Texas COVID-19 Spread and Vaccine

Christopher Stromer

7/26/2020

1. **Introduction**

The COVID-19 outbreak has caused many US states to shut down most economic activity which is severely hurting our economy. Politicians and leaders in office are under pressure to open the country back up again or face a severe lasting economic crisis and potential long-term hardship. The goal of this project is to track the spread of coronavirus through Texas starting from a specific city and also to find the shortest path from Corpus Christi, where a potential vaccine from another country is being imported, to a city with the greatest need for a vaccine.

1. **Data**

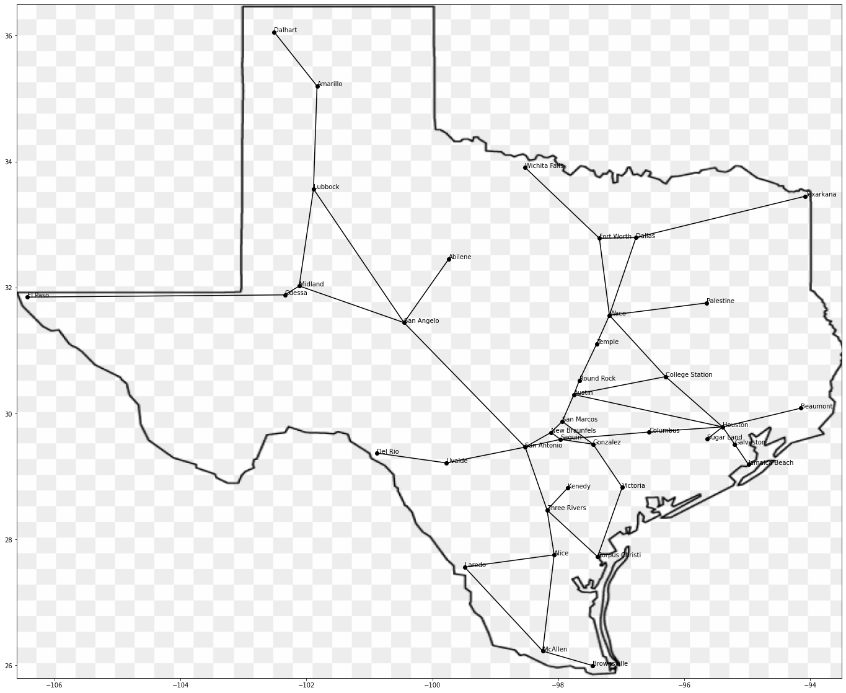
The main data we will be using is coming from The University of Texas at San Antonio Computer Science department. This data contains city and distance information for all of the cities in Texas. I will also be using some location data from Foursquare.

1. **Methodology**

I began by loading all of the data into the notebook. Once I had the data loaded, I labelled the columns so they were easier to understand and reference. After the data was ready to be used, I graphed the data and overlaid it on an outline of a Texas map, so we are able to get the big picture of how the cities are connected. I wrote a function to perform an uninformed search on the data. This function is compatible with Depth First Search and Breadth First Search. In our case we started with Three Rivers, but this method can be used with any other city listed in the dataset. This function returns the entire path of the virus spread and the total distance travelled in order for the virus to reach ever city. In my project, I ran the function using both DFS and BFS. After this function, I created 2 functions to help figure out an informed search strategy on the data so I could determine the best path for the vaccine to take from Corpus Christi to Amarillo. As with the uninformed search function, this one can also use different city combinations. This function returns the path followed and the distance it took just like the uninformed search function.

1. **Results**

Our first image, the map outline of Texas, is provided so you can see all of the cities listed in our dataset and see how all of them are connected so you can understand the path results we received from our search functions.

****

For our uninformed search function, we first tried it starting from Three Rivers and we used Depth First Search. This gave us our total distance calculation to spread from Three Rivers to every other city in Texas in the shortest path using Depth First Search. However, we ran the same function again using Breadth First Search and ended up with a shorter distance and a different path.

Depth First Search Output:

Virus Spread Path:

0, Three Rivers

1, Alice

2, McAllen

3, Brownsville

4, Laredo

5, Corpus Christi

6, Victoria

7, Gonzalez

8, San Marcos

9, Austin

10, Houston

11, Sugar Land

12, Columbus

13, Galveston

14, Jamaica Beach

15, Beaumont

16, College Station

17, Waco

18, Temple

19, Dallas

20, Texarkana

21, Fort Worth

22, Wichita Falls

23, Palestine

24, Round Rock

25, New Braunfels

26, Seguin

27, San Antonio

28, San Angelo

29, Abilene

30, Lubbock

31, Amarillo

32, Dalhart

33, Midland

34, Odessa

35, El Paso

36, Uvalde

37, Del Rio

38, Kenedy

Total Distance Travelled = 3454.6133905600004

Breath First Search Output:

Virus Spread Path:

0, Three Rivers

1, Kenedy

2, San Antonio

3, Corpus Christi

4, Alice

5, New Braunfels

6, Seguin

7, Uvalde

8, San Angelo

9, Victoria

10, Laredo

11, McAllen

12, San Marcos

13, Gonzalez

14, Columbus

15, Del Rio

16, Midland

17, Lubbock

18, Abilene

19, Brownsville

20, Austin

21, Houston

22, Odessa

23, Amarillo

24, Round Rock

25, College Station

26, Beaumont

27, Galveston

28, Sugar Land

29, El Paso

30, Dalhart

31, Temple

32, Waco

33, Jamaica Beach

34, Palestine

35, Fort Worth

36, Dallas

37, Wichita Falls

38, Texarkana

Total Distance Travelled = 3354.3199229800002

For our informed search function, we started at Corpus Christi and ended in Amarillo. The goal of this was to determine the shortest path from Corpus Christi to Amarillo. We started by performing a Greedy Best First Search on the data and came up with a total distance of 674.63. Then, we performed an A\* search on the same data and came up with the same total distance. The only difference between the two search algorithms is that the A\* checked every node to make sure it was taking the best path, while the Greedy Best First Search only chose nodes with the best immediately available path.

Greedy Best First Search Output:

Nodes Visited

Corpus Christi

Three Rivers

San Antonio

San Angelo

Lubbock

Amarillo

Vaccine Distribution Path:

0, Corpus Christi

1, Three Rivers

2, San Antonio

3, San Angelo

4, Lubbock

5, Amarillo

Total Distance Travelled = 674.63197221

A\* Search Output:

Nodes Visited

Corpus Christi

Three Rivers

San Antonio

Victoria

Kenedy

Gonzalez

New Braunfels

San Marcos

Seguin

New Braunfels

Austin

San Marcos

Seguin

San Angelo

New Braunfels

Round Rock

New Braunfels

San Marcos

Austin

Seguin

New Braunfels

San Marcos

San Antonio

Round Rock

San Antonio

Three Rivers

Austin

New Braunfels

San Marcos

Seguin

San Antonio

San Antonio

Lubbock

New Braunfels

Austin

San Marcos

Uvalde

Seguin

Round Rock

San Antonio

New Braunfels

Austin

Austin

New Braunfels

Temple

Seguin

San Marcos

Amarillo

Vaccine Distribution Path:

0, Corpus Christi

1, Three Rivers

2, San Antonio

3, San Angelo

4, Lubbock

5, Amarillo

Total Distance Travelled = 674.63197221

1. **Discussion**

For our uninformed search, we can see different paths and different distances calculated as a result of those different paths. The reason for this is because Breadth First Search visits each of its child nodes before exploring deeper, whereas Depth First Search is the opposite. This causes our algorithm to pick a different order of cities to visit even though the data is the same. For our current data set, we have determined that the best fitting algorithm for uninformed searching is Breadth First Search. For our informed search, we can see that the “Nodes Visited” section of the output is very different, while the total distance travelled remained the same. This is due to A\* search being extremely thorough. It searches through every possible node so it can check every possible path to make sure you are getting the lowest possible total distance calculation. It just so happens that our data was set up in such a way that the Greedy Best First Search was able to produce the correct results in a shorter and quicker time.

1. **Conclusion**

Covid-19 is a rapidly spreading virus that the world was not fully prepared to handle, but we are doing the best we can. It spreads rapidly, so containing it is very difficult. In my project, I gave a few demonstrations on how the coronavirus could spread throughout Texas, and how we could potentially distribute a vaccine throughout Texas, prioritizing the most affected cities first. Seeing as the virus has already spread well throughout the country, we need to focus on the vaccine to prevent more people from falling ill and try to lessen the pressure on our healthcare system. In my project, I demonstrated how some search algorithms could be useful to determine the best path for distributing a vaccine for Covid-19.