CSE 6730 Project 2 Checkpoint: Simulation of Virus Spread

Yue Yu[1] , **Chenjun Tang**[2] and **Tianqi Liu**[3]
School of Computational Science and Engineering, Georgia Institute of Technology
[1] yyu414@gatech.edu, [2] ctang90@gatech.edu, [3] tliu318@gatech.edu

Introduction

Modeling and characterizing the spread of diseases play a vital role in guarding the safety of citizens and maintaining public security. Moreover, the development of simulation modeling techniques makes it possible to estimate the spread of contagious diseases as well as conduct risk assessments for various control measures [Wang and Wu, 2018]. Recently, the outbreak of the coronavirus disease (COVID-19) poses a great threat to public health, as more than 700,000 cases have been confirmed in more than 100 countries, with more than 30,000 deaths[1]. To tackle the COVID-19 crisis, there are numerous literature studying the origin of COVID-19 [Andersen *et al.*, 2020] as well as the potential treatment strategies [Cortegiani *et al.*, 2020]. Apart from them, understanding the spread patterns of COVID-19 is also of great significance. For individuals, knowing the spread characteristics of COVID-19 enables them protect themselves better and decrease the risk of COVID-19 infection. For governments, they can make better public policies to offer guidance to people and reduce their anxiety[2]. To conclude, it is urgent and necessary to model and understand the spread of COVID-19 disease.

Motivated by above, in our project, we aim to simulate the spread of the disease by using several frameworks including cellular automata model, graph-based dynamic model and ordinary differentiation equation-based (ODE) model. Specifically, with the given information about the disease, we employ a simulation model to analyze the dynamics of disease transmission to evaluate the effectiveness of some factors as well as epidemic control measures. Based on the above simulation models, our ultimate goal to investigate several questions: How did the virus get spread from a location to other locations? To what extent do the policies, for instance, social distancing and community quarantine, help controlling the spread of the virus? Moreover, with the simulation result, we aim to offer a few tips on preventive measures for COVID-19.

System Description

Cellular Automata Model: The cellular automata model focus on the discrete spatial community spread of the virus. Key system components include virus characteristics(e.g. incubation period, infection probability, recovery probability) and the community size.

Graph-based Model: The graph based model investigate how different graph typologies accelerate or decelerate the spreading of the virus. The subject of study is the graph typology. Key system components include graph and graph characteristics (e.g. average degrees), virus characteristics (e.g. Incubation Period, Infection Probability).

ODE-based Model: The ODE-based model investigate how the ratio of different type of population change over time with the spread of the disease. Key system components include the propoertion wrt. different kinds of people and several hyper-parameters that determine the properties of the disease.

[1]https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports

[2]https://www.cdc.gov/coronavirus/2019-ncov/hcp/guidance-prevent-spread.html

Model

I. Cellular Automata Model

In this part, we show the spread of the COVID-19 can be modeled over a geometric region using cellular automata (CA). Cellular automata are algorithms that describe the discrete spatial and temporal evolution of complex systems by applying local deterministic or probabilistic transformation rules to the cells of a regular lattice. In this part, the spread of COVID-19 is modeled in a population distributed geographically. This disease is assume to be a derivation of the susceptible-exposed-infectious-recovered (SEIR) model. People are first covertly infected and they will only show symptoms after an incubation period. It is assumed that a person who has the illness does not die from it and eventually recovers. And a recovered person will never suffer from the illness again.

Conceptual Model

1. World:The world is simulated to be a grid of cells that evolve over discrete time (in days). Every cell of the grid is a person.
2. Cell: Each cell is in one of four possible colors and stands for a person is in the corresponding states:
   (a) White Cell, Susceptible (S): This person has never gotten the illness before. If this person comes in close contact with a sick person, he/she is at risk of catching the illness.
   (b) Yellow Cell, Exposed (E): This person has been infected but is not showing symptoms. This person carries the virus and can infect its neighbor cells. After an incubation period of T days, this person will become Infected.
   (c) Red Cell, Infected (I): This person has the illness and showing symptoms.
   (d) Green Cell, Recovered (R): This person had the illness but has fully recovered. He/she cannot become sick again.

Simulation of the Virus Spreading

The evolution rule of the CA system in each iteration is defined as follows:

1. Spreading: The virus spreads from all carriers to their neighbors at rate τ_1, which is uniform and independent for all positions. Healthy person (Susceptible) infected in this progress becomes a virus carrier (Exposed) but will not show symptoms during the incubation period.
2. Develop: The incubation period of virus carriers decrease by 1. If the incubation period reaches 0, that person starts to show symptoms (Infected).
3. Recover: Exposed and Infected people will recover at rate τ_2 and τ_3, respectively. A recovered person will not be infected any more.

II. Graph-Based Model

In this model we represent the social interactions as an undirected graph and investigate how different society topology, culture characteristics, quarantine and social distancing affect the spread of the

virus. In fact, graph can be seen as a special type of cellular automata where the cells are represented by nodes and its neighbors are defined by the connectivity matrix.

Advanced Graph Terminologies

1. Small-world Network: A small-world network is a type of mathematical graph in which most nodes are not neighbors of one another, but most nodes can be reached from every other node by a small number of hops or steps.
2. Clique: A clique is a set of mutually adjacent vertices (or the complete subgraph induced by that set).
3. Cut Edge: A edge whose removal disconnects the graph. Also called a bridge.

Conceptual Model and Simulation

1. Graph: Graph represents the topology of the social interaction.
2. Node: A node represents a person. **Red Node**: A person who is infected and is showing symptom. When a node become a red note, it will be isolated and all the edges will be removed. **Yellow Node**: A person who is covertly infected. The person carries the virus and can spread it to its neighbor nodes. In T days, A yellow node will become Red Node. **Green Node**: A healthy person. When infected, becomes Yellow Node.
3. Edge: A Edge represent interactions between two person. The weight of the edge represents how frequency two persons interact with each other. The weight is ranged between 0 and 1. It is the external factor that affluence the virus contraction.
4. Clique: A Clique represent a family with most frequent interactions.
5. Cut-Edge A Cut-Edge can represent a non-local interaction, such as travel.

Simulation of the Community Network

It is generally recognized that community graph including social network are subclass of 'small world network', which indicates that any two social entities can be reachable in a small number of steps [Amaral *et al.*, 2000]. Based on this fact and literature review on community detection, we built three types of community model: *Connected Caveman Graph* [Kang and Faloutsos, 2011], *Lancichinetti–Fortunato–Radicchi Benchmark Graph* [Lancichinetti *et al.*, 2008] and *Les Miserables Graph* [Knuth, 2009] based on the culture habits and topology for simulation.

Simulation of the Virus Spreading

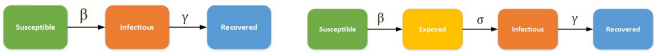
1. Infection Probability: In every iteration, there is a probability that a healthy person can contract the virus thru its infected neighbors. The probability is calculated as

III. ODE-based Model

To model the spreading of the epidemic disease, ordinary differential equation models have been used. There are several kinds of models including SIR, SIS, SEIR and SEIRS, where each particular letter stands for one of the groups of the whole population (Susceptible, Infective, Exposed and Recovered), and the time evolution of these groups is modeled by the set of differential equations with several parameters as external factors. For instance, fig. 1(a) and fig. 1(b) show the flowchart of SIR and SEIR models respectively.

Conceptual Model

Since the original ODE model is the population level model and do not consider geographical information, then the key component is the 'World', which consists of four kinds of people.



(a) The flowchart of SIR model (b) The flowchart of SEIR model

Figure 1: The flowchart of ODE-based simulation model.

(a) Susceptible (S): This person has never gotten the illness before. If this person comes in close contact with a sick person, he or she is at risk of catching the illness.
(b) Exposed (E): This person has been exposed to infected but currently cannot not showing symptoms.
(c) Infected (I): This person has the illness and showing symptoms and can transmit disease to other people.
(d) Recovered (R): This person had the illness but has fully recovered. He or she cannot become sick again.

Simulation of the Virus Spreading

The time evolution of the population compartments in ODE model is described by several nonlinear differential equations. Take the basic SEIR model as an example, it can be written as several equations:

$$\begin{aligned} &\frac{dS}{dt} = \mu(N-S) - \beta\frac{SI}{N},\ \frac{dE}{dt} = \beta\frac{SI}{N} - (\mu+\sigma)E \\ &\frac{dI}{dt} = \sigma E - (\mu+\gamma)I,\ \frac{dR}{dt} = \gamma I - \mu R, \end{aligned} \tag{1}$$

where $N = S + E + I + R$ is the total number of population, β controls how often a susceptible-infected contact results in a new exposure, γ is the rate an infected recovers and moves into the resistant phase, σ is the rate an exposed person becomes infective, μ is the natural mortality rate (unrelated to disease). By solving the equations, the model calculates the number of people in different types, which models the spread of disease in a macro perspetive.

Platform

The project is implemented on the Intel i7 CPU with 8.00GB RAM. All the programs are written in Python. Our simulation platform can be found at https://github.gatech.edu/tliu318/COVID19-Sim.

Related Work

We discuss the work on simulating the spread of epidemiology in the following perspectives. A detailed survey is in [Britton, 2010].

Cellular Automata Model is introduced in [Wolfram, 1983], where each cell is the basic unit and in one state. The new state of each cell is determined by its current state and the states of neighbor cells. Cellular Automata has been applied in cryptography [Tomassini and Perrenoud, 2001] and traffic analysis [Barlovic *et al.*, 1998]. For Disease Spread Simulation, [Pfeifer *et al.*, 2008] adopt Cellular Automata to modeling and justify that geographical barriers may help to slow down the spread of the disease, [Bin *et al.*, 2019] integrate personalized information to evolution rules and achieve better result.

Graph-based Model considers social interactions as an undirected graph where nodes represent individuals and edges represent contacts, which have been widely used for social mobility analysis [Eubank *et al.*, 2004; Cho *et al.*, 2011]. For disease modeling, [Salathé *et al.*, 2010] use contact network data to model the transmission of disease effectively, [Pastor-Satorras *et al.*, 2015] analyze the epidemic spreading in heterogeneous networks theoretically and [Yang *et al.*, 2020] design a dynamic graph model to study COVID-19 transmission and illustrate the efficacy of city lockdown to halt virus spread.

ODE-based Model adopts non-linear dynamical models to simulate the spread of diseases. These models regard population with several compartments: Susceptible (S), Exposed (E), Infectious (I) and Recovered (R) and representative models includes SIS [van den Driessche and Watmough, 2000], SIR [Kermack and McKendrick, 1927] , SEIR [Aron and Schwartz, 1984] and SIRS model [Li *et al.*, 2014]. One drawback of such model is that it neglects external infections due to traveling individuals. Besides, it also does not include variable susceptibility of individuals and complex boundary and initial conditions [Ahmed and Agiza, 1998].

Current Progress

I. Cellular Automata Model

The cellular automata model has been developed on Jupyter Notebook. For now the program does not enable user interaction. The input is the size of community, infection rate and recovery rate. The output includes the timestamp, number of exposed. infected and recovered people. Below is a snippet of the simulation output:

```
Day 61, exposed: 246, infected:171,
recovered:673
Day 62, exposed: 242, infected:178,
recovered:695
Day 63, exposed: 248, infected:176,
recovered:716
Day 64, exposed: 251, infected:178,
recovered:739
Day 65, exposed: 255, infected:169,
recovered:760
```

Also, the spreading of the virus is visualized using OpenCV toolbox and shown in figure 2.

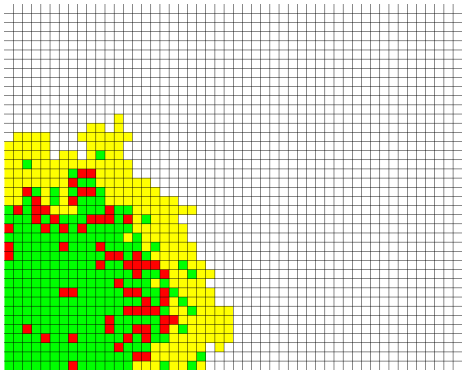


Figure 2: The simulation result of Cellular Automata

II. Graph-based Model

We have a working prototype of the simulation on Google Colab. We are able to simulate the virus spreading on a imported network model. Currently, a healthy person will become covertly infected in the next iteration if he or she is neighbored with a virus carrier. Probabilistic models and self quarantine are not implemented yet, as a result, at the end of the simulation the whole network will become infected. The figure 3 shows the virus spreading on a connected caveman graph, which resembles the social structures in many countries, where a clique of size five resembles a family.

III. ODE-based Model

We've implemented the ODE-based model on Jupyter Notebook with SciPy toolbox. For the simplest settings, we have use SEIR

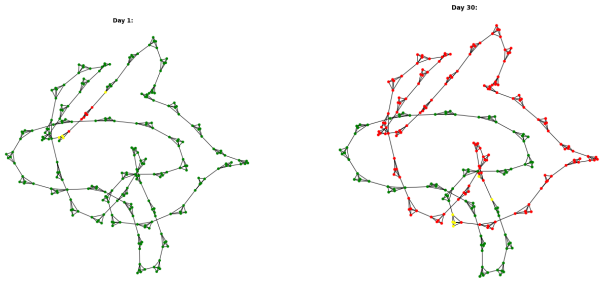


Figure 3: Virus Spreading on a ring-of-cliques network

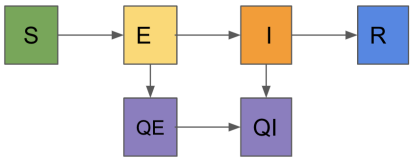


Figure 4: The flowchart of SEIR model with the consideration of quarantine.

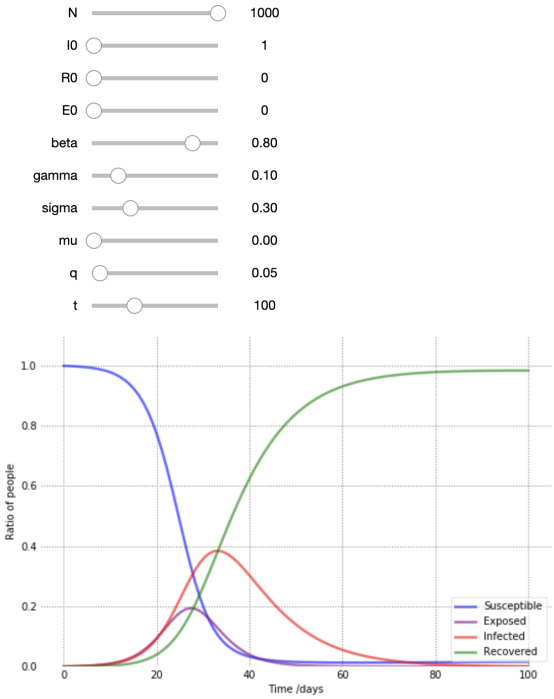


Figure 5: The simulation interface for the basic ODE-based model. By changing the value of different parameters, we can obtain the curve for ratio of different kinds of people.

model to simulate the proportion of different kinds of population (i.e. S, E, I, R) with different parameter settings $\{\beta, \mu, \sigma, \gamma\}$.

Moreover, we also consider the issue of self-quarantine, which have been recommended for preventing COVID-19[3] into the simulation. Specifically, apart from the SEIR modules, we also consider the self-quarantine people, who has a smaller probability to infect others with the disease. Specifically, as shown in Fig. 5, we consider two external variable Q_E, Q_I standing for quarantined exposed and infected people respectively. When people are quarantined, then they will not transimit the disease to others.

Mathematically, compared with Eq. (1), the ODE formulas can be rewritten as

$$\begin{aligned} \frac{dS}{dt} &= \mu(N-S) - \beta\frac{SI}{N}, \ \frac{dE}{dt} = \beta\frac{SI}{N} - (\mu+\sigma+q)E \\ \frac{dI}{dt} &= \sigma E - (\mu+\gamma+q)I, \ \frac{dR}{dt} = \gamma I - \mu R \\ \frac{dQ_E}{dt} &= -\sigma Q_E + qE, \ \frac{dQ_I}{dt} = -\gamma Q_I + \sigma Q_E + qI \end{aligned} \tag{2}$$

where q is the is the proportion of population being self-quarantined.

By using the simulation model in Eq. (2), we would like to justify the effectiveness of quarantine as it can hinder the spreading speed of the disease.

[3] https://www.cdc.gov/quarantine/quarantineisolation.html

Future Work

Currently, the simulation is mainly conducted from three different perspectives separately. However, each model may have its shortcomings. For example, ODE-based models often fail to model the migration of people over different regions, which is a common case in reality. In the future, we plan to take the advantage of these three models altogether and consider both the dynamic information as well as the geographical information for the spread of the disease. Some potential approaches include intergrating dynamics into the social network to evaluate their structural properties relevant to disease propagation [Barthélemy *et al.*, 2005; Bansal *et al.*, 2007] and using cellular automata to capture the individual heterogeneity and improve the ODE-based model [Holko *et al.*, 2016].

Division of Labor

Yue Yu: Developed the ODE-based model, finished introduction, related work and model part of the report.
Chenjun Tang: Developed the Cellular Automata model and helped on report.
Tianqi Liu: Developed the Graph-based model and helped on report.

References

[Ahmed and Agiza, 1998] E Ahmed and HN Agiza. On modeling epidemics including latency, incubation and variable susceptibility. *Physica A: Statistical Mechanics and its Applications*, 253(1-4):347–352, 1998.

[Amaral *et al.*, 2000] Luıs A Nunes Amaral, Antonio Scala, Marc Barthelemy, and H Eugene Stanley. Classes of small-world networks. *Proceedings of the national academy of sciences*, 97(21):11149–11152, 2000.

[Andersen *et al.*, 2020] Kristian G Andersen, Andrew Rambaut, W Ian Lipkin, Edward C Holmes, and Robert F Garry. The proximal origin of sars-cov-2. *Nature Medicine*, pages 1–3, 2020.

[Aron and Schwartz, 1984] Joan L Aron and Ira B Schwartz. Seasonality and period-doubling bifurcations in an epidemic model. *Journal of theoretical biology*, 110(4):665–679, 1984.

[Bansal *et al.*, 2007] Shweta Bansal, Bryan T Grenfell, and Lauren Ancel Meyers. When individual behaviour matters: homogeneous and network models in epidemiology. *Journal of the Royal Society Interface*, 4(16):879–891, 2007.

[Barlovic *et al.*, 1998] Robert Barlovic, Ludger Santen, Andreas Schadschneider, and Michael Schreckenberg. Metastable states in cellular automata for traffic flow. *The European Physical Journal B-Condensed Matter and Complex Systems*, 5(3):793–800, 1998.

[Barthélemy *et al.*, 2005] Marc Barthélemy, Alain Barrat, Romualdo Pastor-Satorras, and Alessandro Vespignani. Dynamical patterns of epidemic outbreaks in complex heterogeneous networks. *Journal of theoretical biology*, 235(2):275–288, 2005.

[Bin *et al.*, 2019] Sheng Bin, Gengxin Sun, and Chih-Cheng Chen. Spread of infectious disease modeling and analysis of different factors on spread of infectious disease based on cellular automata. *International journal of environmental research and public health*, 16(23):4683, 2019.

[Britton, 2010] Tom Britton. Stochastic epidemic models: a survey. *Mathematical biosciences*, 225(1):24–35, 2010.

[Cho *et al.*, 2011] Eunjoon Cho, Seth A Myers, and Jure Leskovec. Friendship and mobility: user movement in location-based social networks. In *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 1082–1090, 2011.

[Cortegiani *et al.*, 2020] Andrea Cortegiani, Giulia Ingoglia, Mariachiara Ippolito, Antonino Giarratano, and Sharon Einav. A systematic review on the efficacy and safety of chloroquine for the treatment of covid-19. *Journal of critical care*, 2020.

[Eubank *et al.*, 2004] Stephen Eubank, Hasan Guclu, VS Anil Kumar, Madhav V Marathe, Aravind Srinivasan, Zoltan Toroczkai, and Nan Wang. Modelling disease outbreaks in realistic urban social networks. *Nature*, 429(6988):180–184, 2004.

[Holko *et al.*, 2016] Arkadiusz Holko, M Mdrek, Zbigniew Pastuszak, and Kongkiti Phusavat. Epidemiological modeling with a population density map-based cellular automata simulation system. *Expert Systems with Applications*, 48:1–8, 2016.

[Kang and Faloutsos, 2011] U Kang and Christos Faloutsos. Beyond'caveman communities': Hubs and spokes for graph compression and mining. In *2011 IEEE 11th International Conference on Data Mining*, pages 300–309. IEEE, 2011.

[Kermack and McKendrick, 1927] William Ogilvy Kermack and Anderson G McKendrick. A contribution to the mathematical theory of epidemics. *Proceedings of the royal society of london. Series A, Containing papers of a mathematical and physical character*, 115(772):700–721, 1927.

[Knuth, 2009] Donald E. Knuth. *The Stanford GraphBase: A Platform for Combinatorial Computing*. Addison-Wesley Professional, 1st edition, 2009.

[Lancichinetti *et al.*, 2008] Andrea Lancichinetti, Santo Fortunato, and Filippo Radicchi. Benchmark graphs for testing community detection algorithms. *Physical review E*, 78(4):046110, 2008.

[Li *et al.*, 2014] Chun-Hsien Li, Chiung-Chiou Tsai, and Suh-Yuh Yang. Analysis of epidemic spreading of an sirs model in complex heterogeneous networks. *Communications in Nonlinear Science and Numerical Simulation*, 19(4):1042–1054, 2014.

[Pastor-Satorras *et al.*, 2015] Romualdo Pastor-Satorras, Claudio Castellano, Piet Van Mieghem, and Alessandro Vespignani. Epidemic processes in complex networks. *Reviews of modern physics*, 87(3):925, 2015.

[Pfeifer *et al.*, 2008] Bernhard Pfeifer, Karl Kugler, Maria M Tejada, Christian Baumgartner, Michael Seger, Melanie Osl, Michael Netzer, Michael Handler, Andreas Dander, Manfred Wurz, et al. A cellular automaton framework for infectious disease spread simulation. *The open medical informatics journal*, 2:70, 2008.

[Salathé *et al.*, 2010] Marcel Salathé, Maria Kazandjieva, Jung Woo Lee, Philip Levis, Marcus W Feldman, and James H Jones. A high-resolution human contact network for infectious disease transmission. *Proceedings of the National Academy of Sciences*, 107(51):22020–22025, 2010.

[Tomassini and Perrenoud, 2001] Marco Tomassini and Mathieu Perrenoud. Cryptography with cellular automata. *Applied Soft Computing*, 1(2):151–160, 2001.

[van den Driessche and Watmough, 2000] Pauline van den Driessche and James Watmough. A simple sis epidemic model with a backward bifurcation. *Journal of Mathematical Biology*, 40(6):525–540, 2000.

[Wang and Wu, 2018] Lin Wang and Joseph T Wu. Characterizing the dynamics underlying global spread of epidemics. *Nature communications*, 9(1):1–11, 2018.

[Wolfram, 1983] Stephen Wolfram. Statistical mechanics of cellular automata. *Reviews of modern physics*, 55(3):601, 1983.

[Yang *et al.*, 2020] Xiaofei Yang, Tun Xu, Peng Jia, Han Xia, Li Guo, Lei Zhang, and Kai Ye. Transportation, germs, culture: A dynamic graph model of covid-19 outbreak. *Available at SSRN 3544816*, 2020.