Urban Simulation

Zhengyong Liu

February 2024

Word Count: 3133

This paper is for the assignment of CASA 0002 Urban Simulation. The **GitHub link** can be found here.

1 Topological Network

This section involves assessing the resilience of London's underground system by identifying stations whose removal could potentially compromise its stability. The following process will help determine which stations are crucial for the underground's operation and which methods are most suitable for identifying these critical stations.

1.1 Centrality Measures

In order to study the resilience of London's underground, we should first determine different centrality measures. We consider Degree Centrality, Betweenness Centrality, and Closeness Centrality. These three network analysis metrics are able to provide an in-depth understanding of the importance of the nodes in the transportation network(Yang et al., 2023). In addition, these metrics can evaluate dynamic network analysis, helping to understand traffic flow and node importance over time(Crescenzi et al., 2020).

1.1.1 Degree Centrality

Degree Centrality is proposed by Freeman et al. (2002) in 1978, and the core concept behind it is that central nodes in a network should have more connections. In a network, the degree of a node is the number of its adjacent edges. A high degree centrality means that the node connects many other nodes and may have more influence in information dissemination or control.

Degree centrality measures the number of direct connections a node has with other nodes in the network. It is defined as:

$$C_D(v) = \frac{\deg(v)}{n-1} \tag{1}$$

where $\deg(v)$ is the degree of node v, and n is the total number of nodes in the network(Freeman et al., 2002). This measure shows the number of direct connections each station has. Stations with high degree centrality are key transfer points and are highly accessible within the network.

Table 1: Ranking for London Underground Degree Centrality

Rank	Station	Degree Centrality
1	Stratford	0.0225
2	Bank and Monument	0.0200
3	King's Cross St. Pancras	0.0175
4	Baker Street	0.0175
5	Earl's Court	0.0150
6	Oxford Circus	0.0150
7	Liverpool Street	0.0150
8	Waterloo	0.0150
9	Green Park	0.0150
10	Canning Town	0.0150

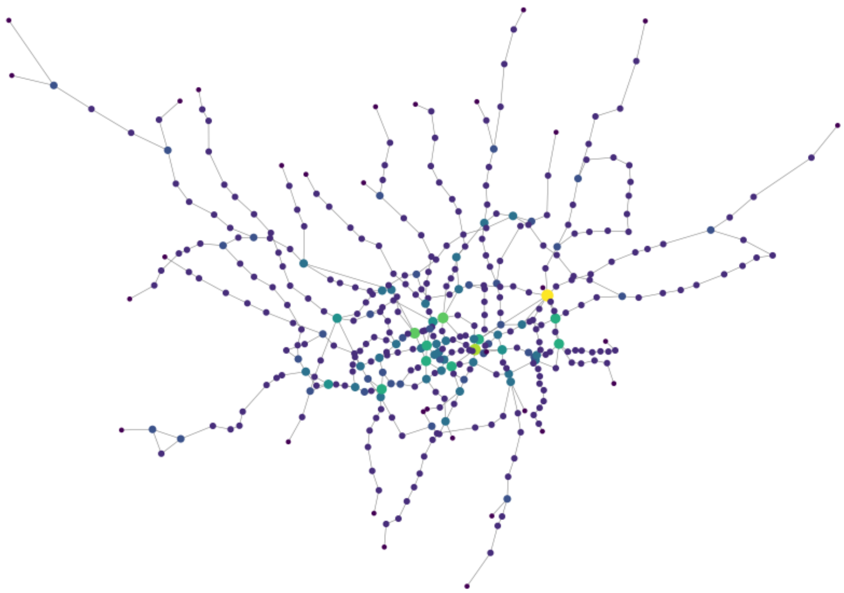

Figure 1: Enter Caption

1.1.2 Betweenness Centrality

Betweenness centrality measures the extent to which a node lies on paths between other nodes(Freeman, 1977). Specifically, it quantifies the number of times a node acts as a bridge along the shortest path between two other nodes. The formula is:

$$C_B(v) = \sum_{s \neq v \neq t} \frac{\sigma_{st}(v)}{\sigma_{st}} \tag{2}$$

where σ_{st} is the total number of shortest paths from node s to node t and $\sigma_{st}(v)$ is the number of those paths that pass through v. This measure reflects the number of times a station acts as a bridge along the shortest path between two other stations. High betweenness centrality indicates a station's critical role in maintaining connectivity across the network(Bloch et al., 2023).

Table 2: Ranking for London Underground Betweenness Centrality

Rank	Station	Betweenness Centrality
1	Stratford	23768.093434
2	Bank and Monument	23181.058947
3	Liverpool Street	21610.387049
4	King's Cross St. Pancras	20373.521465
5	Waterloo	19464.882323
6	Green Park	17223.622114
7	Euston	16624.275469
8	Westminster	16226.155916
9	Baker Street	15287.107612
10	Finchley Road	13173.758009

1.1.3 Closeness Centrality

Closeness centrality measures how quickly a node can access all other nodes in the network. It is defined as the inverse of the average shortest path distance to all other nodes:

$$C_C(v) = \frac{n-1}{\sum_{u \neq v} d(v,u)} \tag{3}$$

where $d(v,u)$ is the distance between nodes v and u. Closeness centrality measures how quickly a station can reach all other stations in the network. A high closeness score indicates a station's strategic location for facilitating fast

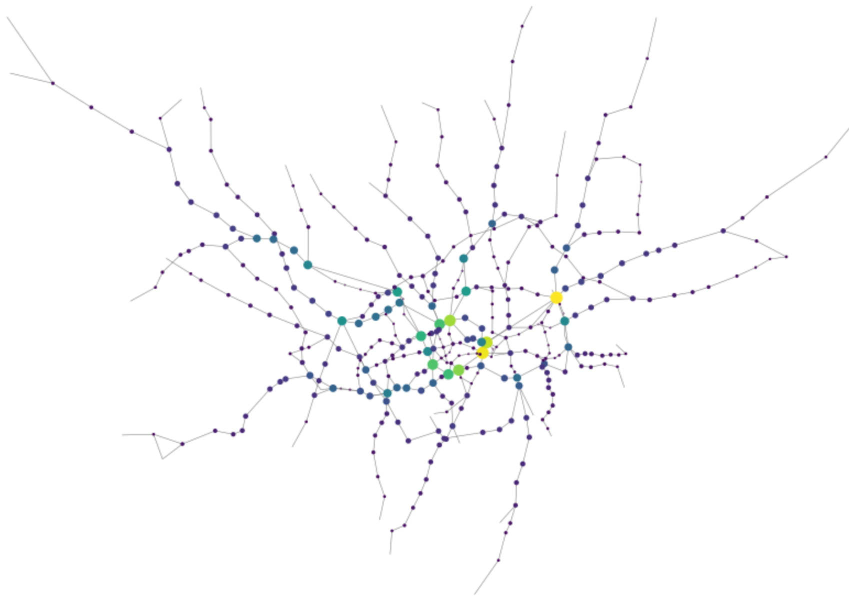

Figure 2: Enter Caption

transport across the network(Borgatti, 2005).

Table 3: Ranking for London Underground Closeness Centrality

Rank	Station	Closeness Centrality
1	Green Park	0.114778
2	Bank and Monument	0.113572
3	King's Cross St. Pancras	0.113443
4	Westminster	0.112549
5	Waterloo	0.112265
6	Oxford Circus	0.111204
7	Bond Street	0.110988
8	Farringdon	0.110742
9	Angel	0.110742
10	Moorgate	0.110314

1.2 Impact Measures

In this section, we will introduce two ways to evaluate the resilience of London Underground, namely Global Efficiency and Largest Connected Component (LCC)

1.2.1 Global Efficiency (network science)

Global Efficiency Network is an index proposed by Latora and Marchiori (2001) to measure the efficiency of complex networks. It reflects the efficiency of information flow among nodes in the network and measures the extent to which different nodes in the network are interconnected by the shortest path. The shorter the path, the more efficient the network. Higher global efficiency means that the network transmits information in fewer steps, reflecting higher connectivity and redundancy. The formula is shown as follows:

$$E_{glob} = \frac{1}{N(N-1)} \sum_{i \neq j} \frac{1}{d_{ij}} \tag{4}$$

N represents the total number of nodes in the network.d_{ij} is the shortest path length between nodes i and j, typically measured by the number of edges traversed. It indicates the shortest distance between nodes in the network. The term $\frac{1}{d_{ij}}$ represents the efficiency between nodes i and j. Shorter paths imply higher efficiency. The

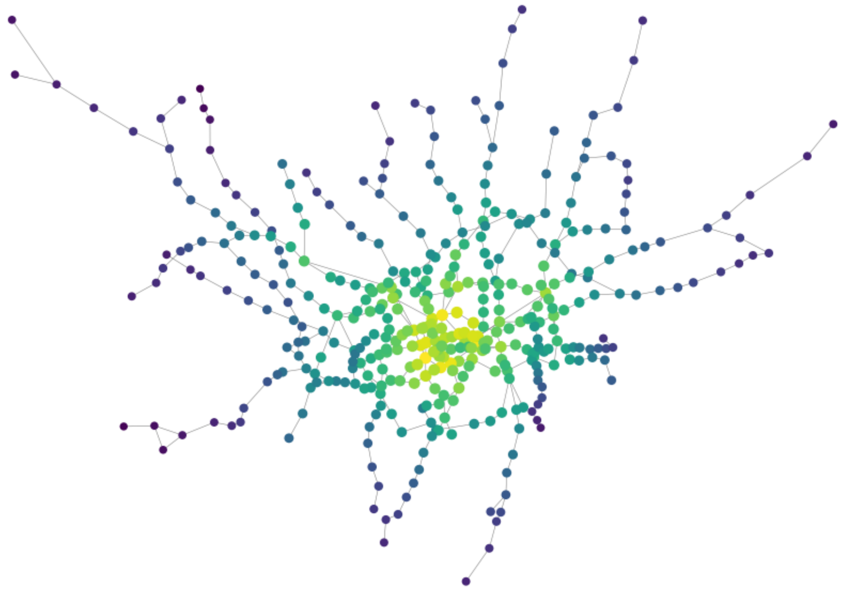

Figure 3: Enter Caption

sum $\sum_{i \neq j} \frac{1}{d_{ij}}$ is the total efficiency between all unique node pairs. The coefficient $\frac{1}{N(N-1)}$ normalizes the result to represent the average efficiency across the network(Ek et al., 2015).

1.2.2 Largest Connected Component, LCC

The Largest Connected Component (LCC) is the largest connected component in a network, and its index is mainly used to describe the connectivity and integrity of the network. Its expression method can be expressed by the following formula:

$$LCC = \max_{C} |C| \tag{5}$$

C represents a connected component, which refers to the part of the network where nodes can be connected to each other through edges. Any node or group of nodes that can be connected to each other by one or more edges belongs to a connected component. $|C|$ represents the number of nodes in the connected component.

In addition, the LCC concept has been applied to identify vulnerable lines within the metro network, integrating ridership data to precisely locate lines whose failures will disproportionately affect the overall connectivity of the network (Tang et al., 2021). This approach highlights the importance of considering the structural and functional aspects of the metro network in maintaining and enhancing its resilience. LCC in metro networks is an important part of analyzing and improving network robustness, efficiency, and resilience to random and directed disruptions (Nistor et al., 2018)(Angeloudis and Fisk, 2006).

1.2.3 Node Removal

In this section, we will evaluate node deletion with two different evaluation methods for different centrality. Two methods of Non-Sequential removal & Sequential removal will be considered for node removal

Non-Sequential removal: According to the table created in I.1., starting from the most important node, a node is removed in order of ranking until at least 10 nodes are removed. After each removal, the impact of the removal is evaluated using the two metrics mentioned in I.2. and then proceed to remove the next node.

Sequential removal: remove the highest ranked node and evaluate the impact using the 2 measures. After removal, re-compute the centrality measure. Remove the highest-ranked node in the new network and evaluate the impact. Continue until removing at least 10 nodes.

Figure 4 shows the results of Non-Sequential removal and Sequential removal under different Centrality results of Global Efficiency.

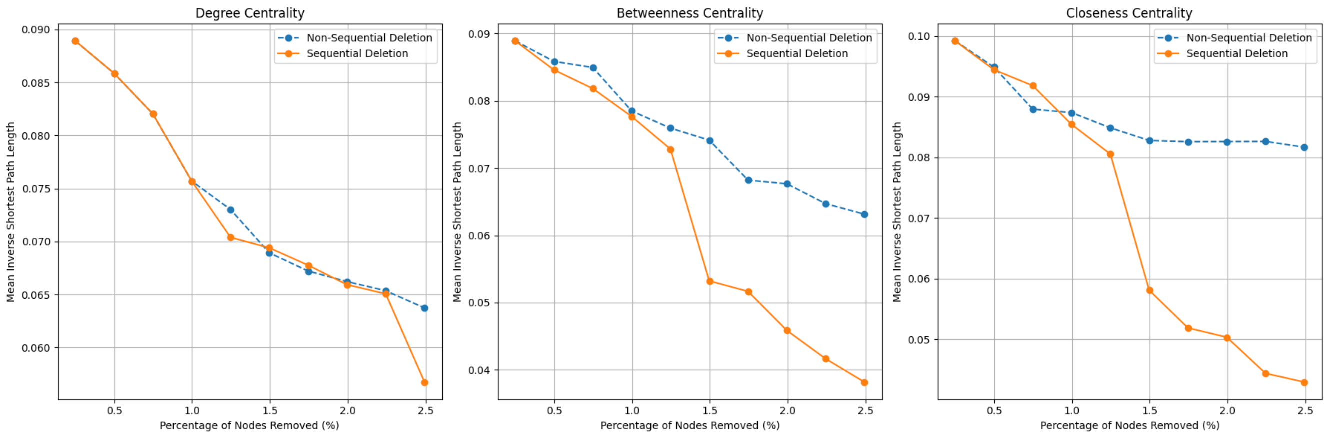

Figure 4: Global Efficiency

Figure 5 shows the results of Non-Sequential removal and Sequential removal under different Centrality results of Normalized Largest Connected Component.

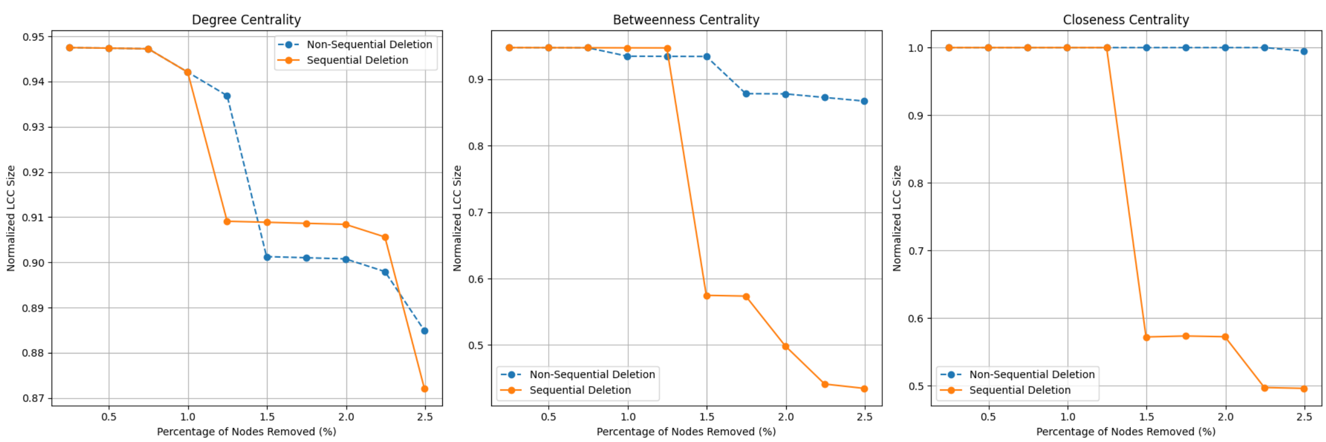

Figure 5: Normalized Largest Connected Component

In the degree centrality chart, we can see that Normalized LCC is relatively stable at the initial stage of node removal regardless of sequential deletion or sequential deletion, but when the proportion of nodes removed exceeds 1%, the curve begins to decline sharply. This suggests that when nodes with higher connectivity in the network are removed, the connectivity of the network is affected, but not much in the initial period until a critical point is reached.

In the diagram of closeness centrality, the situation is similar to degree centrality, but closeness centrality focuses on the average distance of a node from all other nodes in the network, so it measures the access efficiency of a node. The figure shows that the downward trend of the curve is more obvious in the case of sequential deletion, which means that the average efficiency of the whole network decreases when the nodes with high closeness centrality are removed.

In contrast, in the graph of betweinness centrality, we can observe a significantly different phenomenon: as soon as nodes are removed, Normalized LCC Size and Global Efficiency decrease rapidly regardless of sequential deletion or sequential deletion, and the impact of sequential deletion is more significant. This means that stations with high betweenness centrality play an extremely critical role in connecting different lines and routes in the underground transport network. Once these critical nodes are removed, their impact on the network as a whole becomes apparent, leading to a rapid deterioration in transport efficiency and connectivity.

Under the three centrality measurement methods, the decrease of NormalizedLCC size and the increase of Global Efficiency caused by sequential deletion are faster than those caused by disordered deletion, which indicates that in practice, the sequential deletion strategy of removing key nodes step by step can more effectively evaluate the resilience of the network in the face of continuous interference. Therefore, the sequential deletion strategy shows a greater impact on network resilience.

Based on these trends and values, the inverse of Global Efficiency is more sensitive to the impact of node removal. Especially in the case of sequential deletion, it can indicate the decrease of network efficiency more quickly. Although the size of Global Efficiency is also an important indicator, it mainly reflects whether the connectivity of the network is destroyed. Therefore, Global Efficiency would be a better measure if our goal were to assess the impact of subway station removal on overall operational efficiency in a more nuanced way.

2 Flows: Weighted Network

2.1 Weighted Centrality

In the previous analysis, we can conclude that Betweenness Centrality is the most suitable method for analysis among several methods, so we conduct further analysis on it in this chapter. The structure of weighted betweenness centrality and unweighted betweenness centrality is consistent in the overall formula, but from the perspective of the calculation method, unweighted only considers the number of edges on the path and ignores the weight of edges. In this case, all edges are equally weighted. Weighting considers the weight of each edge, and the shortest path may not have the fewest edges, but the lowest weight. In the calculation process, the "length" of the data will be used as the weighting element.

When the edges are weighted, the betweenness centrality is calculated based on these weighted shortest paths. Some routes that were not originally regarded as shortest paths may become new shortest paths due to lower weights. Here are the results:

Table 4: Ranking for London Underground Weighted Betweenness Centrality

Rank	Station	Weighted Betweenness Centrality
1	Bank and Monument	17656.0
2	King's Cross St. Pancras	16693.0
3	Stratford	14548.0
4	Oxford Circus	13561.0
5	Euston	13240.0
6	Baker Street	12150.0
7	Earl's Court	11475.0
8	Shadwell	11128.0
9	Waterloo	10408.0
10	South Kensington	10335.0

From this we can see that some sites change their positions in the weighted and unweighted betweenness centrality rankings. For example, "Bank and Monument" rose from unweighted second place to weighted first place. Some sites that do not appear in the unweighted ranking appear in the weighted ranking, indicating that these sites are important nodes in the network. Some sites dropped in the weighted ranking or did not appear in the top ten at all. This may indicate that these stops appear important in the unweighted case, but are not optimal transit options when actual travel costs or time are considered.

2.2 Adjust the measure for a weighted network

The algorithm for "Global Efficiency" mentioned in 1.2 (Formula 2) is traditionally calculated using node-based metrics. However, given that traffic data is associated with edges rather than nodes, the method is adapted to accommodate such edge-centric data. To weight edges, the calculation method of d_{ij} needs to be modified to make it the weighted shortest path length. This means that each edge will have a corresponding weight (i, j) will have a corresponding weight w_{ij}. At the same time, if we also consider the flow of edges, we can further adjust the weight to reflect the impact of flow. The weight may be inversely proportional to the flow, so that the edge with a large flow has a low weight, which means higher efficiency (Wang et al., 2019). The Modified formula is as follows (modified-global Efficiency) :

$$E'_{glob} = \frac{1}{N(N-1)} \sum_{i \neq j} \frac{w_{ij}}{d'_{ij}} \tag{6}$$

In this modified formula, d'_{ij} is the weighted shortest path length, calculated using weighted shortest path algorithms, like Dijkstra's algorithm. The weight w_{ij} represents the attribute of edge (i, j), which can be calculated as $w_{ij} = \frac{1}{f_{ij}}$, where f_{ij} is the flow on the edge.

Similarly, in the second method of "Normalized Largest Connected Component", the proportion of the most important part in the network is evaluated by comparing the ratio of the connected component with the largest traffic to the total traffic of all connected components in the figure. The specific formula is as follows (Modified-Weighted-NLCC):

$$\text{Modified Weighted Normalized Largest Connected Component} = \frac{\text{MaxFlow}(C_i)}{\sum_{j \in S} \text{Flow}(C_j)} \tag{7}$$

Where C_i is the connected component with the largest flow. $\text{MaxFlow}(C_i)$ is the flow of the connected component that has the largest total flow among all connected components. S is the set of all connected components in the graph. $\text{Flow}(C_j)$ is the total flow of the j connected component (Berche et al., 2009).

2.3 Largest impact of station close

Based on the two modified measures proposed in 2.1.4, we can use the results to evaluate which station closures have a greater impact on the population. Here are some of the results.

Weighted Betweenness Centrality		
Removal-Station-Name	Modified-Global-Efficiency	Modified-Weighted-NLCC
Stratford	0.000053	0.976435
Bank and Monument	0.000050	1.000000
King's Cross St. Pancras	0.000047	1.000000
Removal-all-3-station-above	4.887152920206759e-05	0.9734734323642835
Non-Weighted Betweenness Centrality		
Removal-Station-Name	Modified-Global-Efficiency	Modified-Weighted-NLCC
Stratford	0.000053	0.976435
Bank and Monument	0.000050	1.000000
King's Cross St. Pancras	0.000047	1.000000
Removal-all-3-station-above	4.887152920206759e-05	0.9734734323642835

Table 5: Two Modified Measures evaluate the result of Weighted Centrality and non Weighted Centrality

In the Modified-Global-Efficiency column, we see relatively similar values for all three sites, suggesting that none has a particularly large impact on the overall network efficiency in terms of shortest paths and traffic. In the Modified-Global-Efficiency column, we find a score of 1.0 for both "Bank and Monument" and "King's Cross St. Pancras". This means that removing these sites does not destroy the cohesion of the network -there is still a large connected component. On the other hand, "Stratford" has a modified weighted Global-Efficiency score of 0.976435, which indicates a slight decrease in network cohesion after deletion. Taking these results into account, "Stratford" appears to have the most significant impact on the station network when closed. In addition, we can see the impact of Stratford's deletion on the whole network through counter-verification. Through calculation, it can be found that after its deletion, The efficiency of the entire transport network decreased by 12%, compared to 4.22% in King's Cross and 4.46% in Bank and Monument, and the results are obvious.

Station-Name	Efficiency-Change
Stratford	-12.185756242416284 %
Bank and Monument	-4.4653441451587135 %
King's Cross St. Pancras	-4.222232074112703 %

Table 6: Efficiency Change in Different Node removal

3 Spatial Interaction models

Spatial interaction model, The spatial movement of people is treated as a behaviour, and attempts are made to explain this behaviour through mathematical modelling. The model is based on the concept of Newton's law of universal gravitation, which states that the interaction between two geographical areas is proportional to their population and economic size, and inversely proportional to the distance between them (de Vries et al., 2000).

3.1 Models

3.1.1 The Unconstrained Model

The Unconstrained Model can be expressed as follows:

$$T_{ij} = KO_iD_j \exp(-\beta c_{ij}) \tag{8}$$

T_{ij} is the predicted flow from origin i to destination j. K is a constant for model calibration. O_i represents the capacity of origin i to generate flows. D_j represents the attractiveness or capacity of destination j. $\exp(-\beta c_{ij})$ is the exponential decay function of the cost or distance c_{ij} between i and j. β is a parameter indicating sensitivity to the cost or distance(Wilson, 1971).

This model is subject to the constraint that the sum of all flows matches a total observed flow T:

$$\sum_{i=1}^{n}\sum_{j=1}^{m} T_{ij} = T \tag{9}$$

The constraint states that the sum of all flows T_{ij} in the model must equal some total observed flow T. This is an additional condition that ensures the model's outputs are consistent with observed data for the total amount of flow. Also, the model assumes that O_i, D_j, and c_{ij} are independent, which may not always be the case in real-world scenarios. There might be feedback loops where the flow between locations affects the costs or the opportunities at the origins and destinations(Roy and Thill, 2004).

3.1.2 The Constrained Models

Wilson (1971) provides a theoretical basis for the spatial interaction model based on entropy maximization; Utility maximization provides another rationale that presents the same result. Wilson also developed different specifications of the spatial interaction model: the origin constrained model, the destination constrained model. The origin constraint model ensures that the sum of the estimated outflows is the same as the sum of the observed outflows from each origin. Similarly, the destination constrained model ensures that the estimated inflows to each destination are the same as the sum of the observed inflows (Antolín et al., 2017).

- Origin - Constrained Model can be explained by the following

$$T_{ij} = A_iO_iD_j \exp(-\beta c_{ij}) \tag{10}$$

- Destination - Constrained Model can be explained by the following

$$T_{ij} = B_jO_iD_j \exp(-\beta c_{ij}) \tag{11}$$

- A Doubly Constrained Model ensures that these two pairs sum to the same. A doubly constrained spatial interaction model can be written as :

$$T_{ij} = A_iB_jO_iD_j \exp(-\beta c_{ij}) \tag{12}$$

3.2 Calibrate the Parameter

Based on the existing variables (station-origin, station-destination, flows, jobs, distance), analyze and calculate some corresponding data for the value of β. Wilson (1971) shows how to use the iterative Furness algorithm to estimate his spatial interaction model derived from entropy maximization. This algorithmic estimation is quite different from the linear regression used to estimate simple spatial interaction models, which involves a logarithmic transformation. This is how we use it in this chapter:

Unconstrained Model:

$$\text{flows} \sim \log(\text{jobs}) + \log(\text{distance}) - 1 \tag{13}$$

Origin-Constrained Model:

$$\text{flows} \sim \text{station\_origin} + \log(\text{jobs}) + \log(\text{distance}) - 1 \tag{14}$$

Destination-Constrained Model:

$$\text{flows} \sim \text{station\_destination} + \log(\text{jobs}) + \log(\text{distance}) - 1 \tag{15}$$

Double-Constrained Model:

$$\text{flows} \sim \text{station\_origin} + \text{station\_destination} + \log(\text{jobs}) + \log(\text{distance}) - 1 \tag{16}$$

	Unconstrained Model	Origin-Constrained Model
jobs γ	0.8881	0.7686
distance -β(cost of function)	0.4860	0.8782
CalcRSqaured	0.103	0.3883
CalcRMSE	124.523	102.89
	Destination-Constrained Model	Double-Constrained Model
jobs γ	1.2391	1.2454
distance -β(cost of function)	0.7212	0.9097
CalcRSqaured	0.0921	0.40768
CalcRMSE	125.893	101.335

Table 7: Parameters in Different Models

Based on these metrics, the Origin-Constrained Model and the dual Constrained Model seem to be the best performing of the four models as it has a higher coefficient of determination and a lower RMSE. And both models can explain 40% of the predicted values.

3.3 Scenario A

When we consider a 50% reduction in the number of jobs in Canary Wharf while keeping the total population constant, we usually mean a 50% reduction in the number of jobs arriving at Canary Wharf from all stations, not a 50% reduction from all stations to all stations. Here are the jobs that calculate the new Canary Wharf. Through the calculation, we can find that the number of jobs in London's overall network has decreased from 376,833,285 to 367,899,941, a total reduction of 8,933,344 by about 2.37

$$\text{New\_jobs} = 0.5 \times \text{Desti\_Canary\_Wharf\_jobs} + \text{Rest\_jobs} \tag{17}$$

Considering this scenario, although the R Square of the bidirectional constraint model is good, it is not applicable to this scenario, so we choose the model with the origin constraint (Equation 14).

$$\text{Scenario A Flows} \sim \exp(A_i + \gamma * \log(\text{Reduced\_jobs}) - \beta * \log(\text{distance})) \tag{18}$$

Using the parameter values in Table7, the final calculation formula can be adjusted as follows:

$$\text{Scenario A Flows} \sim A_i \cdot O_i \cdot D_j^{0.7686} \cdot \exp(-0.8782 \cdot d_{ij}) \tag{19}$$

In order to ensure that the total population remains unchanged, we need to adjust the value of $A_i$to ensure that the overall flow remains unchanged. Below is a rendering of visualizing A_i.

Station Origin	Old A_i	New A_i	Change %
West India Quay	1.184739	1.391539	17.455369%
Heron Quays	3.734459	3.968836	6.276076%
Poplar	4.119542	4.266556	3.568703%
South Quay	4.147878	4.264868	2.820476%
Westferry	4.312771	4.425456	2.612802%
All Saints	3.380981	3.461319	2.376162%
Blackwall	3.609648	3.689761	2.219420%
Crossharbour	4.336149	4.410379	1.711896%
Mudchute	4.228312	4.289648	1.450593%
Langdon Park	4.218719	4.279849	1.449018%

Table 8: Top 10 stations with the most A_i changes

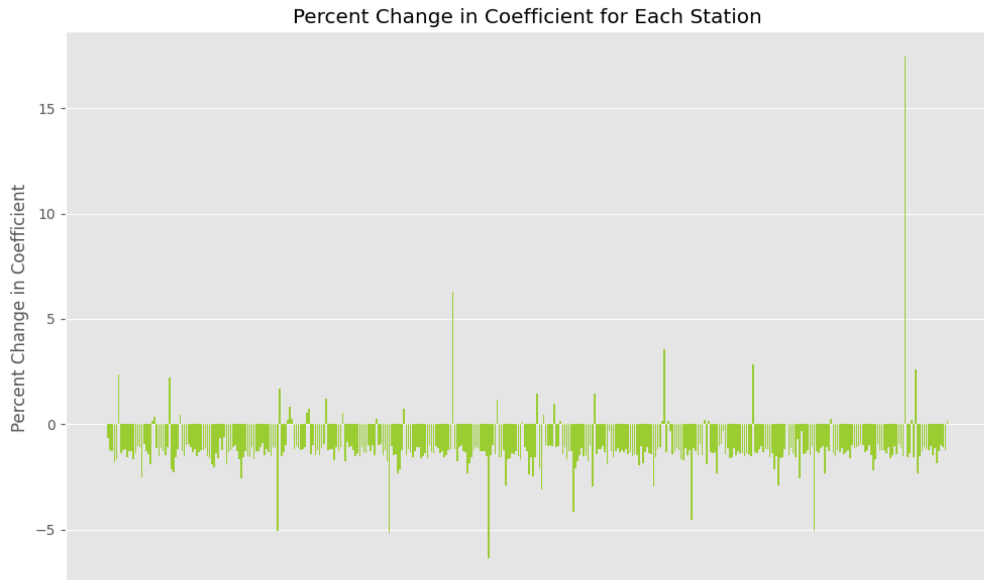

Figure 6: A_i Change % for all Station

According to the above table and picture, after the number of jobs in Canary Wharf is reduced by half, the corresponding $A_i$value of Underground stations around Canary Wharf will increase, and the closer the distance is, the greater the increase of $A_i$will be. This phenomenon is basically in line with the reality. The specific analysis will be put in 3.5 for further explanation.

3.4 Scenario B

In this scenario, βwould need to be recalculation if we were to consider a significant increase in hypothetical transportation costs. We still use the origin constraints model in this model. Since our initial βvalue is 0.8782, we need to double and triple it. B1=β_1=1.7564, B2=β_2=2.6346. The changes in flows and $A_i$can be revisited by adjusting the value of β. Here are the adjusted figures.

Top 10 stations with the largest flow growth		The top 10 stations with the largest flow reduction	
Station Name	%	Station Name	%
Stratford High Street	556.218274	Chesham	-94.238683
Elverson Road	406.194690	Woolwich Arsenal	-87.420958
Gidea Park	263.210702	Epping	-86.546185
Goodmayes	253.587116	Amersham	-84.765625
Rotherhithe	243.613707	Heathrow Terminal 5	-78.461538
Seven Kings	203.398927	West Ruislip	-78.350515
Maryland	201.142857	High Barnet	-77.859391
Abbey Road	166.956522	Clapham Junction	-77.560877
Penge West	164.666667	West Croydon	-76.367188
West India Quay	159.000000	Beckton	-75.113122

Table 9: Top 10 stations with the largest flow growth & reduce in β_1

Top 10 stations with the largest flow growth		The top 10 stations with the largest flow reduction	
Station Name	%	Station Name	%
Stratford High Street	1640.101523	Chesham	-94.650206
Elverson Road	761.061947	Woolwich Arsenal	-93.202349
Gidea Park	362.876254	New Cross	-92.112950
Rotherhithe	353.894081	Beckton	-88.461538
Goodmayes	351.244510	Clapham Junction	-87.864762
Fairlop	321.428571	Tower Gateway	-86.535009
Ruislip Manor	276.715686	Epping	-86.345382
Penge West	254.666667	South Hampstead	-84.745763
St James Street	225.000000	Richmond	-80.266815
Seven Kings	221.466905	High Barnet	-79.853095

Table 10: Top 10 stations with the largest flow growth & reduce in β_2

3.5 Evaluating Scenarios

In Scenario A, the overall impact of job reductions at Canary Wharf on the flow across the vast transport network is not very large. flows decreased in 23% of stations and increased in 15% of stations.

Under Scenario B1, it can be clearly found from the table that the change of beta value (twice) will obviously affect the traffic of the whole network. As shown in the figure below, the change of flows is very obvious. flows decreased in 24.06% stations, increased in 63.74% stations, and remained unchanged in 12.18% stations. In the previous analysis, we looked at the change value of total flows at each station. Now let's dig a little deeper and see that after the change in beta value, It can be found that flows from 'London Bridge' To 'Bank and Monument' increase the most (as shown in the figure below).

Station Origin	Station Destination	Flow Change
London Bridge	Bank and Monument	10522
Liverpool Street	Moorgate	9267
Liverpool Street	Bank and Monument	7299
Waterloo	Southwark	7159
Waterloo	Westminster	5839
Canary Wharf	Heron Quays	5775
Waterloo	Embankment	4985
Victoria	St. James's Park	4906
Bank and Monument	Liverpool Street	4773
Stratford	Stratford High Street	4712

Table 11: Top 10 flows change from station to station in B1

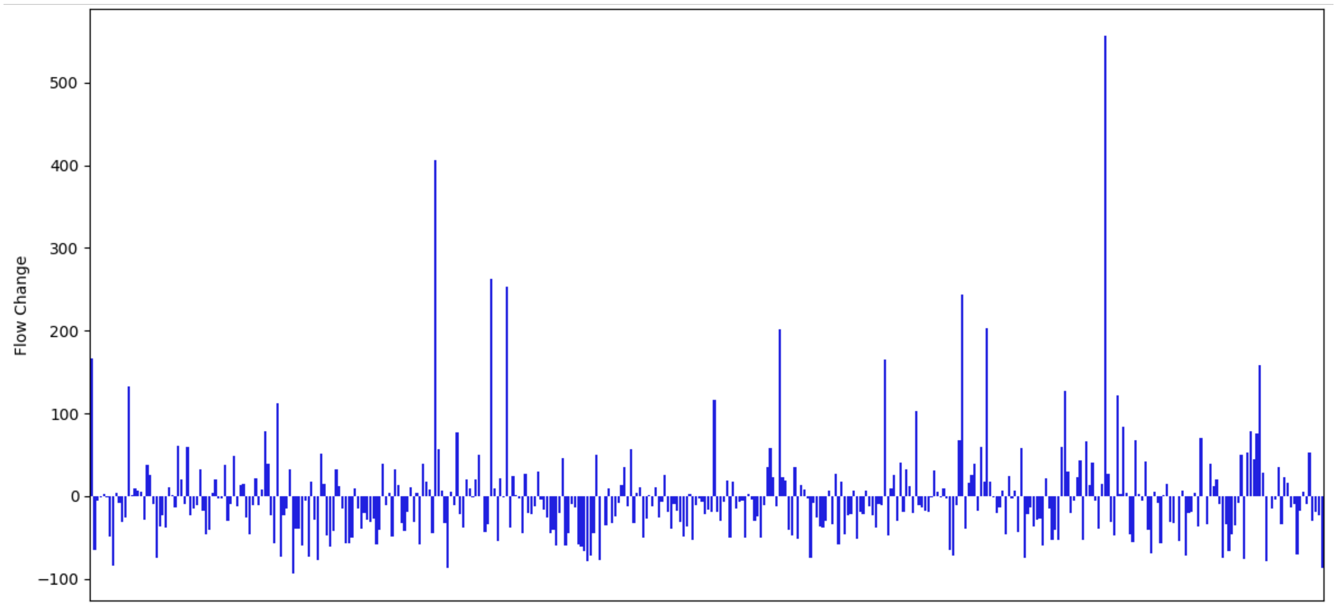

Figure 7: Station Flow change based on Scenario B1

In Scenario B2, where the β value triples, look again at the changes in flows between sites (table below). We can find that although there are slight changes in the order, there are significant changes in the flow change at the same origin and the same arrival.

Station Origin	Station Destination	Flow Change
London Bridge	Bank and Monument	19684
Liverpool Street	Moorgate	15225
Stratford	Stratford High Street	13440
Waterloo	Southwark	13235
Canary Wharf	Heron Quays	11159
Waterloo	Westminster	10134
Victoria	St. James's Park	9918
Liverpool Street	Bank and Monument	9747
Paddington	Edgware Road	8500
Waterloo	Embankment	8210

Table 12: Top 10 flows change from station to station B2

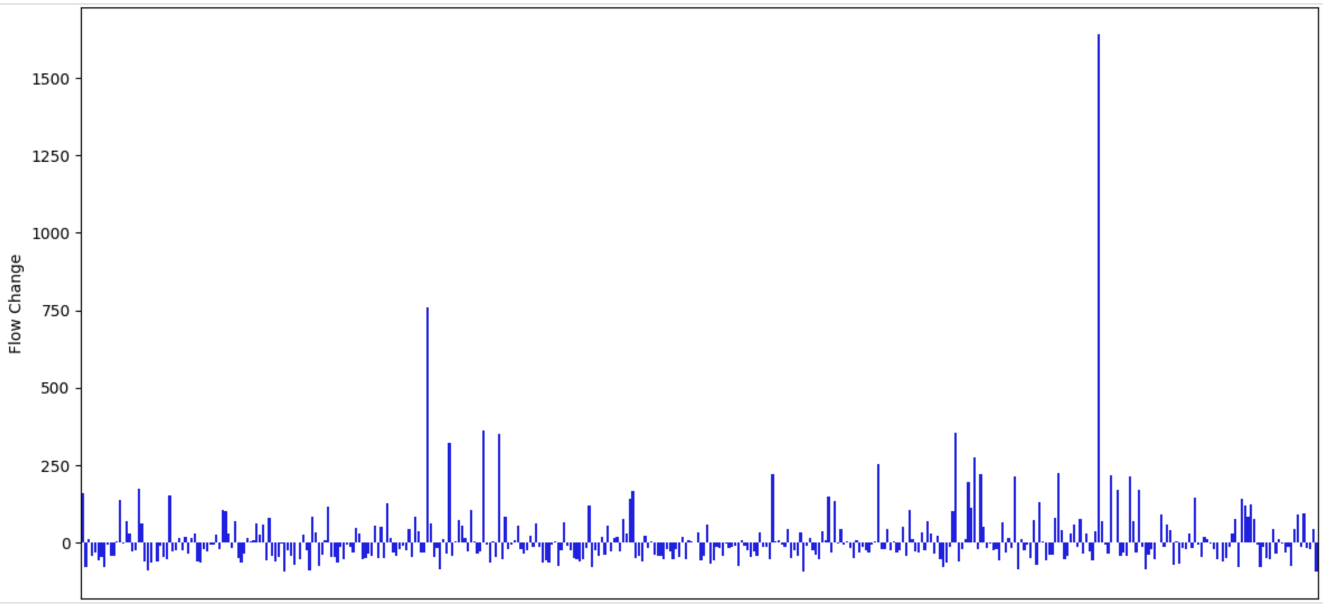

Figure 8: Station Flow change based on Scenario B2

Based on the Figures and Tables in 4.2 and 4.3, we can assume our conclusion that first small changes in β lead to drastic changes in flows. However, as the value of β increases, that is, the commuting cost increases, people will become more cautious in their travel process and are more inclined To take the route that they consider the most important, which in the above study is 'London Bridge' To 'Bank and Monument'. To confirm our conjecture, we can calculate the Standard deviation of prediction error of B1 and B2 for the true flows. The value of B1 is 1473, and the value of B2 is 230, which indeed proves that the change of β will lead to a large range of changes in flows.

References

Angeloudis, P. and Fisk, D. (2006). Large subway systems as complex networks. *Physica A: Statistical Mechanics and its Applications*, 367:553–558.

Antolín, G., Cordera, R., and Alonso, B. (2017). Spatial interaction models. In *Land Use–Transport Interaction Models*, pages 55–67. CRC Press.

Berche, B., Von Ferber, C., Holovatch, T., and Holovatch, Y. (2009). Resilience of public transport networks against attacks. *The European Physical Journal B*, 71:125–137.

Bloch, F., Jackson, M. O., and Tebaldi, P. (2023). Centrality measures in networks. *Social Choice and Welfare*, 61(2):413–453.

Borgatti, S. P. (2005). Centrality and network flow. *Social networks*, 27(1):55–71.

Crescenzi, P., Fraigniaud, P., and Paz, A. (2020). Simple and fast distributed computation of betweenness centrality. In *IEEE INFOCOM 2020-IEEE Conference on Computer Communications*, pages 337–346. IEEE.

de Vries, J. J., Nijkamp, P., and Rietveld, P. (2000). Alonso's general theory of movement: Advances in spatial interaction modeling. Technical report, Tinbergen Institute Discussion Paper.

Ek, B., VerSchneider, C., and Narayan, D. A. (2015). Global efficiency of graphs. *AKCE International Journal of Graphs and Combinatorics*, 12(1):1–13.

Freeman, L. C. (1977). A set of measures of centrality based on betweenness. *Sociometry*, pages 35–41.

Freeman, L. C. et al. (2002). Centrality in social networks: Conceptual clarification. *Social network: critical concepts in sociology. Londres: Routledge*, 1:238–263.

Latora, V. and Marchiori, M. (2001). Efficient behavior of small-world networks. *Physical review letters*, 87(19):198701.

Nistor, M. S., Bein, D., Teodorescu, H. N., and Pickl, S. W. (2018). Finding the maximal day-time dependent component of a subway system. In *Advances in Human Factors in Simulation and Modeling: Proceedings of the AHFE 2017 International Conference on Human Factors in Simulation and Modeling, July 17–21, 2017, The Westin Bonaventure Hotel, Los Angeles, California, USA 8*, pages 562–572. Springer.

Roy, J. R. and Thill, J. C. (2004). Spatial interaction modelling. *Papers in Regional Science*, 83(1):339–361.

Tang, J., Li, Z., Gao, F., and Zong, F. (2021). Identifying critical metro stations in multiplex network based on d–s evidence theory. *Physica A: Statistical Mechanics and its Applications*, 574:126018.

Wang, Y., Wang, S., and Deng, Y. (2019). A modified efficiency centrality to identify influential nodes in weighted networks. *Pramana*, 92(4):68.

Wilson, A. G. (1971). A family of spatial interaction models, and associated developments. *Environment and Planning A*, 3(1):1–32.

Yang, Z., Li, H. L., Yang, Z. S., and Ye, L. Z. (2023). Out-degree betweenness centrality based on betweenness centrality. In *International Conference on Cryptography, Network Security, and Communication Technology (CNSCT 2023)*, volume 12641, pages 103–108. SPIE.