results are shown as a live plot of the locations and disease states of agents over time. The top plot shows the running total of the known and true numbers of cases, along with the number of recovered patients. The second plot shows the instantaneous number of asymptomatic, symptomatic and recovered patients at any point in time. Third plot represents the proportion of agents that are located at the various node types at any point in time. The last plot shows the current number of infected and symptomatic agents located at a hospital node. The horizontal axis for all plots is in units of “time steps” of the model, (one hour in this case). I was unable to find a way to configure the Mesa graphs to customize this or add labels to the charts beyond the legends at the top.

The case shown in Figure 1 demonstrates a run with fairly strict social distancing measures. The result of this is that about 16% of the total population becomes infected, with a peak of slightly fewer than 70 unknown cases active cases at around 54 days into the scenario and just over 40 known active cases at around 72 days into the scenario.

Conclusion & Final Thoughts

The YACM simulation provides an alternative representation of how social distancing behavior impacts disease spread within a population. Agents will congregate at specific nodes, which allows more opportunity for disease spread at those locations. User configuration allows for a variety of scenario configurations based upon movement patterns, types of agents, numbers of relevant locations and factors in disease transmission. These extensible configurations, allows a user a great deal of freedom in modeling various types of cases for both COVID-19 and other types of disease.

All models are wrong, some are useful. YACM provides a certain degree of insight into how effective social distancing can be in the spread of disease, but is not mature enough to be any sort of reliable predictor for a given population. Beyond the current shortcomings and assumptions in the model, for the specific case of COVID-19 there are simply too many physiological and sociological unknowns still to reliably model this behavior.

References

- [1] H. Stevens, "Why outbreaks like coronavirus spread exponentially, and how to 'flatten the curve,'" *The Washington Post*, Mar. 14, 2020.
- [2] 3Blue1Brown, "Simulating an epidemic," *YouTube*, Mar. 27, 2020.
<https://www.youtube.com/watch?v=gxAaO2rsdIs> (accessed Apr. 15, 2020).
- [3] S. A. Lauer *et al.*, "The Incubation Period of Coronavirus Disease 2019 (COVID-19) From Publicly Reported Confirmed Cases: Estimation and Application," *Ann. Intern. Med.*, Mar. 2020, doi: 10.7326/M20-0504.
- [4] P. M. Team, "Mesa: Agent-based modeling in Python 3+," *Mesa*, May 09, 2020.
<https://mesa.readthedocs.io/en/master/> (accessed May 09, 2020).
- [5] T. Harko, F. S. N. Lobo, and M. K. Mak, "Exact analytical solutions of the Susceptible-Infected-Recovered (SIR) epidemic model and of the SIR model with equal death and birth rates," *Appl. Math. Comput.*, vol. 236, pp. 184–194, Jun. 2014, doi: 10.1016/j.amc.2014.03.030.
- [6] L. Gibson and D. Rush, "Novel Coronavirus in Cape Town Informal Settlements: Feasibility of Using Informal Dwelling Outlines to Identify High Risk Areas for COVID-19 Transmission From A Social Distancing Perspective," *JMIR Public Health Surveill.*, vol. 6, no. 2, p. e18844, Apr. 2020, doi: 10.2196/18844.
- [7] W. Guan *et al.*, "Clinical characteristics of 2019 novel coronavirus infection in China," *medRxiv*, p. 2020.02.06.20020974, Feb. 2020, doi: 10.1101/2020.02.06.20020974.

- [8] G. N. Price and E. van Holm, "The Effect of Social Distancing On The Spread of Novel Coronavirus: Estimates From Linked State-Level Infection And American Time Use Survey Data," p. 15, 04/05/20202.
- [9] "DataUSA Columbia, MD," *DataUSA*. <https://datausa.io/profile/geo/columbia-md/> (accessed Apr. 17, 2020).
- [10] "Maryland Department of Health Coronavirus Disease 2019 (COVID-19) Outbreak," *Maryland Department of Health*, May 10, 2020. <https://coronavirus.maryland.gov/> (accessed May 10, 2020).