

SNA-Project

Bust Route Analysis using social networking techniques. This readme consist of how to run various code files and description of major functions. Out Project uses python3 for data processing and Gephi and Google Earth for visualization. Requirements: For Python install the packages in req.txt. These can be installed also using pip install

```
pip install -r req.txt
```

For Gephi we have used "Event Graph Layout" plugin.

Data collection

The [AUTODOWNLOAD.ipynb](#) file contains step by step instruction how we downloaded the data from various websites. This file generates a csv file, [Basicmodtable.csv](#) containing route information and geographical coordinates. We use this file to extract L1 space.

Format of the output file

The output csv file must have the following coulmn heading necessary for further processing. We have called the output table as [Basicmodtable.csv](#).It must have "RouteNumber","Source","Destination" (all without quotes)columns required for further processing by other script files.

Formation of different Spaces

L-Space

The output of [AUTODOWNLOAD.ipynb](#) script [BasicTablemod.csv](#) generates a CSV file which is our L Space. We will derive other spaces based on this space.

P-space

Use the script [PSpace.py](#) to generate the Pspace. Set the variable loc to the location of the CSV file in the above mentioned format.

```
loc = 'BasicTablemod.csv' #Replace this with your filename in the above format
```

This script will store the P-space graph in a networkx variable called PGraph. We run networkx's Flyod Warshall Algorithm to get the minimum hops(route changes required) to reach from one bus station to another. The minimum hops are stored in "fwnp" variable which is a numpy matrix where each row and coumn correspond to a bus station. The script finally uses fwnp matrix to plot the histogram of number of hops.

BiPartite Graph(B-Space)

The [CSVtoGEXF.py](#) This takes input from CSV and edits the standard GEXF to add labels to nodes so that they

can be plotted by "Event Layout" plugin in the gephi. The output of this file is a siple gexf file and the the label are read as string. So one needs to duplicate the "label" column in the "Data Explorer" Tab > "Nodes" to integer values.

C-space

The [CspaceTablemaker.py](#) script takes our [BasicTablemod.csv](#) and process it to fowm various C-space tables and save them in .csv format. Our results can be found in the [Output/Cspace](#) folder.

Misc Scripts

[viz.py](#) A quick script to visualize the the graph on network x