**Network Analysis and Modeling**

**CSCI 5352, Fall 2020**

**Prof. Dan Larremore**

**Project Proposal**

**Qinglu Sun & Tianli Wu & Mengchen Gong**

**Introduction**

The outbreak of COVID-19 that started in Wuhan quickly swept the world in just three months. It is highly contagious, spreads quickly, has a long incubation period, is asymptomatic and remains contagious. Because of the lack of timely control and the high flow of people, the virus quickly spread across China, and then a blowout outbreak began globally. From the winter of 2019 to the spring of 2020, scientists analyzed and confirmed that the virus was a droplet infection. At the end of the summer of 2020, various countries analyzed the new subtypes of the new coronavirus and found that its infectivity increased while its lethality decreased. We hope to use network models to describe the spread of infectious diseases in the population, perceive if the model fits the actual situation/or previous infection cases, thereby predicting the number of patients in the future. Even though there are lots of parameters to consider such as temperature, humidity etc.

**Related work**

Parnell, D.[3] provided a rapid analysis of COVID-19 on mass sports gathering, he treated sports as a fundamentally network game, and showed sports gathering as a mirror to our network society and globalized operation. A few studies have been conducted on analyzing the spreading of COVID-19 using the SIR model since there are things in common between them. Chen et al.[1] proposed a time-dependent SIR model which tracks the probability of one susceptible becoming infected and the probability of one infected becoming susceptible to anticipate COVID-19 tendency. There are other studies focused on characterizing the spread of COVID-19, especially analyzing the impact on different genders, different ages groups and analyzing the average, mode and median incubation periods among the people based on the data they got [2]. Fanelli et al.[5] studied mainly on the susceptible-infected-recovered-deaths (SIRD) model to see the rate of infection, recovery and death in Italy, China and France and found that the recovery rate is about the same in different countries even though the infection rates are different because of the culture factors. Additionally, the study found out the death rate is correlated to average age and health conditions the patients are in, which gave us an insight of the potential factors that we need to take into account before starting the simulations.

**Proposed Work**

Objectives for this project are formulate individual virus infection probability calculation, use random walk for individual movement model and simulate isolation, self-isolate nodes near the source of infection, transmission route of super infected person, community transmission simulation by using epidemiological models. There are several different ways to formulate epidemiological models, such as SIR, SEIR, and etc. We can also make model assumptions based on data such as the incubation period ≈ 7 days, or set the mean incubation period from 2-14 days, according to WHO literature report and research conducted by Wang, P.[2], the onset period can be set to 10 days, and we can make the following naive assumptions when computing the model that the patient in incubation period don’t have the ability to transmit COVID-19 to others the patient in onset period may have the ability to transmit the virus to other with certain infection probability, the patient no longer has the ability to transmit COVID-19 after onset period, and assuming there was only one human infected person at the starting period of COVID-19. We set the infection probability initially dependent on daily confirmed cases, and use weight to represent the distance between individuals where distance can be physical distance, or comprehensive distance between individuals,we can adjust the parameter when simulating the model. By using random walk, we can first initialize individual position and place one individual of source of infection the individual movement step is used to describe the individual and formulate matrix and changing values in for loop within for loop, update the position of the individual with time step until time period we want to analyze is reached. Due to the randomness of parameters, we may need to adjust parameters to make the model fits to the actual model (actual situation). We might simulate two different scenarios: the initial stage of the virus development, where there is one human infected with COVID-19 but there wasn’t anyone notice, where everyone including the infected human is active normally, second scenarios, policy of quarantine was made, any infected human or contact with the infected human is under quarantine.

**Dataset**

Possible dataset to use: <https://www.kaggle.com/kimjihoo/coronavirusdataset>. The datasets has several csv, case.csv described the geological information, patient info have the sex age, infected by, number of contact and date of sympotom which is the dataset we are most interested in. The datasets also include the policy information, weather and etc.

<https://github.com/skissler/haslemere> This dataset consists of the number of pairwise distances between individuals and time step accordingly.

**Works Cited**

[1] Y. -C. Chen, P. -E. Lu, C. -S. Chang and T. -H. Liu, "A Time-dependent SIR model for COVID-19 with Undetectable Infected Persons," in IEEE Transactions on Network Science and Engineering, doi: 10.1109/TNSE.2020.3024723. Retrieved October 05, 2020, from <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=9200529>

[2] Wang, P., Lu, J. A., Jin, Y., Zhu, M., Wang, L., & Chen, S. (2020). Statistical and network analysis of 1212 COVID-19 patients in Henan, China. International Journal of Infectious Diseases. Retrieved October 06, 2020, from <https://reader.elsevier.com/reader/sd/pii/S1201971220302678?token=EBF672F7C4EAA5E46EF47EC5784BFBB9ED1AF3B8CD70A539C66D3BBC85518021FDA092FA497250DF8D4B4189E6C5B27C>

[3] Parnell, D., Widdop, P., Bond, A., & Wilson, R. (2020). COVID-19, networks and sport. Managing Sport and Leisure, 1-7. Retrieved October 09, 2020, from <https://www.tandfonline.com/doi/full/10.1080/23750472.2020.1750100?casa_token=6nP7uRnnM_kAAAAA%3Abc1323hddDAdbMUnOQZVCd0TT6ex9vmydGTQqvY-C7aFHINVIVrGfFHoop_Mmbza9YCKdkzCSA8n#aHR0cHM6Ly93d3cudGFuZGZvbmxpbmUuY29tL2RvaS9wZGYvMTAuMTA4MC8yMzc1MDQ3Mi4yMDIwLjE3NTAxMDA/bmVlZEFjY2Vzcz10cnVlQEBAMA==>

[4] World Health Organisation. (2020a).Who, Coronavirus disease (COVID-19) outbreak.(2020). Retrieved October 07, 2020, from <https://www.who.int/emergencies/diseases/novel-coronavirus-2019>

[5] Fanelli, D., Piazza, F., (2020). Analysis and forecast of COVID-19 spreading in China, Italy and France. Retrieved October 08, 2020, from <https://doi.org/10.1016/j.chaos.2020.109761>

[6] Firth, J., Hellewell, J., Klepac, P., Kissler, S., Kucharski, A., & Spurgin, L. (2020, August 07). Using a real-world network to model localized COVID-19 control strategies. Retrieved October 09, 2020, from <https://www.nature.com/articles/s41591-020-1036-8>