

Simulation and Analysis of COVID-19 Outbreaks addressing asymptomatic transmission by a modified SEIR model

Muyang Guo, Dayu Zhu, Yibo Wang
{mguo34,dzhu71,ywang3534}@gatech.edu
Georgia Institute of Technology

ABSTRACT

Since the first coronavirus (COVID-19) outbreaks in Wuhan, China, the COVID-19 viral disease has swept into at least 184 countries and killed more than 94,000 people. The US has become a new "center" of the outbreaks as the number of confirmed cases climbed over 450,000 and the number of death reached to 16,000 by 9th April 2020. It is crucial for the public to understand the transmission pattern of the disease and to take proper measures accordingly.

Groups of scientists in China, German claimed that people who are infected but do not have symptoms, or have not yet developed symptoms, can also infect others. This transmission pattern distinct among other epidemic transmission patterns such as SARS or MERS. Traditional models such as SIR or SEIR may not describe the spread correctly. Thus, an extended SEIR model was proposed and analyzed in this project paper to address the asymptomatic transmission during the incubation period. ODE, cellular automata and Markov Chain are the three approaches being used to deploy the conceptual models. The preliminary results are collected by this project checkpoint.

1 DESCRIPTION OF SYSTEM

Different from most infectious disease, the COVID-19 [9] has a very long incubation period, which can be up to 14 days. During the incubation periods, the patients have no serve symptom, or even no symptom at all, but are still able to deliver the virus to susceptibles. Thus, a traditional susceptible-infected-recovered (SIR) model [8] may not strictly describe the spread of COVID-19. Thus, we introduce the susceptible-exposed-infected-recovered (SEIR) [7] model to simulate the circumstances of the pandemic.

We start from the very basic SIR model. In the SIR model, each individual has three distinct states: susceptible, infected, and recovered states. In SIR model, we first need to define a time-dependent variable for S, I and R state, and the variables are the fraction of the states in the overall population. That is, let S_t be the fraction of the population that is susceptible at time t ; I_t be the fraction that is infected at t ; and R_t be the fraction that is recovered at t , where $S_t + I_t + R_t = 1$.

Since COVID-19 has a very long incubation period, which makes it hard to handle. And the no-symptom patients is one of the major channel of the spread of the virus. To model the incubation period, we introduce another state, exposed (E), to the SIR model, and the modified system is the so-called SEIR model.

In SEIR model, the patients who get the virus will first get into the Exposed state. Then, part of the exposed patients will show symptoms and turn to the Infectious state. We use the parameter σ_0 to model the transition. Then, the system can be describe as

$$\frac{d\vec{y}}{dt} = \frac{d}{dt} \begin{pmatrix} S(t) \\ E(t) \\ I(t) \\ R(t) \end{pmatrix} = \begin{pmatrix} -\tau_0 I(t) S(t) \\ \tau_0 I(t) S(t) - \sigma_0 E(t) \\ \sigma_0 E(t) - \frac{1}{\kappa_0} I(t) \\ \frac{1}{\kappa_0} I(t) \end{pmatrix} \equiv \vec{F}(\vec{y}), \quad (1)$$

where $\vec{y}(t)$ is the state vector and both τ_0 , σ_0 and κ_0 are now rate parameters, having units of "fractions per unit time."

2 APPROACHES

2.1 Ordinary Differential Equations (ODE)

solving the ODE is the most straight forward scheme to simulate the system. Here, we use the ODE solver odeint from scipy package to solve the ODE in Eq.(2). The simulation goes across a whole month, 30. At the very beginning, a few individuals in the pool are already infectious, which can be regarded as the seed of the spread, and all the rests are susceptibles. As the spread proceeds, some close contacts of the infectious are infected unfortunately, and the probability of the infection reflects the capability of the spreading. Some infectious will recover and get the immunity and never get infected any more, while some will loss their lives. However, since the dead cases will no longer affect the distribution of S, I and R groups, and since the fatality rate is not too high to change the overall population, we will not take the death into our model.

The continuum model, which is the 'mean field' model, will only care about the overall density of the S, I and R groups without taking a close look at the exact states of individuals. Then, the continuum model will provide a general idea of the spread, so it is a handy method to model a huge population and get the mean behaviour of the spread. On the contrary, since the model is a over-simplified dynamical system, it is not straightforward to describe every details of the real world. Compared with other two methods of our study, the Markov chain and the cellular automaton methods, the dynamical system is not bounded by the size of the pool, which is easily scalable with acceptable computational expenses. For another, this method cannot model each cells in the system, so it is not always precise. Since COVID-19 has a very long incubation period, which makes it hard to handle. And the no-symptom patients is one of the major channel of the spread of the virus. To model the incubation peroid, we introduce another state, exposed (E), to the SIR model, and the modified system is the so-called SEIR model. The preliminary results and some code are displayed in the project update part.

2.2 Cellular Automata - a modified SEIR model

In this simulation, a SEIR model (Susceptible, Exposed, Infectious, Recovered) **fig 1** was selected to describe the spread of COVID-19. However, as the study suggested, it is unknown yet how large in scale can the people during the incubation period (Status: EXPOSED) be infectious yet showing no symptoms. To address this concern, an extra probability layer with a tunable variable was added to describe how soon and how much of people in "E" could become "I", and how people in "E" could affect neighbors become "E" in a similar fashion like people in "I" behave **fig 2**. And an assumption made that the recovered individual will have total immune to the COVID-19. Thus, a modified SEIR model addressing the asymptomatic virus shedding was created.

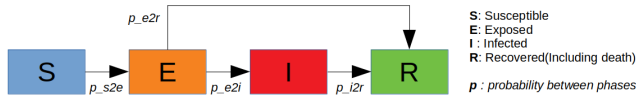


Figure 1: Phase transition pattern of SEIR model

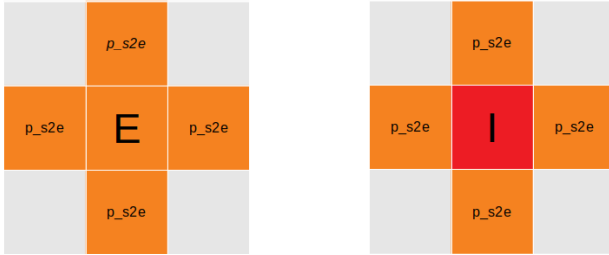


Figure 2: Modified Infectious Pattern for SEIR model

A cellular Automata method was applied, to model the spread, putting one person in each grid cell of a $n \times n$ matrix, the four surrounding neighbor cells representing the persons in close contact. Each person has S,I,E,R four possible states.

2.3 Markov Chain with a modified SEIR model

Markov chain is a stochastic epidemic model for S-I-R and S-E-I-R infectious diseases. Similar to the cellular automata model, Markov chain provides another point of view to analyze the SEIR-CA simulation[2]. Markov chain model can not only simulate the process following the time line, it can also provide the analysis of steady state. COVID-19 progression is assumed to follow a discrete time Markov chain with stationary transition probabilities. The model contains key parameters like finite state space $S = S_1, S_2, \dots, S_N$ and the transition matrix. The Markov Chain model satisfies the rules that the probability of the random variable X being in the state S_j at time t depends on the X 's state at time $t-1$. Due to Markov Chain model has a poor performance in large scale problem and the computation is expensive, an one-dimensional system is built in this project. There are n people in the system and they are distributed in a one dimensional array. Based on the review of COVID-19, a modified SEIR model is implemented. Σ is the state space and $\sigma \in \Sigma$ represents a state. $\sigma = (g_1, g_2, \dots, g_n)$ and

$g_i \in (S, E, I, R)$. S means that the person is susceptible. E represents the person has been exposed but not tested. I means that the person has been infected and tested positive. R means that the person has recovered. In the system, an n array are used to represent people. Each person in the matrix will has their own state, which could be susceptible, exposed, infected or recovered. Then the matrix that contains everyone's state is called a state of Markov Chain.

3 PLATFORM OF DEVELOPMENT

The computaion models were developed in Python 3 enviroment, in the format of Jupyter Notebooks with results visualized and analyzed.

4 LITERATURE REVIEW

4.1 Facts and Data Acquisitions

An analysis of COVID-19 publicly reported confirmed cases in the early stage of the pandemic, from Lauer SA et al,[4], estimated that that COVID-19 has a mean incubation period approximately of 5 days, and 97.5 percent of people will develop symptoms within 11.5 days of infection.

The incubation period is the time between exposure to the disease and starting to show symptoms, usually considered before the contagious period. However, there are cases reported by Rothe, C. et al.[9] showing that asymptomatic persons are potential sources of 2019-nCoV infection, which means the COVID-19 could be contagious even during the late stage of incubation period. Research by Zou, Lirong, et al.[12] on viral load of asymptomatic infected patients also support this, and warrants a reassessment of transmission dynamics of the current outbreak.

The R_0 value is the reproduction number of an epidemic disease. It describes what would happen if an infectious person were to enter a fully susceptible community, and therefore is an estimate based on an idealized scenario. The popularly acknowledged R_0 numebr for COVID-19 is about 2.28 from works by Zhang, Sheng, et al.[10]

All global and US COVID-19 actual data for validation of the conceptual model is collected from the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University (JHU) database archived at its github public repository.[1]

4.2 SIR and SEIR model

The SIR model [9] is the most widely used method to analyze the spread of infectious diseases. The the states, susceptibles, infectious and recovers, provide enough degree of freedom to model the dynamics of a complex system [3]. Besides, the transitions of the states are physically explicable with real-world parameters, making the model interpretable to public health analysis. Since 1976, researchers keeps updating and revising the model to make it closer to their own topics [11]. As for COVID-19, the main difference between it and common infectious disease is the unexpected long incubation period. The incubation period can be as long as 14 days, with a mean of around 5 days. During the incubation period, the patient has no symptom but still infectious, which add difficulty to the spreading control. Thus, we turn to SEIR [5, 7], which add a state, exposed, to model the incubation period. The susceptibles first turn to exposed state, then some of them will unfortunately

become the infectious[6]. Adding a state will provide more room to better fit the outbreak of COVID-19, and provide better ideas of how to cut down the spread.

5 PROGRESS CHECKPOINT

5.1 Current status

5.1.1 ODE. The ODE method is based on the odeint solver, and the system is formulated as:

```
import numpy as np
import matplotlib.pyplot as plt

def F_seir(x, tau, sigma, kappa):
    # x = (s, e, i)
    x_next = x.copy()
    ### BEGIN SOLUTION
    S, E, I = 0, 1, 2
    x_next[S] = x[S] * ( - tau*x[I])
    x_next[E] = x[S] * tau * x[I] - sigma * x[E]
    x_next[I] = sigma * x[E] + x[I] * ( - 1.0/kappa)
    ### END SOLUTION
    return x_next
```

Figure 3: code realization: ODE method

And the current preliminary output of the system is:

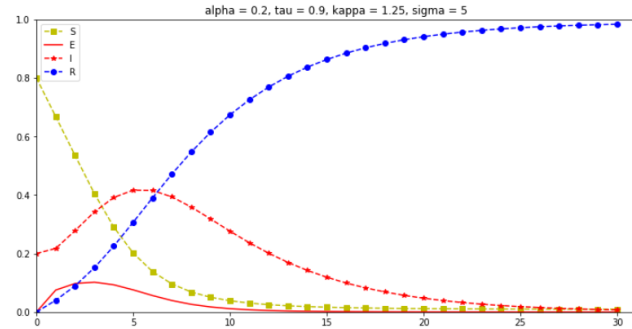


Figure 4: preliminary results by ODE model

Although the result does not strictly reflect the real-world spread of COVID-19, the system runs successfully. We will further tune the parameters of the model or use automatic regression method to get the optimized model. Besides, we will also modify the SEIR model to reflect the truth that some exposed cases may directly recover. The detailed discussions and tutorials can be found from our github.

5.1.2 Cellular Automata. The modified SEIR model addressed the asymptomatic transmission has been fully deployed in the cellular automata frameworks. The model parameters are shown as fig 5. The preliminary results running of a 100×100 a total population of 10000 CA grids shown below in fig 6.

```
### CA framework utilities
EMPTY = -1
SUSCEPTIBLE = 0
EXPOSED = 1
INFECTED = 2
RECOVERED = 3

### Probability Setting
PROB_Exposure = 0.90 #conditional probability of getting sick, given any sick neighbors
PROB_Exposure2Recovered = 0.8
PROB_Exposure2Infectious = 1- PROB_Exposure2Recovered

### Days counter threshold
E_time_thresh = 14 # days
I_time_thresh = 28 # days

### random seed (for debugging purpose)
np.random.seed(1602034230) #
```

Figure 5: Parameters from CA model



Figure 6: Preliminary Result from CA model

We scaled the projection results multiplied to the 832.2 million population of the New York. And we corp-ed the New York real data from the day that has a similar amount of our models day 1 infected cases number, and plotted the comparison shown in fig 7 below, the fatality rate was used a national wide rate of 3.56%. We could easily see the projected confirmed cases from the modified ca model aligned with the actual data closely.

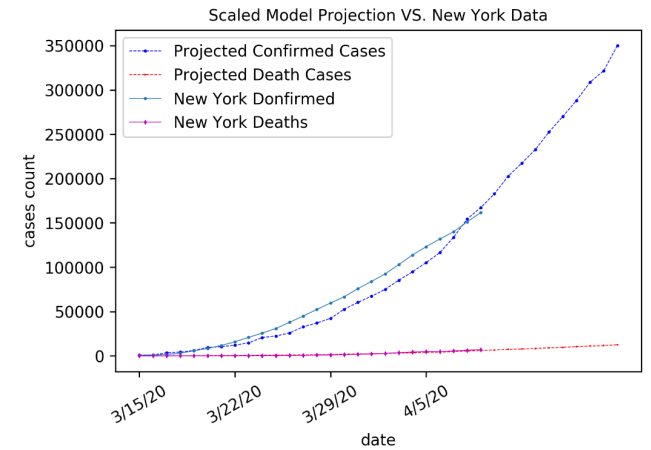


Figure 7: A scaled projection results comparison with the New York real data

5.1.3 Markov Chain. In the very beginning, I tried to implement Markov Chain model on a two-dimensional system. However, after developing a sample code, there will be stack overflow when running the code. Because the time complexity and space complexity are extremely expensive. It can not be run recursively.

The transition matrix is shown below and the steady state are also calculated. As you can see, the final steady state are states with non-infected person.

The parameters like conditional probabilities are predefined. These parameters will be modified based on the real world data in the future.

In this test run, the probability a susceptible is exposed is 0.9 and the probability a infected or exposed recovered is 0.1. To simplify the problem, we choose three people as our population set. The size of population set can be easily changed in the code.

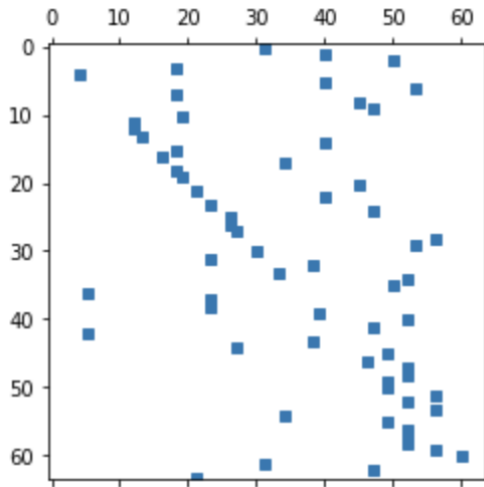


Figure 8: Visualization of Transition Matrix

In the test run, there are 97.7% chance that all three person will get infected and recovered.

```
Initial State: (0, 0, 2)
Steady State:
state: (3, 3, 3) probability: 0.9774143721204103
state: (0, 0, 3) probability: 0.010989010989010986
state: (0, 3, 3) probability: 0.010865851807807258
```

Figure 9: Steady States with Non-zero Probability

We will try to scale the problem for the next step. Two-dimensional system will be studied.

5.2 Changes since proposal and future tasks

Previously in the proposal we tended to model the effect of human inventions, such as social distances, wearing face masks, etc. Since these actions are not directly visualizable by the parameters in our model, we turn to modify the SEIR model to better fit the circumstances of COVID-19. Since the COVID-19 is also infectious

during the incubation period, and the exposed cases can also directly recover without performing any symptoms, so we can modified the equations as:

$$\frac{d\vec{y}}{dt} = \frac{d}{dt} \begin{pmatrix} S(t) \\ E(t) \\ I(t) \\ R(t) \end{pmatrix} = \begin{pmatrix} -\tau_0 I(t) S(t) \\ \tau_0 I(t) S(t) - \sigma_0 E(t) - \gamma_0 E(t) \\ \sigma_0 E(t) - \frac{1}{\kappa_0} I(t) \\ \frac{1}{\kappa_0} I(t) + \gamma_0 E(t) \end{pmatrix} \equiv \vec{F}(\vec{y}), \quad (2)$$

, where $\gamma_0 E(t)$ reflects the transition from exposed states to recovered states. We will add this term into our model in our final reports.

5.3 division of labor

All team members have contributed a similar amount of effort. Muyang Guo is mainly responsible to the Cellular Automata method, Dayu Zhu works on the ODE method, and Yibo Wang studies on Markov chain. All authors contribute equally to the checkpoint report.

6 GITHUB REPO

The preliminary efforts and works are documented in the github repo <https://github.gatech.edu/mguo34/COVID-19-Simulation>.

REFERENCES

- [1] Gardner L. Dong E, Du H. 2019. An interactive web-based dashboard to track COVID-19 in real time. (2019). [https://doi.org/10.1016/S1473-3099\(20\)30120-1](https://doi.org/10.1016/S1473-3099(20)30120-1)
- [2] Jean-Pierre Gabriel, Claude Lefvre, and Philippe Picard. 1990. Stochastic epidemic models for S-I-R infectious diseases: a brief survey of the recent general theory. *Stochastic Processes in Epidemic Theory* (1990), 1–12.
- [3] Herbert W Hethcote. 1976. Qualitative analyses of communicable disease models. *Mathematical Biosciences* 28, 3-4 (1976), 335–356.
- [4] Stephen A. Lauer, Kyra H. Grantz, Qifang Bi, Forrest K. Jones, Qulu Zheng, Hannah R. Meredith, Andrew S. Azman, Nicholas G. Reich, and Justin Lessler. 2020. The Incubation Period of Coronavirus Disease 2019 (COVID-19) From Publicly Reported Confirmed Cases: Estimation and Application. *Annals of Internal Medicine* (03 2020). <https://doi.org/10.7326/M20-0504> arXiv:https://annals.org/acp/content_public/journal/aim/0/aime202005050-m200504.pdf
- [5] Phenyo E Lekone and Bärbel F Finkenstädt. 2006. Statistical inference in a stochastic epidemic SEIR model with control intervention: Ebola as a case study. *Biometrics* 62, 4 (2006), 1170–1177.
- [6] Michael Y Li, John R Graef, Liancheng Wang, and János Karsai. 1999. Global dynamics of a SEIR model with varying total population size. *Mathematical biosciences* 160, 2 (1999), 191–213.
- [7] Michael Y Li and James S Muldowney. 1995. Global stability for the SEIR model in epidemiology. *Mathematical biosciences* 125, 2 (1995), 155–164.
- [8] C Connell McCluskey. 2010. Complete global stability for an SIR epidemic model with delay—distributed or discrete. *Nonlinear Analysis: Real World Applications* 11, 1 (2010), 55–59.
- [9] Camilla Rothe, Mirjam Schunk, Peter Sothmann, Gisela Bretzel, Guenter Froeschl, Claudia Wallrauch, Thorbjörn Zimmer, Verena Thiel, Christian Janke, Wolfgang Guggemos, et al. 2020. Transmission of 2019-nCoV infection from an asymptomatic contact in Germany. *New England Journal of Medicine* 382, 10 (2020), 970–971.
- [10] Sheng Zhang, MengYuan Diao, Wenbo Yu, Lei Pei, Zhaofen Lin, and Dechang Chen. 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 (2020), 201–204.
- [11] Kai Zhu and Lei Ying. 2014. Information source detection in the SIR model: A sample-path-based approach. *IEEE/ACM Transactions on Networking* 24, 1 (2014), 408–421.

- [12] Lirong Zou, Feng Ruan, Mingxing Huang, Lijun Liang, Huitao Huang, Zhongsi Hong, Jianxiang Yu, Min Kang, Yingchao Song, Jinyu Xia, et al. 2020. SARS-CoV-2 viral load in upper respiratory specimens of infected patients. *New England Journal of Medicine* 382, 12 (2020), 1177–1179.