An Agent-Based Model of COVID-19 Transmission

An Overview-Design Concepts-Details Description

Jonathan M. Gilligan and Kelsea B. Best

2020-05-01

Contents

1 Introduction

This model is a simple agent-based model of COVID-19 transmission in a population connected by social networks. The transmission of COVID-19 is modeled using a compartmental SEIR model in which individual agents progress through disease stages of susceptible, exposed, infected or recovered. Agents are connected to one another through non-spatial social networks which are modeled as Watts-Strogatz small world networks (Watts and Strogatz 1998) or Barabási-Albert scale-invariant networks (Albert and Barabási 2002; Barabási and Albert 1999). The disease may be transmitted through these network links.

2 Model Overview

2.1 Purpose

The purpose of this model is to use agent-based modeling to simulate the spread of COVID-19 in communities. The model is used to simulate the effects of different types of network connectivity (including changes to the network connections through interventions such as "social isolation") on the spread of infections.

The model does not attempt to provide reliable predictions of the future spread of disease, but focuses on making comparisons between different patterns of social network connections.

2.2 Entities, State Variables, and Scales

There are two kinds of entities in this model: **agents** and **links**.

2.2.1 Agents

Agents represent individual people. Each agent is initialized with the following characteristics:

- *age*, which is represented as both a number and as a bracket.
- *sex*, a binary variable representing male or female sex phenotype. This model does not currently treat intersex individuals.
- *comorbidity*, a binary variable that represents whether the indivisual has a comorbidity that could complicate susceptibility to new infections.
- *symptomatic*, a binary variable that controls whether the individaul will be symptomatic if they are infected. This is important, because symptomatic individuals are reported to shed greater amounts of virus than asymptomatic ones. Future variations on this model may also account for behavioral differences (symptomatic individuals may be more likely to self-quarantine or seek medical attention, whereas asymptomatic ones may continue to interact with others and thus spread the disease)
- *health status*, an ordinal factor that represents the status of the individual within a four-compartment *SEIR* model, where *S* represents *susceptible* individuals, *E* represents *exposed* individuals who are incubating an infection but are not yet infectious, *I* represents *infectious* individuals who can spread the disease, and *R* represents *recovered* individuals who are no longer contagious and who have acquired immunity against being reinfected.
- *ticks*, an integer representing the number of time steps since the individual entered the current *health status*. This affects the progression from *E* to *I* and from *I* to *R*.
- *shedding factor*, a real number (positive or negative) that represents the relative degree of viral shedding (e.g., symptomatic individuals are reported to have higher shedding factors than asymptomatic ones).
- *susceptibility factor*, a real number (positive or negative) that represents the relative susceptibility of an individual to contracting an infection when exposed to the virus.
- *EI scale*, *EI shape*, *IR scale*, and *IR shape*, real numbers that parameterize progress of the disease. See further details under *Submodels*

2.2.2 Links

Links represent social connections. These can be *household* connections (people who live together), *social* connections of people who live in different households, but are connected through friendship, church, community groups, etc., and see one another regularly, and *work* connections.

Links have parameters:

- *contact frequency*, represents the mean frequency of this kind of contact, in number of meetings per time step.

- *contact intensity*, represents the intensity of close interaction (e.g., visits to a health-care professional will generally be more intense than interactions with cashiers at a store)

The characteristics of the network is described in greater detail under *collectives*, below.

2.2.3 Spatiotemporal scales

This model does not explicitly represent space, and agents interact through social networks rather than spatial proximity.

The time step represents one day. The number of ticks in a model run can be specified by the user.

2.2.4 Environment

Initially, we will not model the environment. We will have agents connected by social networks with no representation of the physical environments they inhabit. Later versions will incorporate infrastructure to represent connections via transportation, and spatial distributions of housing units and workplaces.

2.3 Process Overview and Scheduling

At each time step, two things happen:

1. Each agent in the *E* or *I* state progresses stochastically toward the next state (*I* or *R*, respectively). The probability of progression may depend on the number of time steps the agent has been in this status. See the *disease progression* submodel for details.
2. Every agent in the *I* state stochastically infects its immediate neighbors on the network (those with whom it shares a link) who are in the *S* status. See the *infect* submodel for details.

3 Design Concepts

3.1 Basic Principles

The transmission of COVID-19 is represented using a four-compartment SEIR model. The SEIR model categorizes agents into *susceptible*, *exposed*, *infectious* or *recovered*.

Infection (transition from *susceptible* to *exposed* status) occurs stochastically. At each time step, every *susceptible* agent that is connected to an *infectious* agent by a network link has a probability of transitioning to an *exposed* state. The probability depends on the characteristics of the two agents and of the network link.

Agents are heterogeneous, so two agents with the same age, sex, etc. can have different shedding and susceptibility factors. These factors are drawn at random from probability distributions that are parameterized by the agent characteristics (age, sex, comorbidities, and symptomatic status)

Exposed and *infectious* agents stochastically transition to the next stage (*infectious* and *recovered*, respectively), with time-dependent probabilities that follow gamma or Weibull distributions. The parameters of these distributions may depend on the agent's age, sex, comorbidities, etc.

3.2 Emergence

The spread of the disease emerges from individual interactions on the network.

3.3 Adaptation

Currently, the agents do not adapt their behavior to changing conditions. Future versions may allow agents to change their social interactions when they get sick or in response to public policies, such as stay-at-home orders.

3.4 Objectives

The agents do not pursue objectives.

3.5 Learning

The agents do not learn.

3.6 Prediction

The agents do not engage in prediction.

3.7 Sensing

The agents do not currently use sensing. In the future, they may sense aspects of their own or other agents' health.

3.8 Interaction

Agents interact through links. These are how infections are transmitted.

3.9 Stochasticity

Agent initialization is stochastic with characteristics (age, sex, comorbidities, future symptomatic response to infection, shedding intensity, and susceptibility) drawn from distributions that can be specified at run-time.

Disease progress (E to I and I to R) are stochastic, with probabilities that vary with the amount of time an agent has been in that status.

Disease transmission is stochastic. Disease is transmitted across links that connect *infectious* to *susceptible* agents. The probability of transmission depends on the *shedding intensity* of the *infectious* agent, the *susceptibility* of the *susceptible* agent, and the contact characteristics of the link.

3.10 Collectives

Multiple overlapping social networks (household, social, and work) connect agents. These networks can have different topologies that are specified at runtime when the agents are initialized.

Currently available topologies are Strogatz-Watts *small-world* (Watts and Strogatz 1998) and Barabási Albert *preferential attachment* (Albert and Barabási 2002; Barabási and Albert 1999). The big difference between these is that the degree of connection is fairly uniformly distributed in the Strogatz Watts model, but is very unequally distributed in Barabási-Albert networks, with a few highly connected nodes that may be able to simulate super-spreaders.

3.11 Observation

At each time step we record the number of agents in each health status (S, E, I, or R).

4 Details

4.1 Initialization

Table of parameters for disease transmission probabilities and disease progression is loaded at run time and used to initialize the agents.

Agents are initialized using user-specified distributions of age, sex, comorbidities, and whether they will become symptomatic if they are infected.

Networks are initialized at runtime using either Strogatz-Watts small-world or Barabási-Albert preferential attachment topologies.

4.2 Input Data

There is no input data during a model run. Everything is generated by the model from the initialization.

4.3 Submodels

4.3.1 Infection

Infection probabilities are a function of shedding intensity, susceptibility, and the frequency and intensity of contact:

$$P_{\text{infection}} = 1 - (1 - p_0)^{\text{contact frequency}},$$

where

$$p_0 = \text{logit}^{-1}\left(x_0 + f_{\text{shedding,source}} + f_{\text{susceptibility,subject}} + f_{\text{contact intensity}}\right),$$

- x_0 is a baseline infection probability parameter, which is then modified up or down by the following factors:
- $f_{\text{shedding,source}}$ represents the intensity of viral shedding by the infectious contact, accounting for things like whether the source is symptomatic or asymptomatic.
- $f_{\text{susceptibility,subject}}$ represents the susceptibility of the susceptible contact, accounting for things like age, sex, and comorbidities.
- $f_{\text{contact intensity}}$ represents the intensity of close personal contact (e.g., a medical visit is likely to entail longer and closer contact than an encounter with a cashier at a store).
- *contact frequency* represents the number of times per time-step (i.e., per day) that people will have contact for this type of link.

4.3.2 Disease Progression

Disease progression from E to I and from I to R is stochastic, with a probability of transition at each time step. The probabilities follow either gamma or Weibull distributions (so the probability of transition is time-dependent, as opposed to a time-independent exponential distribution).

For either distribution, the shape parameters are labeled `ei_shape` and `ir_shape`, respectively and the scale parameters are labeled `ei_scale` and `ir_scale`, respectively.

Parameters for E to I have been estimated from empirical data for both gamma and Weibull representations.

There are no good studies for the progression from I to R, and we may need in the fiture to account for changing viral shedding over the course of the I stage of the disease. Right now, we're using best guesses for this distriubtion, but it will be easy to update it with empirically based numbers.

Right now, disease progression is the same for all demographic groups, but there is the flexibility to have disease progression vary depending on age, sex, comorbidities, etc.

5 References

Albert, Réka, and Albert-László Barabási. 2002. “Statistical Mechanics of Complex Networks.” *Reviews of Modern Physics* 74 (1): 47–97. https://doi.org/10.1103/RevModPhys.74.47.

Barabási, Albert-László, and Réka Albert. 1999. “Emergence of Scaling in Random Networks.” *Science* 286 (5439): 509–12. https://doi.org/10.1126/science.286.5439.509.

Watts, Duncan J., and Steven H. Strogatz. 1998. “Collective Dynamics of ‘Small-World’ Networks.” *Nature* 393 (6684): 440–42. https://doi.org/10.1038/30918.