STEVENS
INSTITUTE of TECHNOLOGY
THE INNOVATION UNIVERSITY

Logistics Optimization of Freight Management and Preferences using Social Network Analysis

BIA 658 WS

Submitted By ,
Ruchi Hiralal Mendhegiri

Index

Problem Overview and Motivation

- Ripple effect of the pandemic (covid-19), the global supply chain industry is facing many facing difficulties in smooth flow of goods in multiple sectors including consumer products, industrial products, electronics, automotive, life sciences and many more.
- There is a need of redesigning supply chain strategies to be more collaborative, resilient and agile with customers, suppliers and other stake holders.
- In this project I would like to work on logistics management/ freight management in USA.
- For this project, the scope of project is limited to the eight different domestic modes of transportation and 51 shipping hubs located in different USA states.
- Using the historical data and tools of Social Network Analysis, the project aims to address few of these issues that are laid out below:

Most Influential Shipping Hub

Maximum Connectivity

Shortest/optimal Path

Model Layout

Goal 1: Most influential Hubs

- **Nodes**: 51 Shipping Hubs
- **Edges**: Number of direct transportation modes/options they have running between them.
- So weights on the edges will indicate number of modes.

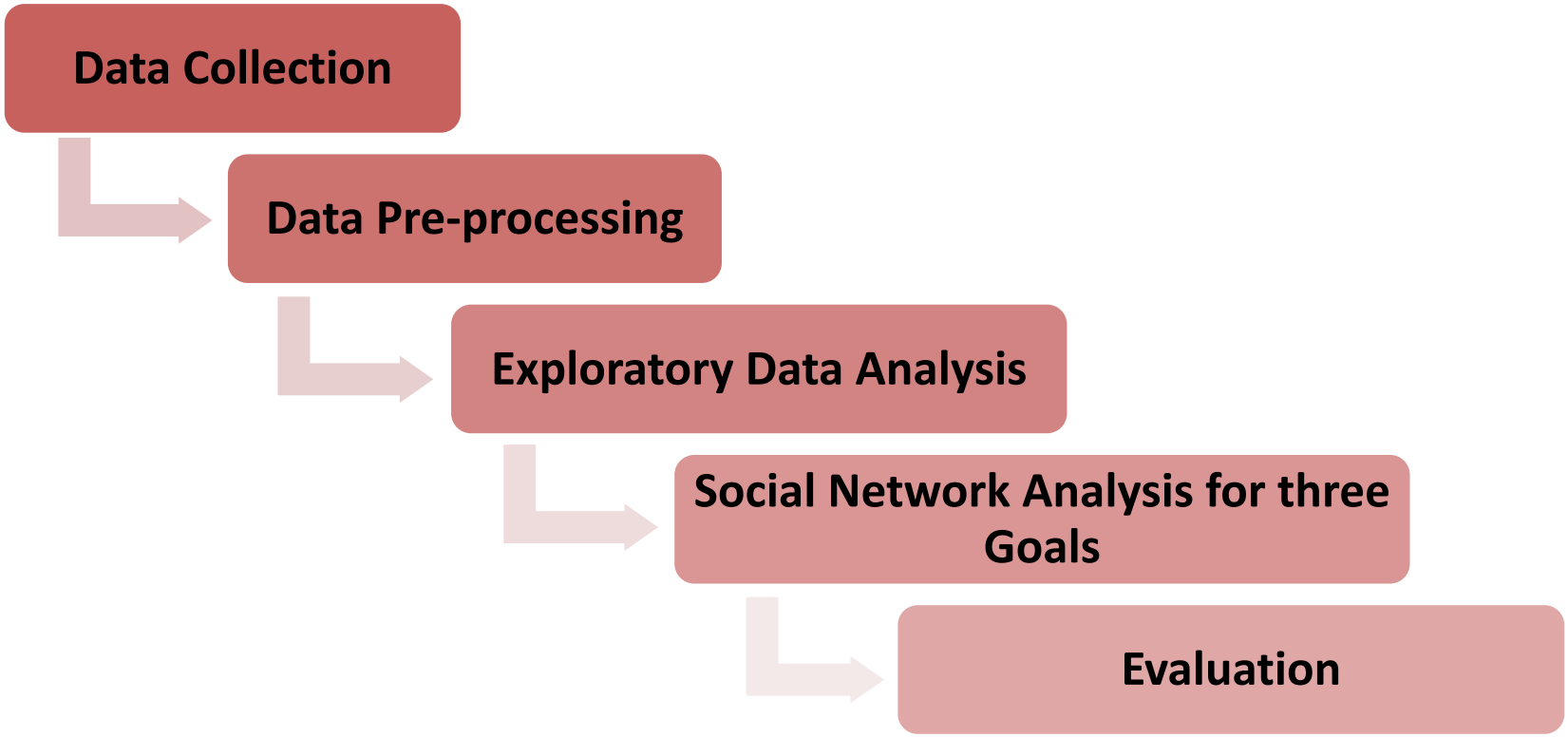
Goal 2: Central shipping hub with maximum road connectivity

- **Nodes**: 51 Shipping Hubs
- **Edges**: Whether they have a direct edge (transportation option) running between them. (0 – 1 indicator).
- For this analysis the focus will be only on mode of transportation as truck and commodity as electronics and other electric equipment.

Goal 3 : Optimal shipping path between origin and destination

- **Nodes**: 27 Shipping Hubs
- **Edges**: Average distance between each pair of origin and destination shipping hub
- For this analysis the focus will be only on mode of transportation as pipeline and commodity as crude petroleum.

Methodology



- For all the 3 focused objectives, the nodes csv file remains same. However, the edges csv file differs as per the problem statement
- I will be using excel tool and Gephi for this project

Data Overview and Attributes

- I will be using one main dataset which will give Freight analysis information and secondary freight databases if necessary.
- Dataset has been downloaded from FAF (Freight analysis framework version 5) website and using data tabulation tool, I have custom select the required data.

Original Dataset

Column Name	Description of Column	Number of Entries	Missing Values	Data type
dms_org	Domestic state origin where freight movement begins: It can take values between 1 to 56	550321	5	Integer
dms_dest	Domestic region destination : It can take values between 1 to 56	550321	5	Integer
dms_mode	Domestic mode : There are total 8 different types of transportation modes available for shipping at domestic level. (Values 1 to 8)	550321	0	Integer
sctg	Commodity : It can take any value from 1 to 99 which represents different types of commodities	550321	0	Integer
value_year	Freight Value in $million	550321	0	Integer
weight_year	Freight Weight in thousand tons	550321	0	Integer

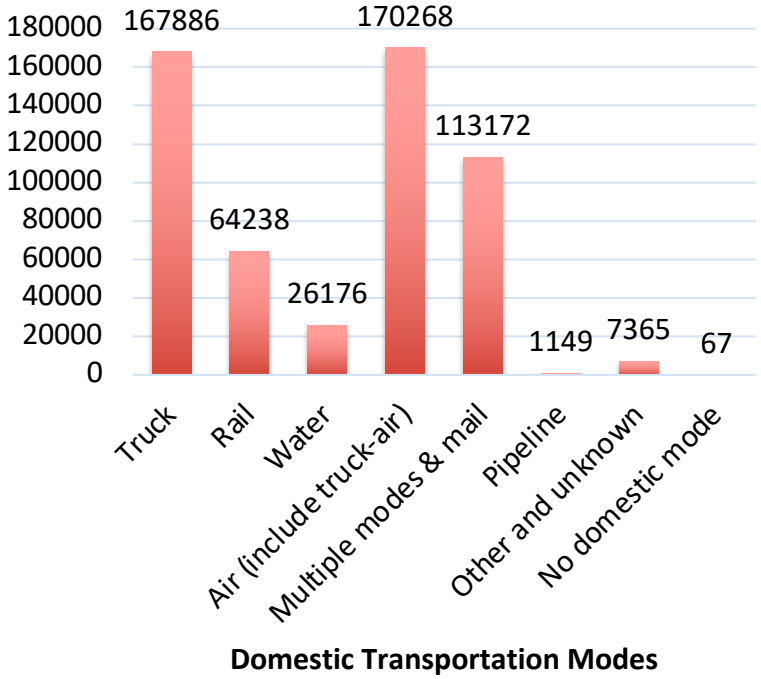
Dataset Link:

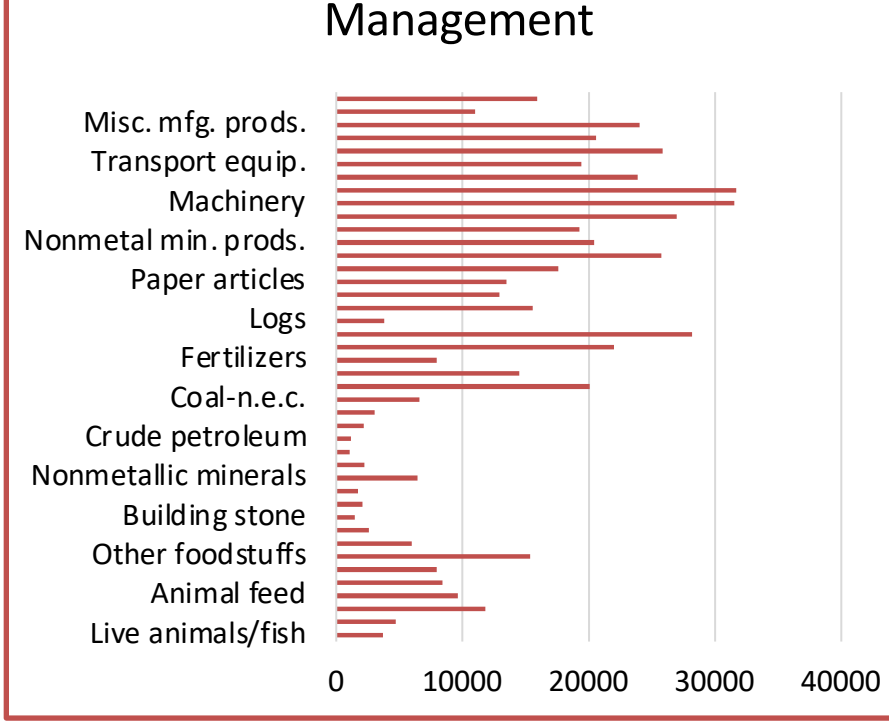
1. https://faf.ornl.gov/faf5/dtt_domestic.aspx
2. https://www.kaggle.com/usdot/freight-analysis-framework

Data Cleaning and Observations

- The columns with missing values have been dropped to eliminate the faulty data.
- There are no duplicate values present in the dataset.
- For initial network analysis, I have filtered dataset in terms of mode of transportation and type of commodity.
  - ➢Number of Nodes : 51
  - ➢Commodity (sctg values)
- After filtering, the new dataset consist of 380053 records of Freight analysis information.
- In these experiments the focus will be on directed graphs.
- The nodes files consist of Shipping Hub ID and its Name.
- The edges file differs as per the problem statement.

Exploratory Data Analysis

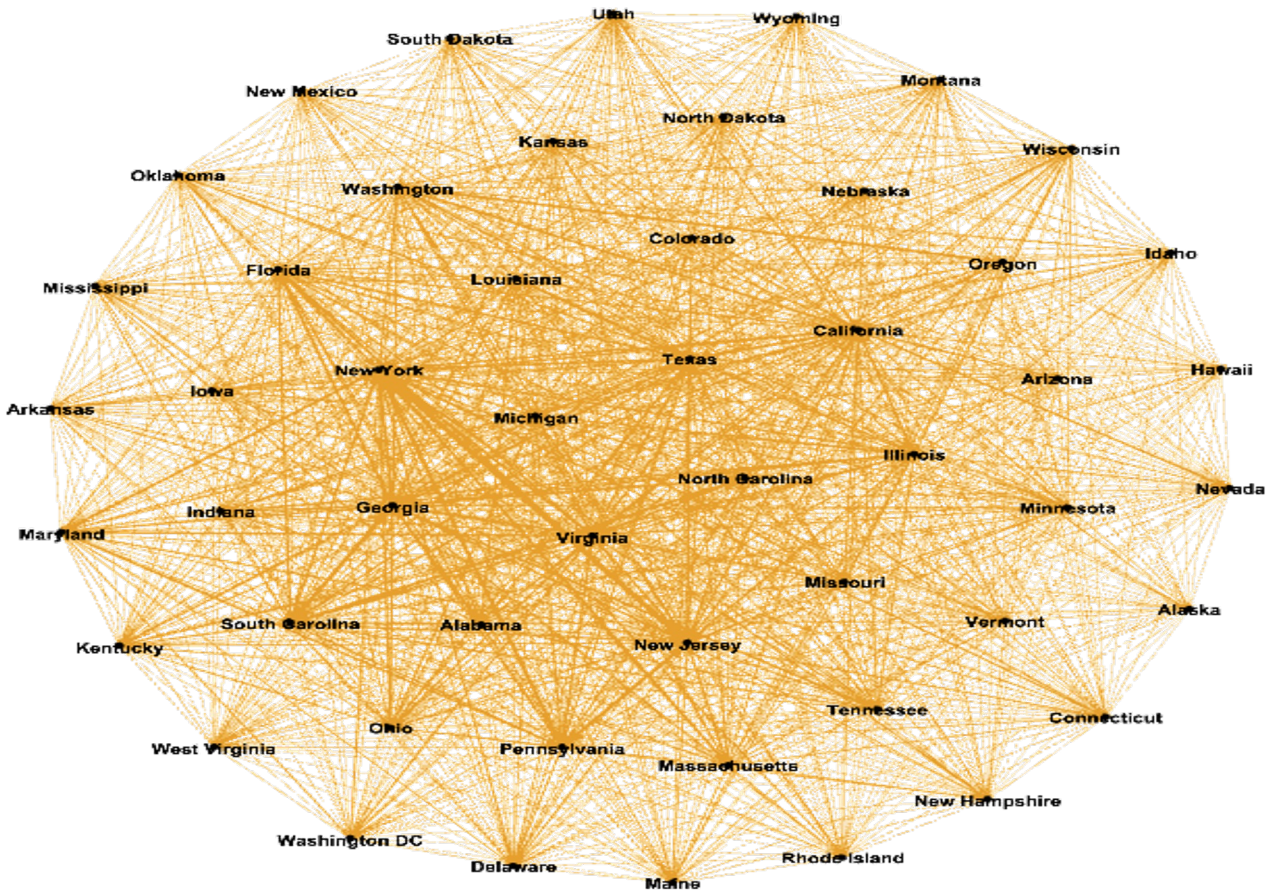




Out of 8 domestic modes of Freight management, Air and Truck Air is most popular which captures more than 85% of total Freight

Most Freight transportation carried out of Basic chemicals, pharmaceuticals, plastic/rubber, machinery and Electronics

Overall Social Network Graph : Gephi



This diagram represents 51 domestic shipping ports with all possible modes of transportation and all available combinations of inbound and outbound traffic.

Social Network Analysis for 3 Goals & Evaluation

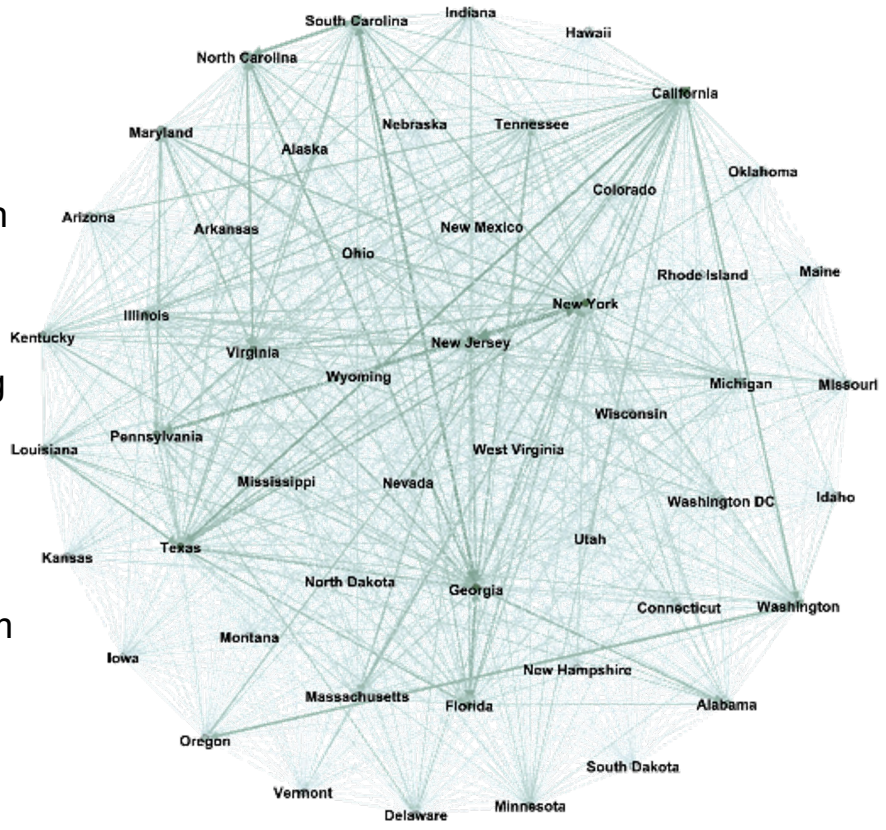
Goal 1 : Most influential Hubs

- The focus of this experiment is to track down, total number of direct transportation options hubs have running between them.
- For first objective, we will create separate edge file with all possible modes of transportations to all connected shipping hubs options (Use of Excel COUNTIFS function).
- Hence, weights on the edges will indicate transportation flow options (1 represents direct connection between shipping hubs).
- The figure shows the network flow diagram with ranking as per weighted degree.

Weighted Degree:

- The weighted degree of a node is like the degree but pondered by the weight of each edge. It's doing the sum of the weight of the edges.
- Since shipping port with high influence have great connectivity, they can be considered to be centre of operations.
- We will consider both outflow and inflow options for this experiment.

Gephi Social Network Diagram:



Top 5:

Shipping Hub	Weighted In-Degree
New York	33214
California	28295
Georgia	28020
Texas	27384
New Jersey	23127

Shipping Hub	Weighted Out-Degree
California	27383
Georgia	26763
Texas	25539
New Jersey	25308
New York	23546

Shipping Hub	Weighted Degree
New York	56760
California	55678
Georgia	54783
Texas	52923
New Jersey	48435

- Comparing all above statistics, It can be concluded that even though there are highest number of options available to reach New- York shipping hub, more outgoing options available in case of California.
- Considering the geographical distribution of these shipping hubs and weighted degree, it will be beneficial to create 3 operations hubs:
  - New York for East Coast Operations
  - California For West Coast Operations
  - Texas for Mid – West Operations
- **This distribution of freight management into three parts will ensure smooth flow of goods with minimum lead time.**

Goal 2: Central Shipping Hub With Maximum Road Connectivity

- For first objective, we will create separate edge file with Mode of transportation as truck (dms_mode = 1)
- The shipping port with maximum connectivity can be determined by marking the shipping port with high degrees or high eigenvalue centrality

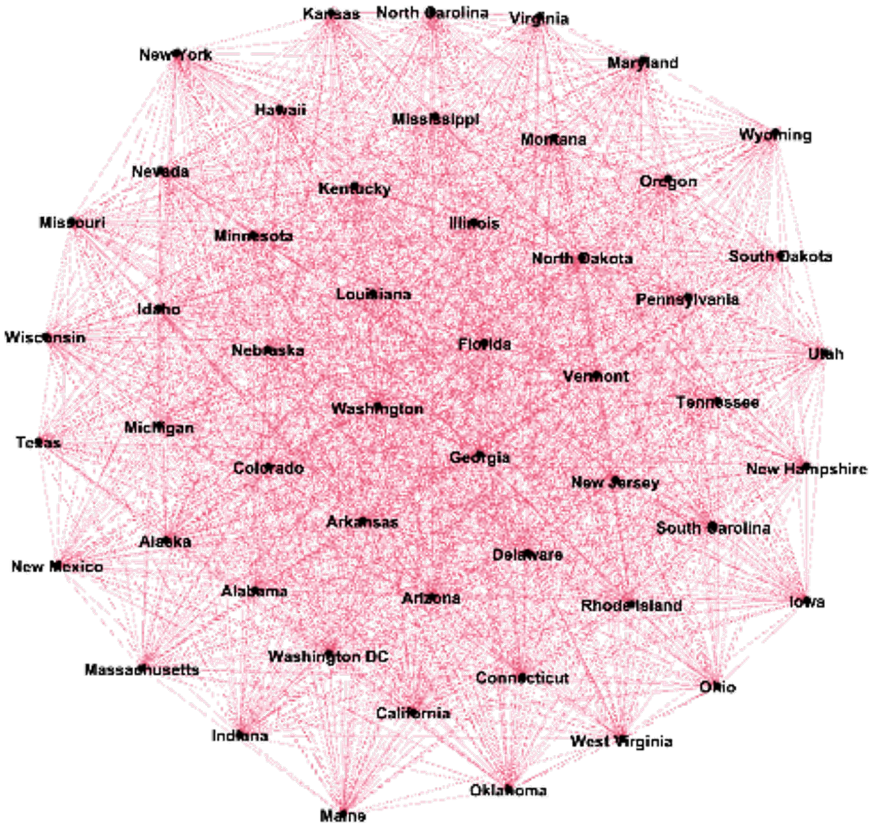
Degree of a node:

- The degree of a node/vertex is the number of edges connected to the node. The higher the degree, the more connected the node.
- In this context, the shipping port with the highest degrees would be more connected, providing supply to maximum nearby ports

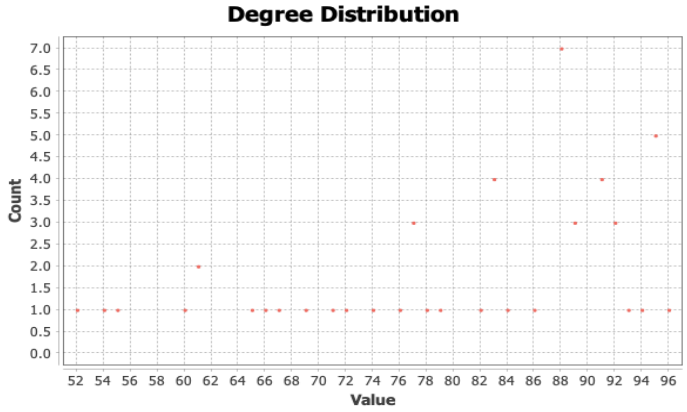
Eigenvalue Centrality:

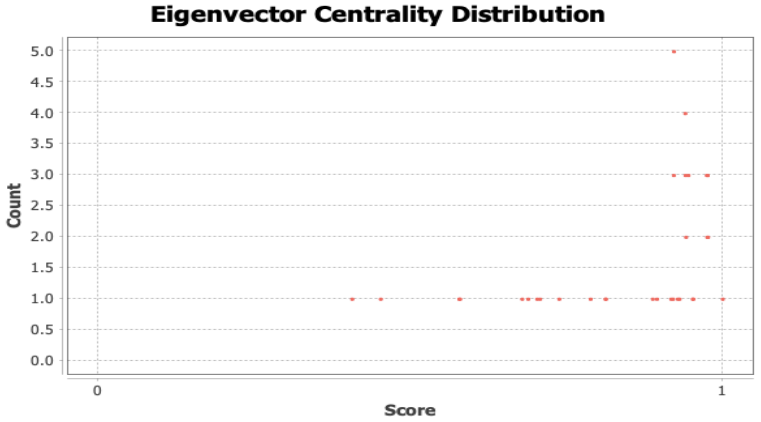
- Eigenvector centrality is a measure of the influence of a node in a network. A high eigenvector score means that a node is connected to many nodes who themselves have high scores.
- Since shipping port with high influence have great connectivity, they can be considered to selected as central hub for road transport for smooth flow of electronics goods

Gephi Social Network Diagram:



Average Degree: 40.745





Top 5:

Shipping Hub	Degree
Colorado	96
Washington DC	95
Delaware	95
California	95
Arkansas	95

Shipping Hub	Eigenvector Centrality
Colorado	1
Washington DC	0.9765
Delaware	0.9765
California	0.976
Arkansas	0.976

It can be observed that Colorado has maximum road connectivity among all 51 shipping hubs.

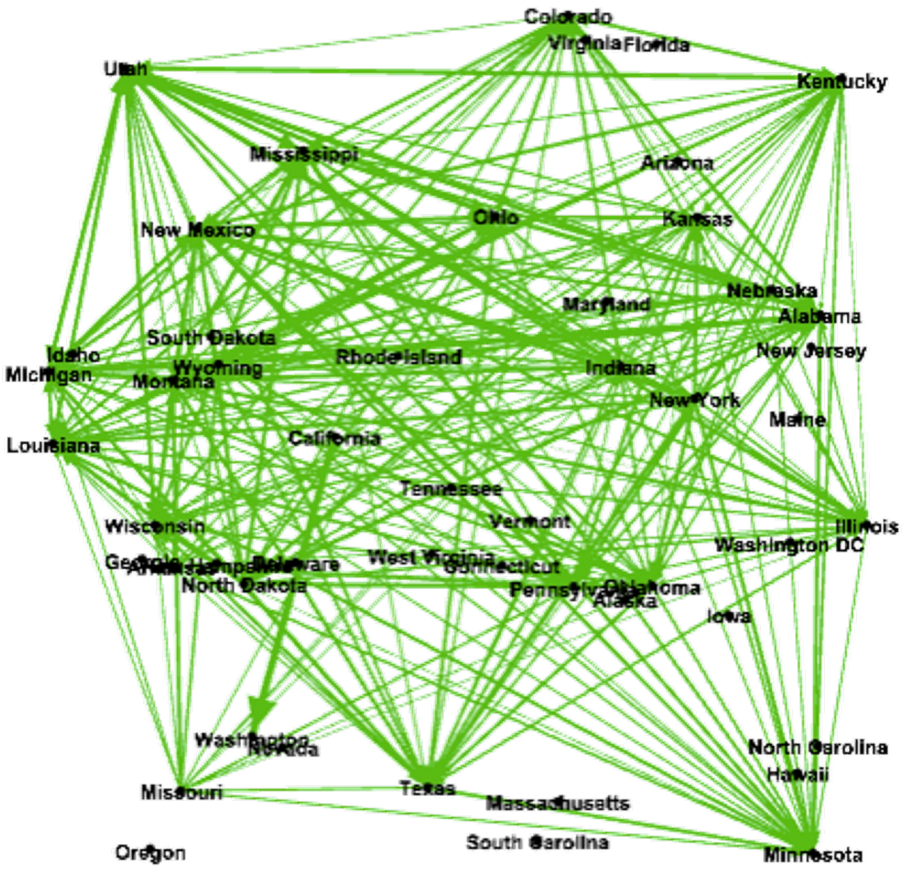
Goal 3: Optimal shipping path between Origin and Destination

- The focus of this experiment is to find out, influential hubs that play an important part while identifying the shortest distance between any two shipping hubs.
- For this analysis the focus will be only on mode of transportation as pipeline (dms_mode=6) and commodity as crude petroleum (sctg2= 16).
- The weights in Edge file signifies the distance between shipping hubs.

Betweenness Centrality:

- It is measure of centrality in graph based on shortest graphs. It is used to track down the most influencial nodes(key responsible node) in the network.
- In this context, the shipping port with the highest Betweenness Centrality would be essential to optimize shortest path between source shipping hub and target shipping hub.

Gephi Social Network Diagram:



Non- Zero Betweenness Centrality Hubs:

Shipping Hub	Betweenness Centrality
Indiana	14.08
Illinois	13
Kansas	5.58
Kentucky	5.58
Michigan	5.58
Oklahoma	5.58
Ohio	4.22
Texas	3.72
Louisiana	3.01
Montana	2.47
Wyoming	2.47
Colorado	1.94
Utah	1.94
New Mexico	0.78
Alabama	0.53
Mississippi	0.53

- The statistics shows out of 27 Crude Petroleum Shipping source, **Indiana** followed by **Illinois** are most influential hubs.
- This can be justified due to geometric location significance as both Shipping Hubs are in central USA.
- If we look at the pattern, the hubs located at extreme ends of USA (for example New York from East Coast or Alaska from North- West) has zero betweenness centrality and can be omitted while identifying the shortest path.

Conclusion & Future Work

- Using the concept of degree of node and eigenvalue centrality, we can identify the connectivity of shipping hub and its influence in Social Network.
- The weights assigned to each edge in network is taken to consideration while calculating the weighted degree of nodes.
- Betweenness centrality is used to track down the most influential shipping hubs(key responsible node) in the network while identifying the shortest paths.

Future Work :

- ➢ Using dynamic data to understand the impact of Covid-19 on Logistics/ Freight management and help select an optimized path that would ensure smooth flow of goods.
- ➢ Using trade policies data, we adjust the edge weights in the logistics network minimize the shipping cost of the freight.

References

- https://www.ey.com/en_us/supply-chain/how-covid-19-impacted-supply-chains-and-what-comes-next
- https://www.bloomberg.com/news/articles/2021-09-10/port-congestion-index-shows-u-s-import-rush-clogging-trade-hubs
- https://www.joc.com/port-news/us-ports
- https://ops.fhwa.dot.gov/freight/freight_analysis/faf/
- http://matthieu-totet.fr/Koumin/2013/12/16/understand-degree-weighted-degree-betweeness-centrality/

Dataset Link:

1. https://faf.ornl.gov/faf5/dtt_domestic.aspx
2. https://www.kaggle.com/usdot/freight-analysis-framework

STEVENS

INSTITUTE *of* TECHNOLOGY

THE INNOVATION UNIVERSITY®

stevens.edu

Ruchi Hiralal Mendhegiri