1 Problem Setting

In this project, we want to investigate how to develop marco-network to connect a nation. More specifically, given some OD flow data, we want to solve for the optimal network which satisfies all demand while minimize construction cost. Next, we will discuss several problem formulations, from the easiest to the hardest, along with algorithms to solve them and the results.

1.1 Static Formulation

First, let us tackle the simplest question: if the OD demand are invariant across time, how can we design optimal network? In this section, we use the yearly average OD demand as input.

Assume we are given OD pairs $e \in \{1, \cdots, N\}$, along with distance $l_e \geqslant 0$ and demand $d_e \geqslant 0$ (notice that pairs (x, y) and (y, x) are encoded in different indices). We want to find a network construction plan $x_n^m, y_e^m \in \{0, 1\}$ with capacity $c_e^m \geqslant 0$ for each mode m, edge e and node n that can satisfy all demand. More specifically, we use f_{oe}^m to represent the flow of mode m from origin o passing through edge e. The full formulation is shown below:

$$
\begin{aligned}
\min \sum_m (\sum_n C_p^m x_n^m + \sum_e C_c^m l_e y_e^m + \sum_e C_t^m l_e c_e^m) \\
s.t. \sum_m \sum_{e \in I(n)} f_{oe}^m = \sum_m \sum_{e \in O(n)} f_{oe}^m + d_{on},\ \forall o, n \\
\sum_o f_{oe}^m \leqslant c_e^m,\ \forall e, m \\
M y_e^m \geqslant c_e^m,\ \forall e, m \\
0.5(x_{e_1}^m + x_{e_2}^m) \geqslant y_e^m,\ \forall e, m = Aviation \\
f_{oe}^m \geqslant 0, c_e^m \geqslant 0, x_n^m, y_e^m \in \{0, 1\}
\end{aligned}
\tag{1}
$$

where d_{on} is the total amount of demand from o to node n, Cs are the unit cost, and M is the big-M coefficient to ensure that the problem is feasible. d_{oo} is defined as $-\sum_{n \neq o} d_{on}$. Currently, we set the cost for each mode to be:

Aviation: point construction fee $C_p = 500$, transportation fee $C_t = 0.001$ per distance metric;

Rail: line construction fee $C_c = 250$ per distance metric, transportation fee $C_t = 0.0003$ per distance metric;

Road: line construction fee $C_c = 100$ per distance metric, transportation fee $C_t = 0.001$ per distance metric.

1.2 Robust Formulation

The above formulation has a major drawback: the design will plan with regard to the average flow, so in many circumstances when the flows are more than expected, the design network will not be feasible to provide service to all the demand. The best way to resolve this is to use robust optimization with uncertainty set learned from the data. However, as such uncertainties are in the demand which is the right hand side of our formulation, the corresponding robust problem will be very hard to solve. On the contrary, we simply use the maximum observation for each OD pair as input. If this is proven to be too conservative, we can find the balance between this demand vector and the average one later.