<u>Final Project Goals</u>

Team members

Halim Park (halimp2) / JeIn You (jeinyou2) / Wan Yoo (wanyoo2)

We will be using OpenFlights data and use Breadth-First Search (BFS) as our traversal. We will use Dijkstra’s Algorithm to determine the shortest path from one point to another. We would like to simulate the spread of COVID-19 within a given time(days) and graph the output using a Force-directed graph drawing algorithm, and the following are the assumptions to reduce the complexity.

Since there are too many airports, we are going to limit our scope of simulation to airports in China only. Since China only has one uniform time-zone, time-zone difference will not be taken into account.

Input / Output / Default settings

1. Input : Time (days)
2. Output :
    - # of people infected after T days (T = input days)
    - Graphic Output
        - Visual representation of # of infected people in each country by color
        - Size of the nodes determined by the # of routes that an airport has
3. Default settings
    - Default starting airport: Beijing
    - Default starting # of infected people: 1
    - Default infection rate: 2 people infected per 1 infected person.

A list of assumptions we made:

- airports and flights are open and available at all times.
- We assume that by default, one person in Beijing is infected (at time = 0day)
- We will randomly assign destinations for each infected person.
- In the case where there are no direct flights, an infected person will make a transfer. .
    - We will also assume that no one is being infected during the flight, and the disease is spread only after arrival.
- We are planning to represent each airport as a node and edges are connected according to the following rule:
    - the nodes (airports) will be loaded into the graph in the order of the dataset given. When the new edge (to be connected) creates an overlap over the existing one, the new edge will not be created.
- Each infected person must be assigned a destination.
- When an infected person reaches his/her destination, that person will not be considered.
- A plane must fly and land at the destination node (airport) or transfer node(shortest path) of its original airport.
    - During this process, there is always a spread of disease, with the probability that is determined by the description below.

Probability of the spread of disease

- While the time zone in China is universal, we are going to take the moment of time to allocate different infection rates at different times of the day.
    - As a default, we are going to infect 2 people per 1 infected person. And we are going to set different probabilities of the infection rate depending on the time period. If the infected person arrives between 11am~7pm , we will set the probability higher than the average, and at other times, we will take that lower than the average. This takes into account the crowdedness of the airports.
    - The process will include random sampling of the population in order to render causality in our experiment.
    - We will take into account the normal behavior of COVID-19, including the patient having no symptoms but still spreads the disease.

Design of resulting Graph

**standards may change as we investigate more

- The airports will be displayed as dots varying in sizes and colors depending on
    - Size: # of routes
        - Largest: # of routes > 500
        - Medium: 100 < # of routes <= 500
        - Small: # of routes <= 100
    - Colors: # of infected people
        - Red: # of infected people > 100,000
        - Yellow: 50,000 < # of infected people <= 100,000
        - Green: # of infected people <= 50,000