SOCIAL AND ECONOMIC NETWORK ANALYSIS

PROJECT REPORT

ROAD NETWORK ANALYSIS

TEAM - 13

19Z315 – Gowtham S

19Z318 – Harish J

19Z347 – Shashaank R

19Z356 – T Naveen Kartik

19Z361 – Vikram Krishna A

PROBLEM STATEMENT:

The problem statement is to analyse a road transport network. In this project we perform the analysis using the links (roads) and nodes (places) in the network. The dataset is loaded, visualized and network metrics are found. The analysis that are performed on the network is finding the shortest path between source and the destination, finding the optimal path where the average speed and the traffic is taken into consideration and a variation of the same analysis where time is taken into consideration.

DATASET DESCRIPTION:

The dataset consists of 4 features and length of the dataset is 170 records. The features in the dataset are Source, Destination, Distance and Traffic. Distance is given in the form of weights, in the dataset is given from 1 to 20. Traffic is also given in the form of weights from 1 to 50. All these features are used for different analysis. For shortest path source, destination and distance is used. For optimal path using traffic, speed and time is also used. Also perform link analysis on how the formation of a new link will affect the neighbouring node's traffic and the travel time.

TOOLS USED:

Gephi: Gephi is the visualization and exploration software for all kinds of graphs and networks. We have used Gephi to visualize the dataset in the form of a graph

NetworkX: NetworkX is a Python package for the creation, manipulation, and study of the structure, dynamics, and functions of complex networks. With respect to our project we have used NetworkX for finding the shortest path. We have used functions like all_shortest_path(), neighbors(), non_neighbors(), add_edge().

Pandas, numpy and matplotlib: Pandas is a python library and it is a fast, powerful, flexible and easy to use open source data analysis and manipulation tool. In our project we have used pandas to load and pre-processing of dataset. Numpy is used for dataset generation, we used the random function to produce the weights for distance and traffic. Matplotlib is used for plotting the shortest distance graph and highlighting the shortest distance.

CHALLENGES FACED:

- The initial dataset taken was too lengthy (around 26,000 records). Visualizing this large dataset using Gephi was not possible so the dataset was trimmed down and modified.
- When choosing the possible nodes for link formation. Deriving at a formula for finding the estimate values of the traffic, distance and time was very hard. A lot of time, trial and error methods had to be performed to find a normalized formula.
- Finding an accurate measure for dynamic traffic redistribution after the formation of new link was hard. We have taken 5% as the static traffic redistribution value. But it can be either 2 or 3 or even 10 from different links.

CONTRIBUTION OF TEAM MEMBERS:

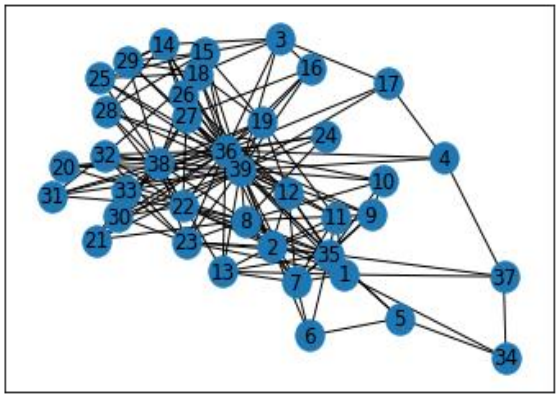
TEAM MEMBER	CONTRIBUTION
19Z315 – Gowtham S	Time and distance graph analysis,Report
19Z318 – Harish J	Dataset collection, pre-processing and shortest route analysis
19Z347 – Shashaank R	New link formation Analysis
19Z356 – T Naveen Kartik	New link formation Analysis
19Z361 – Vikram Krishna A	Visualization,Distance graph analysis and Report

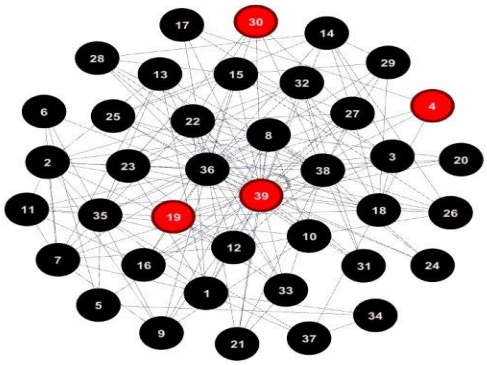
ANNEXURE I: CODE

https://github.com/19z347/SENA_Project---Road_Network_Analysis/blob/main/Road_Network_Anlaysis.ipynb

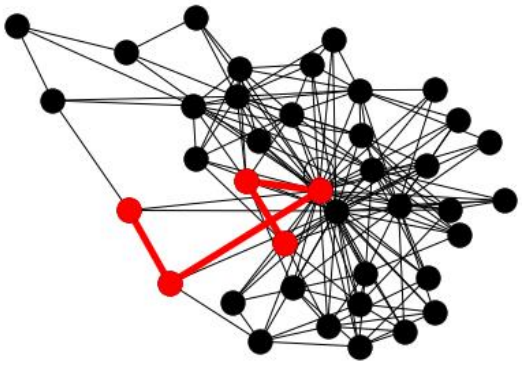
ANNEXURE II: SNAPSHOTS OF THE OUTPUT

1. **Visualization using networkx and Gephi (Distance graph):**





2. **Shortest path using networkx (from node 4 to 19)**
   [4, 17, 39, 12, 19]



3. **Time measure and distance measure in distance graph given above. For the path 4-19.Here the shortest path will be 4-39-12-19, since cost is taken as the weight**

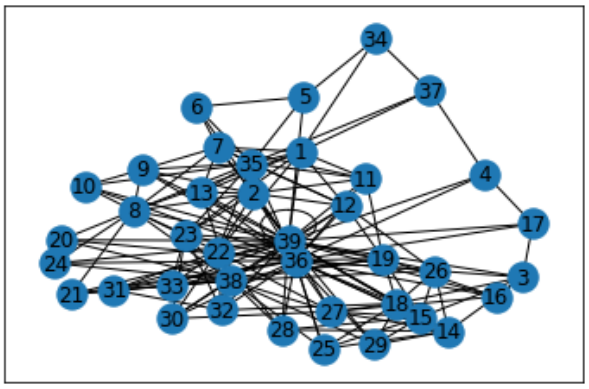
Time measure:

```
Time to travel from 4 to 39 is 27.20
Time to travel from 39 to 12 is 6.38
Time to travel from 12 to 19 is 10.20
Total time to travel from 4 to 19 is 43.77
```

Distance measure:

```
Distance to travel from 4 to 39 is 16.00
Distance to travel from 39 to 12 is 2.00
Distance to travel from 12 to 19 is 6.00
Total distance to travel from 4 to 19 is 24.00
```

4. **Time Graph:**



5. **Time measure and distance measure in time graph given above. For the path 4-19. Here the shortest path will be 4-39-12-19, since cost is taken as the weight**

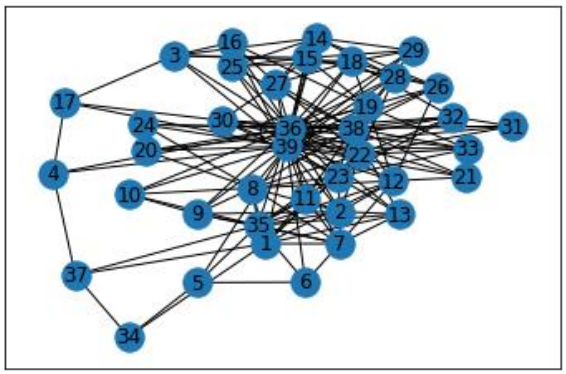
Time measure:

```
Time to travel from 4 to 17 is 34.00
Time to travel from 17 to 39 is 9.37
Time to travel from 39 to 12 is 6.38
Time to travel from 12 to 19 is 10.20
Total time to travel from 4 to 19 is 59.94
```

Distance measure:

```
Distance to travel from 4 to 17 is 4.00
Distance to travel from 17 to 39 is 6.00
Distance to travel from 39 to 12 is 2.00
Distance to travel from 12 to 19 is 6.00
Total distance to travel from 4 to 19 is 18.00
```

6. **Traffic graph:**



7. **Link analysis in the traffic graph:**

In this analysis we analyse how will a creation of a link between two nodes will affect the traffic in the existing links. The nodes are chosen on the basis on the links which have the maximum traffic. The process of choosing the nodes involves calculating the traffic between a node and all its non-neighbouring nodes and the highest traffic is taken. In our case the two selected nodes are 19 and 39. The traffic between 19 and 39 will be reduced due to formation of the node between 19 and 31.

Time measure between link 19 and 31 through 36 (before direct link formation):

```
Time to travel from 19 to 36 is 40.50
Time to travel from 36 to 31 is 153.00
Total time to travel from 19 to 31 is 193.50
```

Time measure between link 19 and 31 through 36 (after direct link formation between 31 and 19):

```
Time to travel from 19 to 36 is 39.34
Time to travel from 36 to 31 is 107.10
Total time to travel from 19 to 31 is 146.44
```

REFERENCES:

[1] https://networkx.org/documentation/stable/tutorial.html

[2] https://snap.stanford.edu/data/roadNet-CA.html

[3]https://www.sci.unich.it/~francesc/teaching/network/geodesic.html#:~:text=A%20shortest%20path%2C%20or%20geodesic,geodesic%20distance%20or%20shortest%20distance

[4] https://www.sciencedirect.com/topics/social-sciences/network-analysis

[5] https://stackoverflow.com/questions/24024411/highlighting-the-shortest-path-in-a-networkx-graph

[6]https://networkx.org/documentation/stable/reference/algorithms/shortest_paths.html

[7] Working with networkx

[8] https://www.analyticsvidhya.com/blog/2020/01/link-prediction-how-to-predict-your-future-connections-on-facebook/

[9] https://towardsdatascience.com/road-transportation-optimization-with-python-part-1-visualisation-costing-698eadcdce0b

[10]https://www.youtube.com/results?search_query=Lect+06+Network+Analysis+with+Networkx+Working+with+Data+sets