

UNIVERSITY COLLEGE LONDON
FACULTY OF THE BUILT ENVIRONMENT



BENVGSC5 Urban Simulation

Understanding Urban Modelling Methodologies

Qiqing Huang, 17014982

Network: 991 words
ABM: 980 words
Spatial Model: 986 words

Centre for Advanced Spatial Analysis
25th April, 2018

Contents

1	Networks	2
1.1	Introduction	2
1.2	Identifying Key Stations	2
1.2.1	Centrality Measures of Key Stations	2
1.2.2	Results of Node Centrality	3
1.3	Different Assumptions of Attacks	4
1.3.1	Disrupting the Most Connected Stations	4
1.3.2	Blocking the Most Pathways	6
1.3.3	Fragmenting the Whole Tube Network	8
1.4	Final Results and Conclusion	9
2	Agent Based Models	10
2.1	Introduction	10
2.2	Experiments	10
2.2.1	Change of Initial Number of Infected	10
2.2.2	Multiple Parameter Settings	11
2.3	Explore of Determining Minimum Runs	12
3	Spatial Interaction Models	13
3.1	Initial Spatial Interaction Models	13
3.1.1	Unconstrained model	13
3.1.2	Production Constrained Model	16
3.1.3	Attraction Constrained Model	17
3.1.4	Doubly Constrained Model	18
3.2	Conclusion	18

1 Networks

1.1 Introduction

This article aims at understanding and predicting the potential terrorist attack strategies in case of London underground system using network analysis. Section 1.2 introduces measures of identifying key stations; Section 1.3 derives hypotheses based on different attack aims and numbers of target stations. The final section 1.4 compares all the results and get an optimal result.

The London underground networks are treated as undirected graphs where nodes represent stations and undirected links represent line segments. In this case, there are 306 nodes and 410 links.

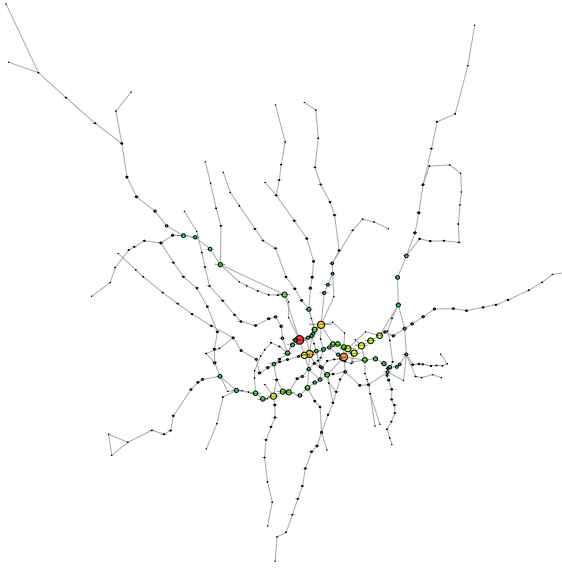


Figure 1: The topographical map of the London underground network. Stations of larger betweenness centrality have larger vertex size and more obvious color (Red > Green > Blue).

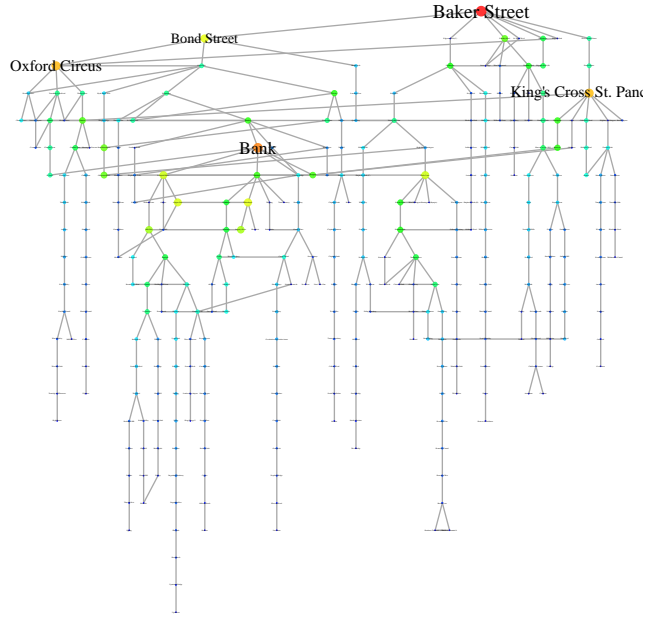


Figure 2: The tree topology of the London underground network. The top five stations of betweenness centrality are marked by larger label size

1.2 Identifying Key Stations

How to identify keyplayers in a network has been widely researched. Among all measures, betweenness centrality is the most popular one as it sums the proportion of shortest paths of all connected pairs that pass by a given node (Freeman 1978).

Other centrality measures are also checked. The goal of finding targeted attack stations is to find keyplayers that the removing of which would disrupt the network most, which is also known as KPP-NEG problem (Borgatti 2006).

1.2.1 Centrality Measures of Key Stations

1. Degree Centrality

Degree centrality measures a node's direct connectedness with other nodes in a network. Its equation is

given as follows (Freeman 1978; An & Liu 2016):

$$D_i = \sum_j w_{ij} + \sum_j w_{ji} \quad (1)$$

where w_{ij} means whether it is connected or not from node i to node j and vice versa.

2. Betweenness Centrality

Betweenness centrality usually measures a node's brokerage power in a network. Its equation is given as follows (An & Liu 2016) :

$$B_i = \sum_{jk} \frac{g_{jk}^i}{g_{jk}} \quad (2)$$

where g_{jk} is the number of shortest paths between nodes j and k , and g_{jk}^i is the number of those paths that pass node i . When the network consists only isolates which is to say $g_{jk} = 0$, the corresponding betweenness centrality B_i equals 1.

3. Eigenvector Centrality

Eigenvector centrality measures the importance of the neighbors of a give node. Its equation is given as follows (An & Liu 2016):

$$E_i = \frac{1}{\lambda} \sum_j w_{ij} E_j \quad (3)$$

4. Fragmentation Centrality

Fragmentation Centrality defines a network's fragmentation level based on the reciprocals of distances (after removing a set of nodes from the network) (Borgatti 2006):

$$F = 1 - \frac{2 \sum_{j,k \neq i} \frac{1}{d_{jk}}}{d(n-1)(n-2)} \quad (4)$$

Where d_{jk} is distance between node j and k after node i is removed. In the case of $d_{jk} = 0$, the corresponding fragmentation centrality F is 1.

1.2.2 Results of Node Centrality

Table 1: Top 10 ranked stations based on degree value, unweighted betweenness and unweighted eigenvector centrality value, respectively (All values are normalized)

Unweighted Degree Centrality		Unweighted Betweenness Centrality		Unweighted Eigenvector Centrality	
King's Cross St. Pancras	0.0393	Waterloo	0.3358	King's Cross St. Pancras	1.0000
Baker Street	0.0328	Bank	0.3204	Farringdon	0.8022
Embankment	0.0262	Baker Street	0.3195	Euston Square	0.7593
Paddington	0.0262	Westminster	0.3026	Barbican	0.6474
Liverpool Street	0.0262	Green Park	0.2980	Great Portland Street	0.5595
Moorgate	0.0262	Liverpool Street	0.2392	Moorgate	0.5275
Earl's Court	0.0230	Stratford	0.2344	Euston	0.3968
Waterloo	0.0230	Bond Street	0.2335	Baker Street	0.3896
Oxford Circus	0.0197	Mile End	0.2237	Liverpool Street	0.3428
Bank	0.0197	Bethnal Green	0.2145	Edgware Road (C)	0.1814

Table 2: Top 10 ranked stations based on weighted betweenness, weighted fragment centrality value and unweighted fragment centrality value, respectively (All values are normalized)

Weighted Betweenness Centrality		Weighted Fragment Centrality		Unweighted Fragment Centrality	
Baker Street	0.2846	Baker Street	0.9176	Euston	0.9081
Bank	0.2503	Euston	0.9174	King's Cross St. Pancras	0.9079
Oxford Circus	0.2340	Stratford	0.9172	Baker Street	0.9074
King's Cross St. Pancras	0.2328	King's Cross St. Pancras	0.9169	Stratford	0.9070
Bond Street	0.2070	Camden Town	0.9166	Camden Town	0.9069
Whitechapel	0.2049	Leyton	0.9159	Leyton	0.9057
Aldgate East	0.2013	Leytonstone	0.9155	Leytonstone	0.9052
Earl's Court	0.1985	Finchley Road	0.9149	Paddington	0.9048
Liverpool Street	0.1983	Canning Town	0.9146	Finchley Road	0.9045
Stepney Green	0.1866	Wembley Park	0.9145	Canning Town	0.9042

From table 1 and 2, it can be seen that King's Cross St. Pancras has the most direct connections to other stations which indicates it is the most connected traffic hub and the importance is also strengthened by its neighbors as it gets the highest eigenvector value. Besides, its removal leads to fragmentation of the network into two separate parts. As for binary network, Waterloo has the highest betweenness value and Euston gets the highest fragment value. For distance weighted network, these two highest value holders are all Baker Street station. In the next section, this article will focus on distance weighted situation because it relates more to the real world.

1.3 Different Assumptions of Attacks

As Jordan (2008) notes, there are different objectives of terrorist attacks including to destroy infrastructure, to hurt normal people, to create panic, etc. Various aims of attacks lead to changing selection of targets.

1.3.1 Disrupting the Most Connected Stations

In terms of destroying the most connected stations, the measurement is based on degree centrality. Therefore, if terrorists choose to damage one station, King's Cross St. Pancras would be targeted. If two stations are selected, King's Cross St. Pancras and Baker Street would be on the list based on the table 1. The assumption behind this decision is that the more neighbor stations around the target station, the higher potential of more traffic will bring to it in the case that actual traffic data is not available.

In figure 3 and figure 4, after removing King's Cross St. Pancras, the network is separated into two parts with obvious discrete trend of the change of the structure. The adding of Baker Street does not contribute much in terms of change of whole structure (see figure 4 and 6).

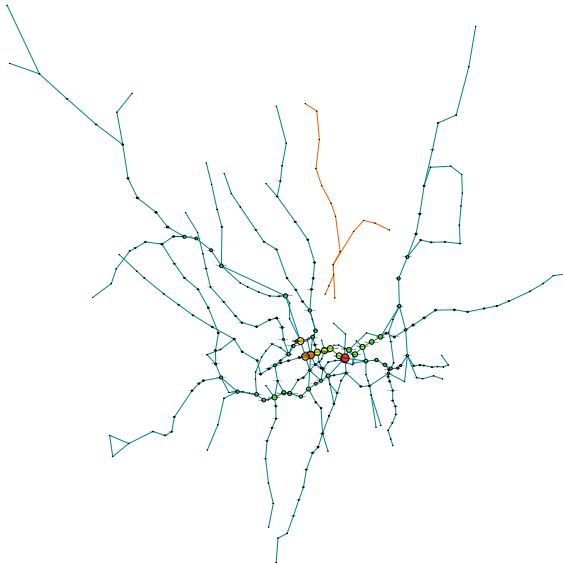


Figure 3: The topographical map of the London underground network after removing King's Cross St. Pancras

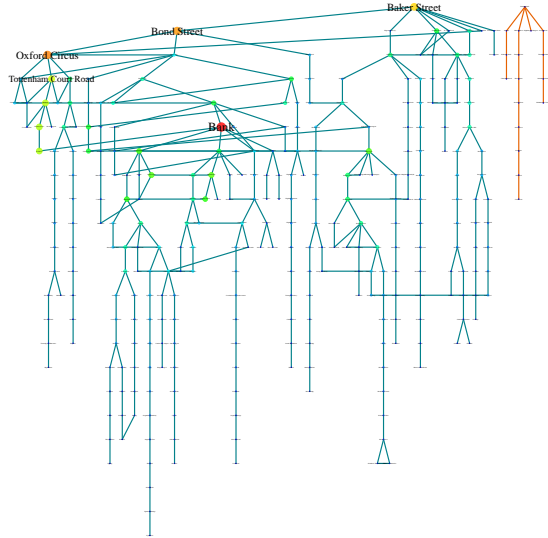


Figure 4: The tree topology of the London underground network after removing King's Cross St. Pancras

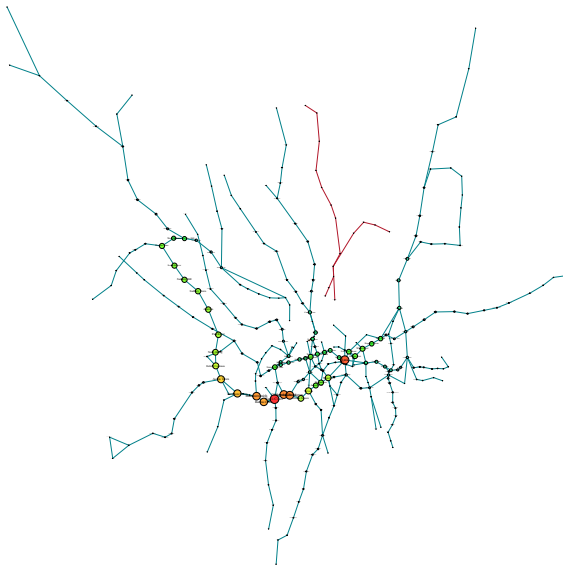


Figure 5: The topographical map of the London underground network after removing King's Cross St. Pancras and Baker Street

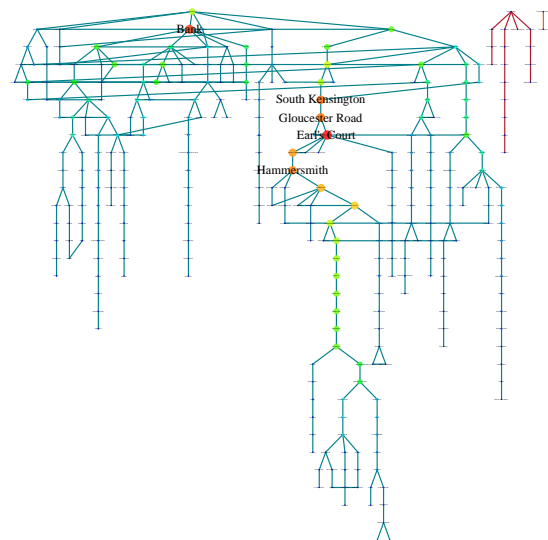


Figure 6: The tree topology of the London underground network after removing King's Cross St. Pancras and Baker Street

1.3.2 Blocking the Most Pathways

If terrorists' aim is to disrupt stations that allows the most shortest paths to go through, then Baker Street would be the No.1 target. If two stations are chosen, how to choose the combination is difficult. The first option is the top 2 stations on the list: Baker Street and Bank. However, if they choose to disrupt sequentially, the station that has the highest betweenness centrality would change to Earl's Court after destroying Baker Street. A major change of the network the removal of Baker Street causes is the move of high betweenness area from center London to the southern part. The combination of two top ranked stations (see figure 7 and 8) turns out adding very weak influences compared to just removing one station. The choose of removing two stations sequentially gets better results by creating an addition components (See figure 11 and 12).

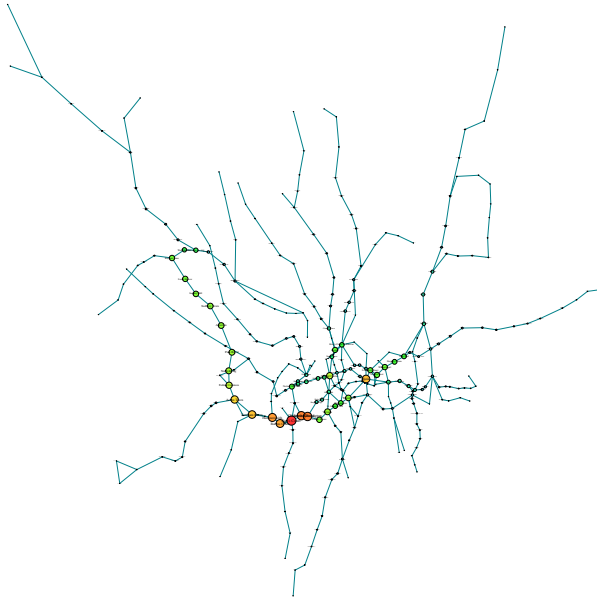


Figure 7: The topographical map of the London underground network removing Baker Street

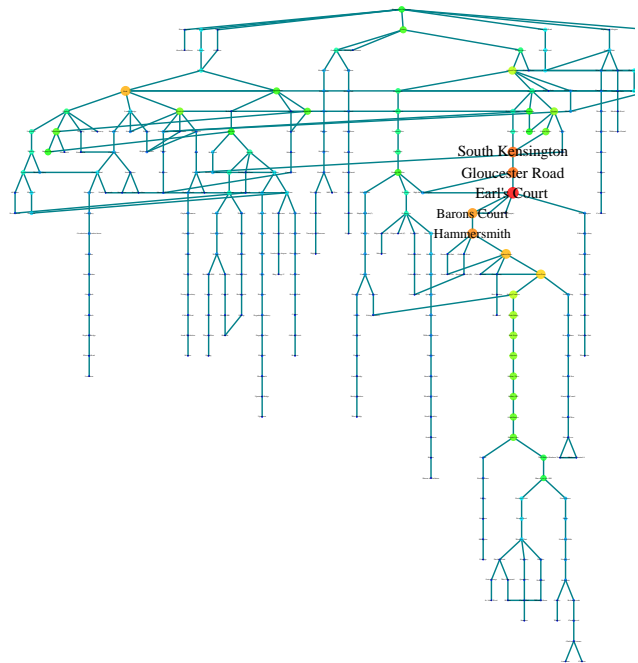


Figure 8: The tree topology of the London underground network after removing Baker Street

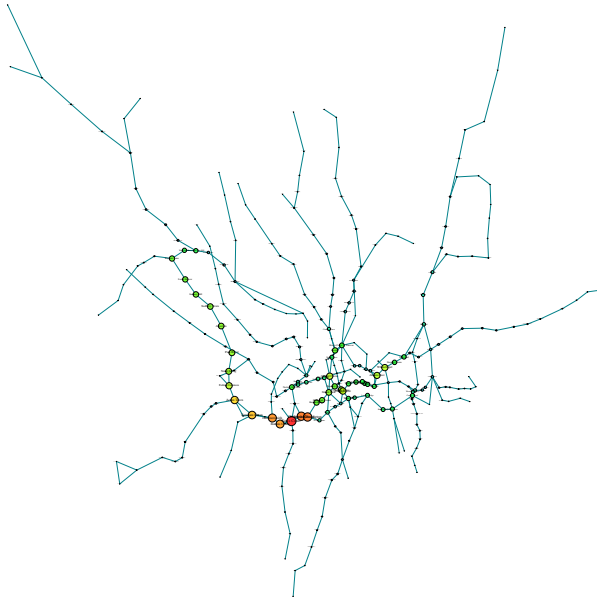


Figure 9: The topographical map of the London underground network removing Baker Street and Bank

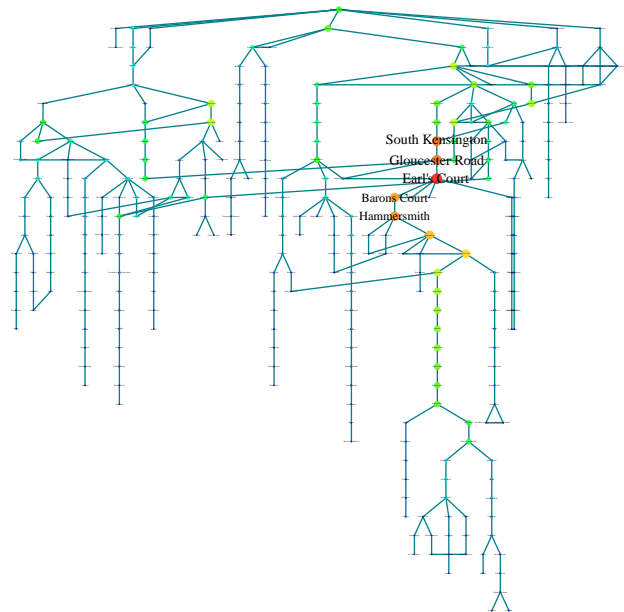


Figure 10: The tree topology of the London underground network after removing Baker Street and Bank

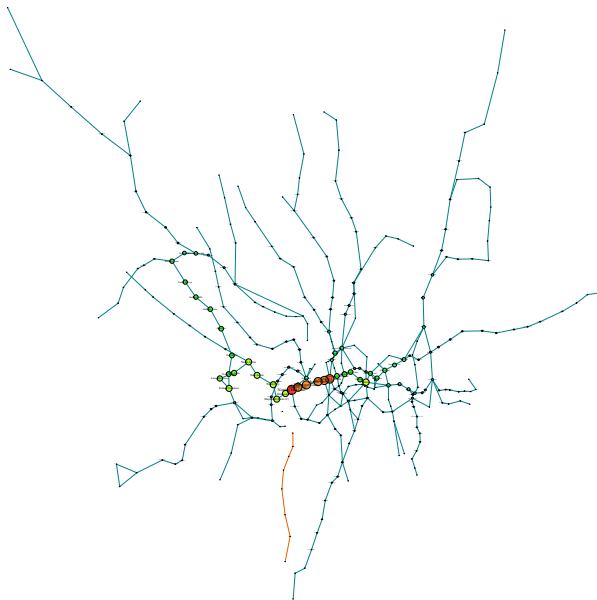


Figure 11: The topographical map of the London underground network after removing Baker Street and Earl's Court

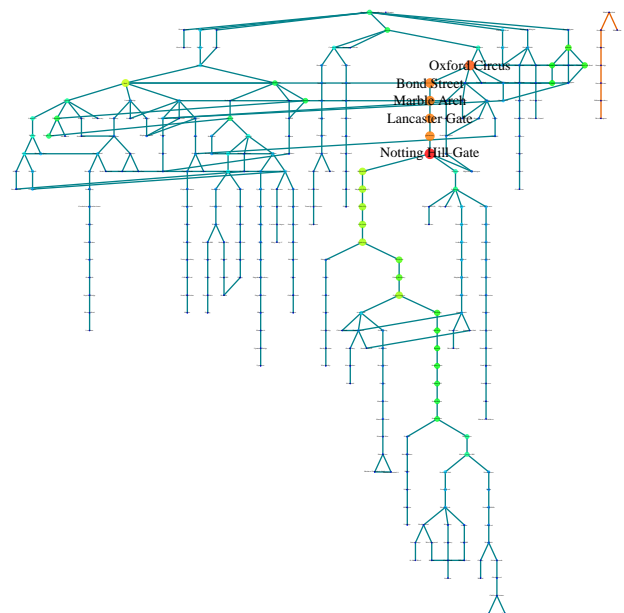


Figure 12: The tree topology of the London underground network after removing Baker Street and Earl's Court

1.3.3 Fragmenting the Whole Tube Network

If terrorists plan to fragment the whole tube system as much as possible, several aspects are required to be taken into consideration. First criteria is the ability of node to divided the network into parts. Second, it is better to disconnect the network into two rough equal size; Third, the internal structure of each components is also important. The more discrete it is, the better. Therefore, here fragement centrality is introduced as the measure by calculating the sum of reciprocals of distances (see table 2).

The first choice is Baker Street which has already been analysed above. When it comes to two targets, there are 46665 options to pick up two different nodes from all the 306 nodes. Kpset method is applied by using keyplayer package in R. It turns out that the combination of Baker Street and Ealing Common is the optimal one.

If three stations are chosen, the former two stations together with King's Cross St. Pancras is the best result that fragments the whole underground network most severely. From figure 13 and 14, it can be seen that one critical change compared to the previous two-station selection cases is the size of the smaller components gets bigger, thus making the network more discrete.

The last model removes three stations dividing the network into four sub-networks, the structure of which is the most discrete (see figure 15 and 16).

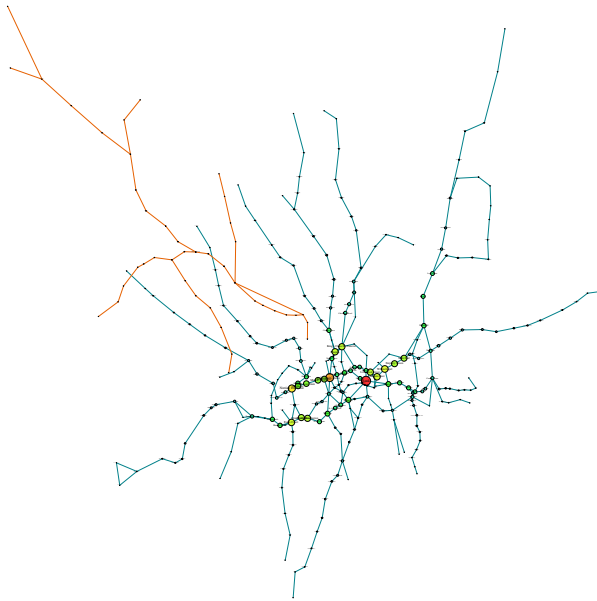


Figure 13: The topographical map of the London underground network after removing Baker Street and Ealing Common

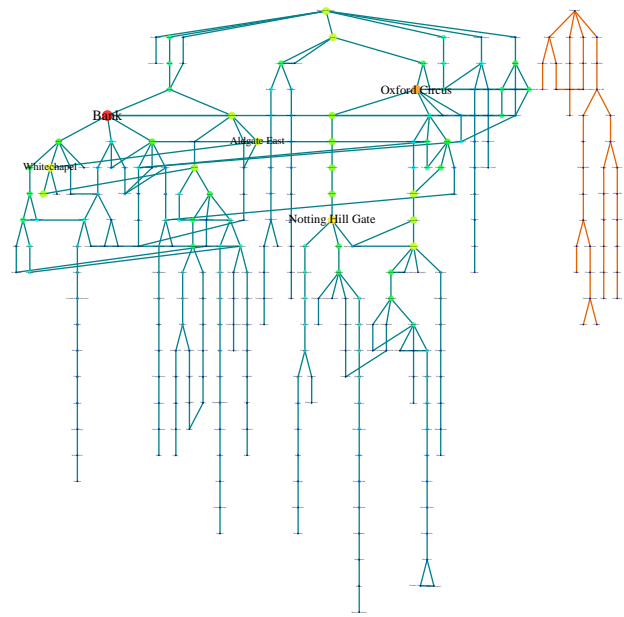


Figure 14: The tree topology of the London underground network after removing Baker Street and Ealing Common



Figure 15: The topographical map of the London underground network after removing Baker Street, Ealing Common and King's Cross St. Pancras

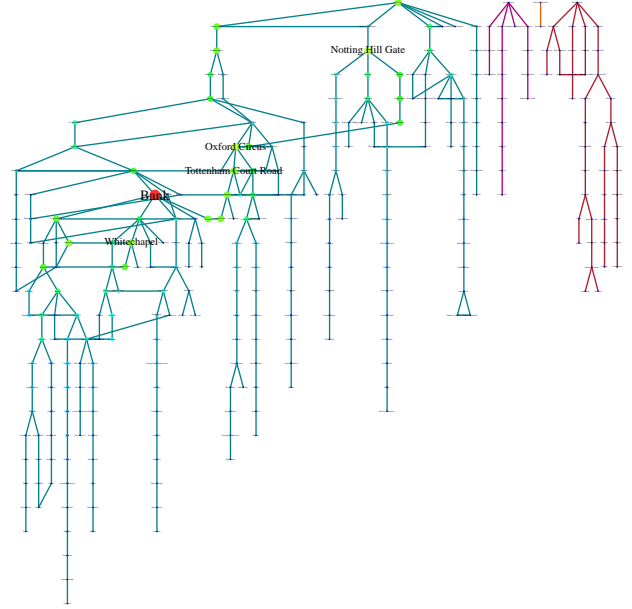


Figure 16: The tree topology of the London underground network after removing Baker Street, Ealing Common and King's Cross St. Pancras

1.4 Final Results and Conclusion

Romoved Nodes	Degree	Betweenness	Eigenvector
King's Cross St. Pancras	0.02431	0.35867	0.98432
King's Cross St. Pancras and Baker Street	0.01798	0.34598	0.98471
Baker Street	0.03085	0.31748	0.98157
Baker Street and Bank	0.03105	0.31540	0.98498
Baker Street and Earl's Court	0.03107	0.39435	0.98151
Baker Street Ealing Common	0.03101	0.21971	0.98151
Baker Street Ealing Common King's Cross St. Pancras	0.01810	0.27022	0.98466

Table 3: Network Centralization Scores of Each Scenario

Table 3 shows the results of different network centrality of all the mentioned models. If we choose network betweenness centrality as the measurement of final impact, we can see that the optimal sets of nodes that disrupts the whole system most is the combination of Baker Street and Ealing Common, which is selected based on node fragment centrality.

In conclusion, based on given evidences (e.g. underground lines, topology of tube stations, distances between pairs of stations) and certain assumptions the above attack strategies were given. In the future, as more information is provided, such as the connection to other kind of transport systems and passenger flux data, these strategies would be optimized. For example, Jordan (2008) indicates that the Heathrow station would become an important target if the global air transport system is taken into consideration.

2 Agent Based Models

2.1 Introduction

Agent-Based Models (ABM) and Cellular Automata Models (CA) are two commonly used urban models for complex spatial system. ABM generally simulates dynamic processes of individuals or objects while CA is based on finer spatial scales in regards to physical morphology (Batty 1997). Therefore, ABM is more suitable for cases containing explicit and dynamic activities or behaviors while CA deals with issues in particular related to land development, built-environment, urban growth, etc.

This article aims at assessing the sensitivity of the Epidemic ABM models to varying parameter settings. This model simulates the spread of illness among a certain amount of people. Parameters like the initial population, number of infected individuals, immune chance and recovery chance are examined.

2.2 Experiments

2.2.1 Change of Initial Number of Infected

Here we set initial number of infected people to 1, 10, 50, 100, 280, respectively (see figure 17). It will then randomly infect this number of people if they do not have any immunity. Thus, as ticks increase, the average percentage of infected people go up (see figure 18). The line range represents 95% confidence interval for each time of all repetitions.

The growing trend tends to be stable afterwards. The higher initial infected number, the less time needs to reach stable infected rate level. The minimum time needed to run the simulation is around 100 times through exploratory analysis.

When there is only 1 infected people at the start the stable average value seems much lower than the others. However, it is due to the many cases that this infected individual recover at the first time leaving the world with no source of infection. If we ignore this phenomena, it can be conclude that under the conditions with 10% recovery chance and 10% immune chance, no matter how many people are initially infected (at least works for range [10:280]), the infected rate will always get stable around 45%. Therefore, in this case, infected rate is not sensitive to the initial number of infected.

To find the minimum number of runs, experiments are set up with 500 repetitions and 10 number of the initial infected while other settings remain the same as before (see figure 19). The black line represents the average of additive infected rate within a certain amount of runs. The red line range indicated 95% confidence interval. Figure 20 shows the corresponding standard errors after certain runs of experiments. If we set minimum standard error to 0.1 as the evidence of getting statistical meaningful results, therefore the minimum runs is supposed to be 103. However, if two of all samples are compared by Welch Two Sample t-test assuming that it's normally distributed, the mean of the results of the first two runs is significantly same (with p value equals to $0.34 > 0.05$) from that of the first three runs. In this case, at least two runs are required to get significant results. Similar results for more runs and the t test result p value for the comparison among 103 runs and 104 runs are around 0.9. In conclusion, the minimum runs differs as the measurement of significantly meaningful differs.

Parameters	
num-infect	1 10 50 100 280
recovery-chance	10
initial-people	300
immune-chance	10
Repetitions	15
Time limit	500

Figure 17: Experiment Setup

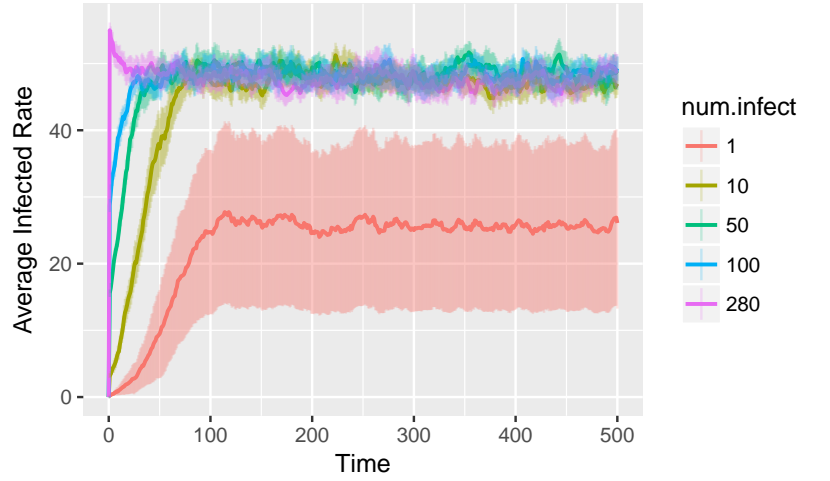


Figure 18: Time vs Average Infected Rate

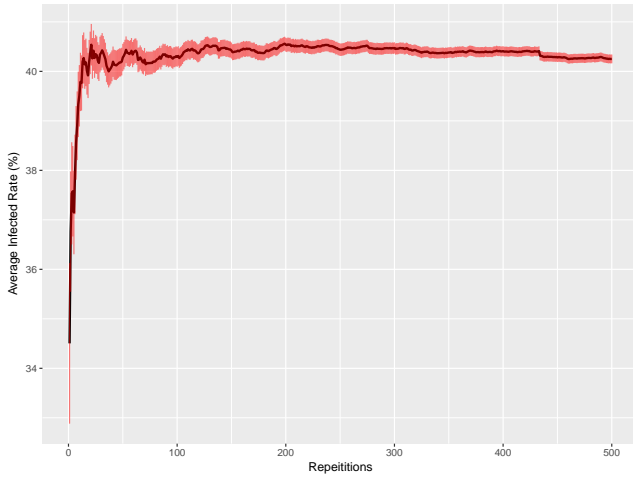


Figure 19: Runs vs Average Infected Rate

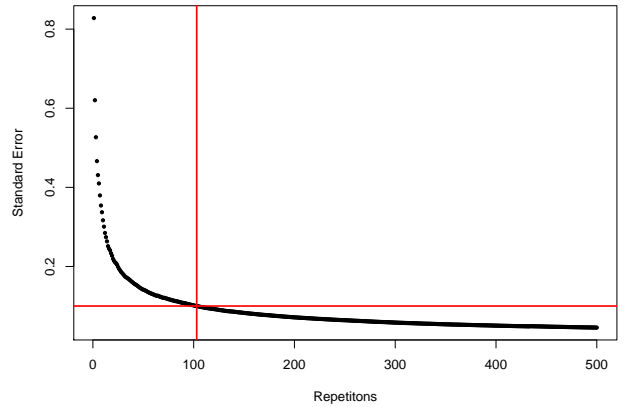


Figure 20: Runs vs Std. Error Infected Rate

2.2.2 Multiple Parameter Settings

By referring to the previous results, we set repetitions to 100 and time limit to 150. We change multi parameters (see figure 21) at the same time to see the combined effect on infected rate with extreme scenarios like 3000 population and one percent recovery rate and immune rate. From figure 22, it can be seen that recovery chance has much stronger effect on the change of infected rate than immune chance, which indicates that improvement on recovery chance has better efficiency on stopping the spread of diseases. In the case with 3000 initial population, what ever the changes of recover rate and immune rate are, the reached stable average infected rates are always quite high, more than 75 percent, which indicates the importance of avoiding high public population density to control diseases spread. Besides, the confidence interval of these curves is very narrow and can be hard tell in the figure.

Parameters	
num-infect	10
recovery-chance	1 5 10 20
initial-people	300 500 3000
immune-chance	1 5 10 20
Repetitions	100
Time limit	150

Figure 21: Experiment Setup

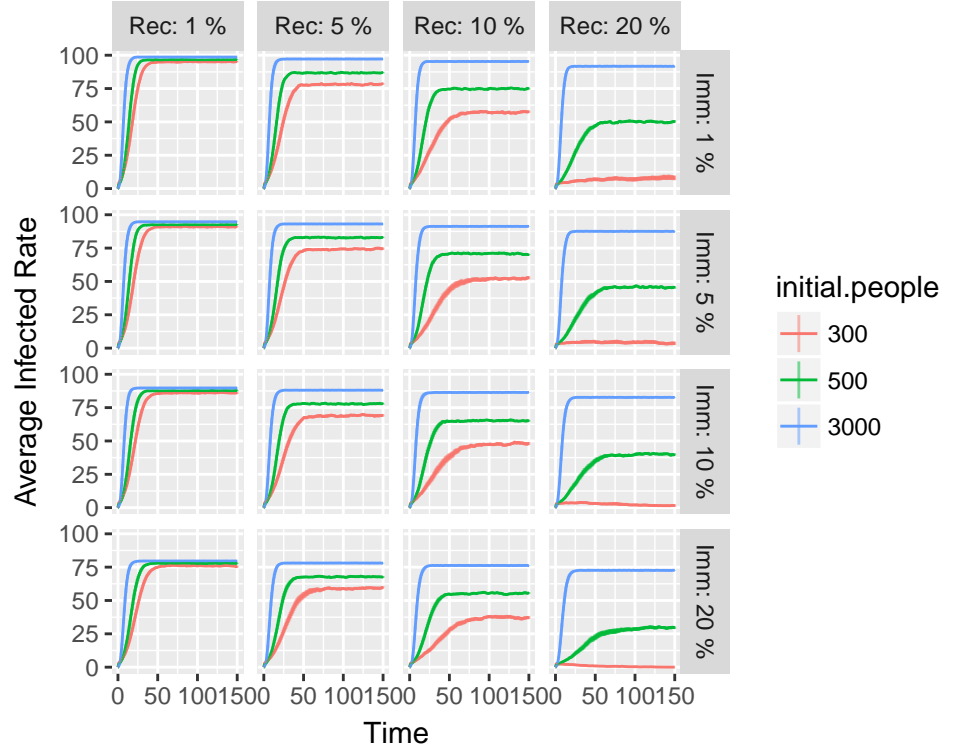


Figure 22: Multiple Parameter Settings

2.3 Explore of Determining Minimum Runs

In conclusion, a statistical exploratory analysis of the outcomes of ABMs especially in the form of means and standard deviations is usually necessary. Since the shape of outcomes' distribution can not be accessed previously, the time or runs at which mean and standard errors of the result reach a stable is important. However, different criteria for deciding the moment of stability suffers and falls to make a wise selection (Lee *et al.* 2015). In the former subsection, we use standard error as an indicator to check the point of stability. Here, Lee *et al.* introduced a scaling one which is the ratio of standard deviation of a sample (σ) to its mean (μ): the coefficient of variation (C_v). Therefore, the method can be improved by examining the C_v of different sized samples. When the sample size meet the conditions that the difference between two nearby C_v continuously falls and under a criteria E , it is considered as the minimum runs. These C_v stability points are contained for all interested ABM outcomes and thus the minimum number of runs is the maximum number of these C_v points.

3 Spatial Interaction Models

First, in section 3.1, this article briefly introduces four core members of spatial interaction models family, namely unconstrained models (subsection 3.1.1), production constrained models (subsection 3.1.2), attraction constrained models (subsection 3.1.3) and doubly constrained model (subsection 3.1.4). In each subsection, it starts with a brief introduction of each model's equation and assesses the role of the constraints to the model. Followed by experiments on testing the effect of varying parameters and inputs on spatial interactions. The model is based on the commuting data among 33 boroughs in London.

3.1 Initial Spatial Interaction Models

Family of spatial interaction models were initially put forward by Wilson 1971. In this article the basic gravity model is used (Dennett 2012, Oshan 2016 and Rodrigue *et al.* 2016).

3.1.1 Unconstrained model

The basic unconstrained model is as follows:

$$T_{ij} = KO_i D_j c_{ij}^{-\beta} \quad (5)$$

where K is a proportionality constant related to the rate of the event which can be calculated under the constrain that total number of trips T holds unchanged overtime. Therefore, K equals to:

$$K = \frac{T}{\sum_{ij} O_i D_j c_{ij}^{-\beta}} \quad (6)$$

The β relates to travel cost friction. The bigger it is, the less trips are generalised and the weaker spatial interaction of the model is. In addition, the α as attractiveness parameter and the λ as emissiveness parameter are added:

$$T_{ij} = KO_i^\lambda D_j^\alpha c_{ij}^{-\beta} \quad (7)$$

In this experiment circumstances, two attributes of origin i noted as O are introduced, namely origin populations and origin average salaries. Attributes of destination j noted as D contains destination populations and destination average salaries. Besides, cost friction c is represented by distance. The only constrain is the fixed sum of all trips generated, making the model flexible to be altered by different combinations of the parameters sets to some extent. However, the drawback is that it leads to a low accuracy of the final results.

In order to calibrate parameters, logarithm transformation is one of the methods that make the process easier. Thus, it ends up with the following equation, and followed by poisson regression:

$$\log T_{ij} = K + \lambda \log O_i + \alpha \log D_j - \beta \log c_{ij} \quad (8)$$

1. Effects of changing K , β , λ and α

Figure 23 shows the scatter plot of observed and estimated travel flows. It is very obvious that although these parameters change only a bit, the target increases exponentially. We examined both positive and negative value as well as 0 for K value. It is found that if K increases or decreases a certain value, estimated travel flows will correspondingly multiple e to the power of this value times.

Second, from figure 24 it can be seen that, the β changes the influence of distance on spatial interactions. The higher value the β is, the steeper decrease spatial interactions have which means the more important β is to the change of travel flows. Figure 25 and 26 show the similar positive effect of λ and α on spatial interactions. The bigger they are means the more influent origin or destination attributes are to the result.

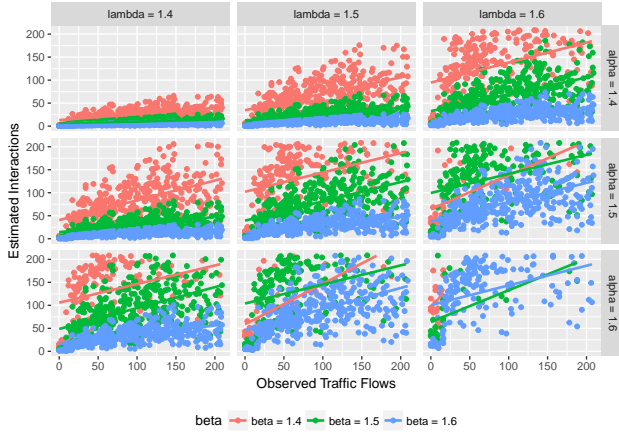
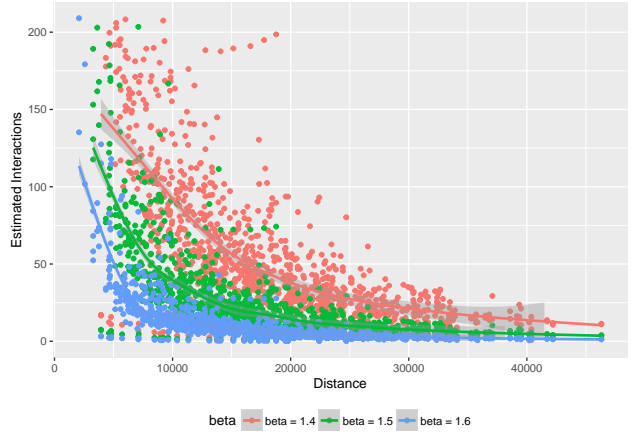
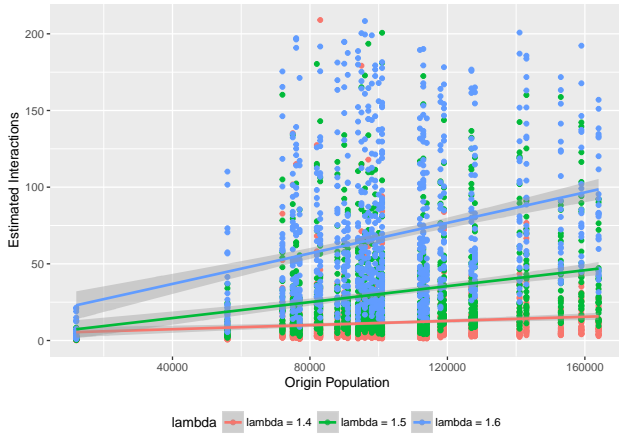


Figure 23: Observation vs Estimation

Figure 24: Effects of β on spatial interactionsFigure 25: Effects of λ on spatial interactionsFigure 26: Effects of α on spatial interactions

2. Effects of Changing Inputs

In the case that all of a sudden a number of well paid jobs appeared in Brent and the average salary doubled, we updated the average salary data, reuse the model and calculate results again. Finally, we find that trips from boroughs to Brent all increased to around 2 times than before while other flows with increase no more than 1.15 times. From figure 27 and 28 , it can be seen that the most popular destination based on increasing flows is undoubtedly Brent while City of London, Kensington and Chelsea, and Richmond upon Thames become top three destinations with severe depopulation, which was former top three destinations with high average salary. This is all because the doubled average salary makes Brent take replace of Kensington and Chelsea becoming the second highest ranking destination.



Figure 27: Change of Travel Flows After Brent's Job Increase. Blue Line: Less Than 0.92 Times Trips Compared to Before; Red Line: More Than 2 Times Trips Compared to Before

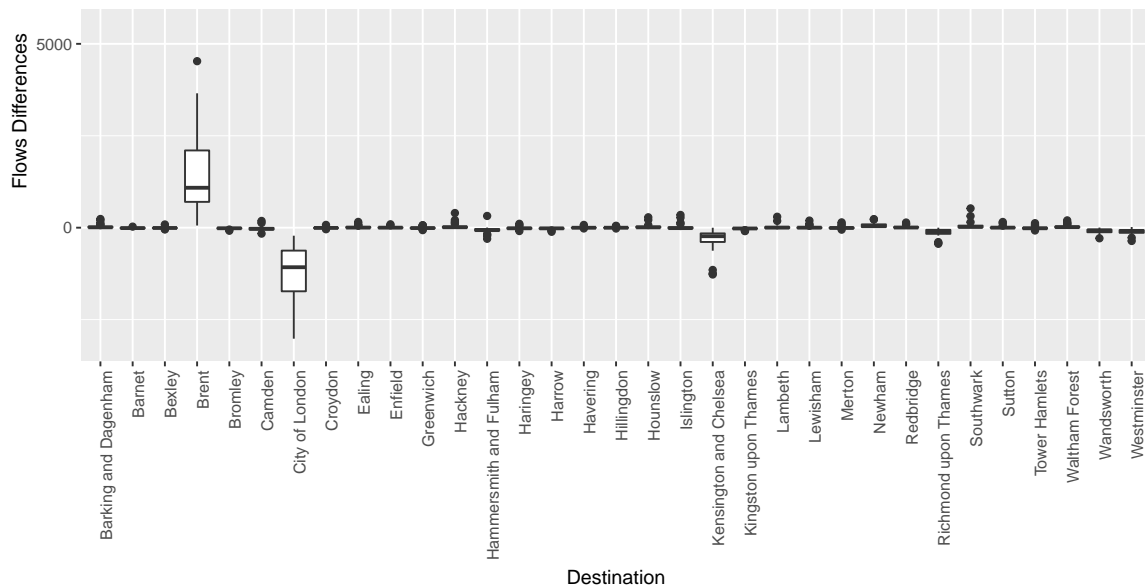


Figure 28: Destinations' Changes in Travel Flows After Brent's Job Increase - Unconstrained Model

3.1.2 Production Constrained Model

In production constrained model, the total number of journeys starts within a given area, which can be stated as :

$$O_i = \sum_j T_{ij} \quad (9)$$

Thus the former model is rewritten out as:

$$T_{ij} = A_i O_i D_j^\alpha c_{ij}^{-\beta} \quad (10)$$

where

$$A_i = \frac{1}{\sum_j D_j c_{ij}^{-\beta}} \quad (11)$$

The main difference between production constrained model and unconstrained model is that in origin constrained model, the sum of all flows for each origin is fixed. This condition to a larger extent improves the accuracy of the model. To make a comparison, we adopt the same scenario which is doubled salary in Brent to see whether there is any difference. The situation for destination estimates' change are similar as unconstrained model (see 28). However, for origins, the differences are clear (see figure 29) with Ealing, Harrow and Barnet becoming the top three ranked origins which lose the most trips while the result for unconstrained model is not clear (see figure 30).

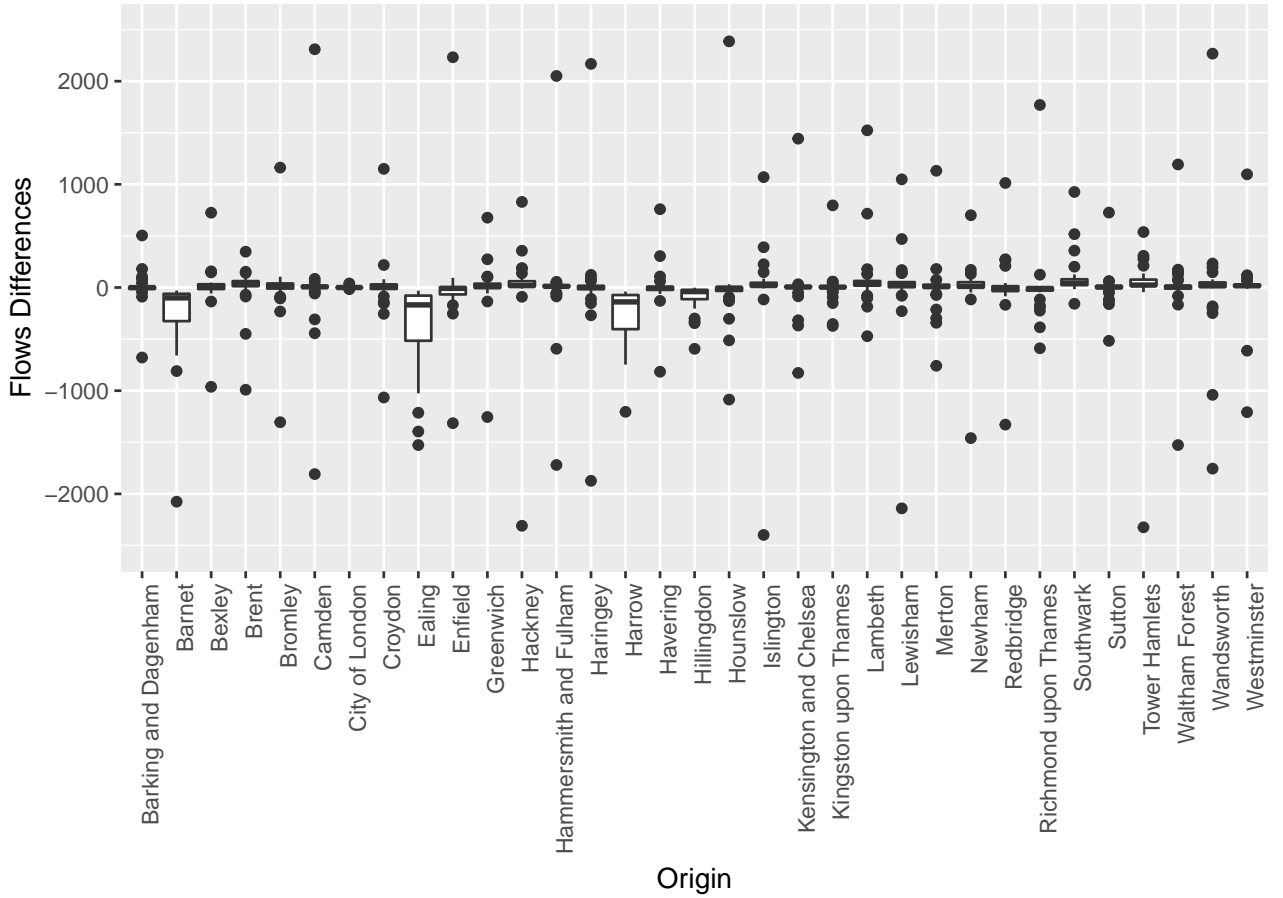


Figure 29: Origins' Changes in Travel Flows After Brent's Job Increase - Production Constrained Model

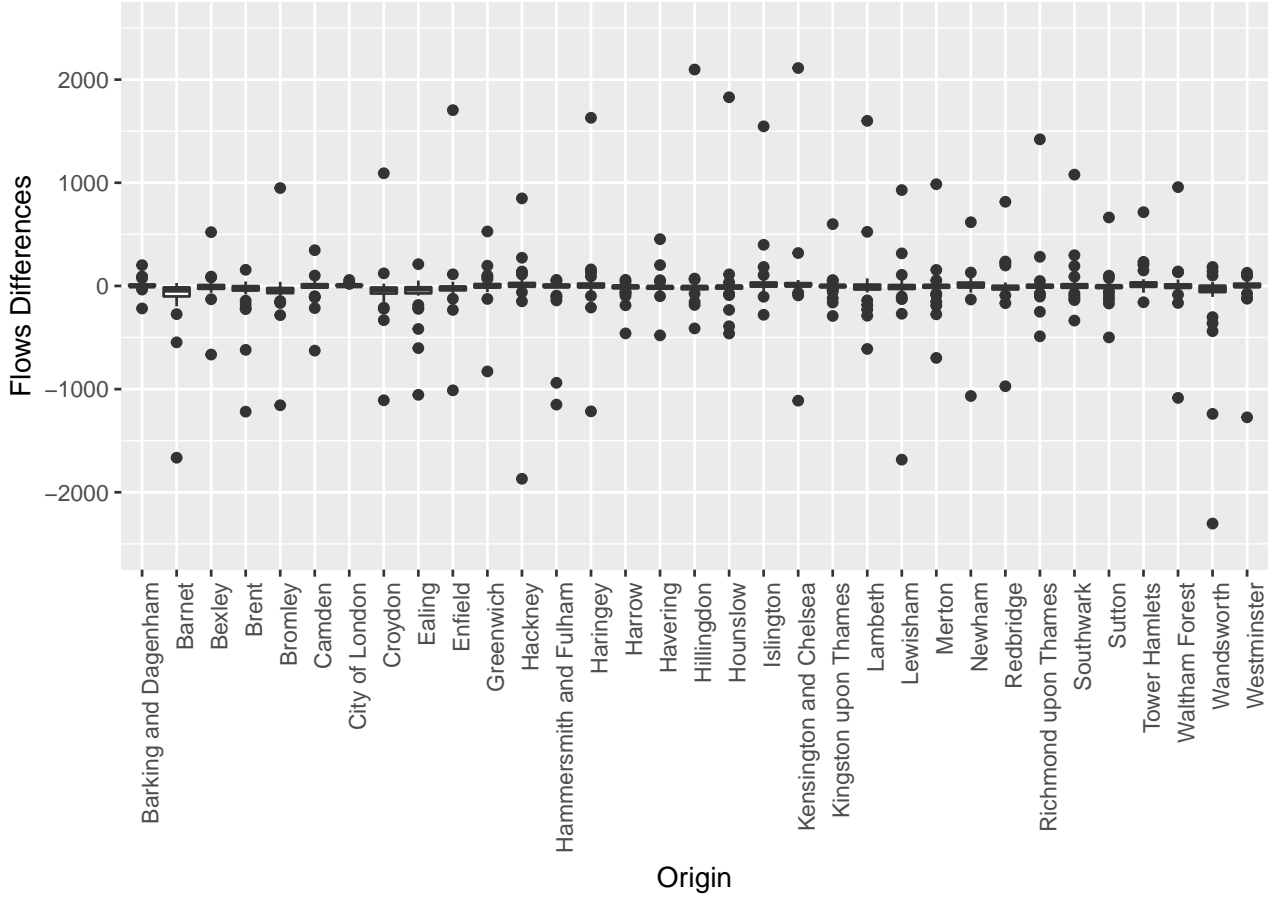


Figure 30: Origins' Changes in Travel Flows After Brent's Job Increase - Unconstrained Model

3.1.3 Attraction Constrained Model

In attraction constrained model, the total number of journeys ends within a given area, which can be stated as :

$$D_j = \sum_i T_{ij} \quad (12)$$

The formula is structured as :

$$T_{ij} = B_j D_j O_i^\lambda c_{ij}^{-\beta} \quad (13)$$

where

$$B_j = \frac{1}{\sum_i O_i^\lambda c_{ij}^{-\beta}} \quad (14)$$

The attraction constrained model has a fixed sum of all flows for each destination. It is used under scenarios like new construction or development is finished and would attract a certain number of people to work here which indicates increase in its inbound flows.

3.1.4 Doubly Constrained Model

In doubly constrained model, the constraints are :

$$\begin{aligned} O_i &= \sum_j T_{ij} \\ D_j &= \sum_i T_{ij} \end{aligned} \quad (15)$$

Thus the former model is rewritten out as:

$$T_{ij} = A_i O_i B_j D_j c_{ij}^{-\beta} \quad (16)$$

where

$$\begin{aligned} A_i &= \frac{1}{\sum_j B_j D_j c_{ij}^{-\beta}} \\ B_j &= \frac{1}{\sum_i A_i O_i c_{ij}^{-\beta}} \end{aligned} \quad (17)$$

In this model, total numbers of both trip production and trip attraction are constrained, which makes its implementation in practicals complex and hard to process. However, it usually has the best performance among all the mentioned four models with the highest R square value and smallest RMSE value.

3.2 Conclusion

As Rodrigue *et al.* (2016) indicates, the process of finding a set of ideal and accurate parameters are time consuming but useful. Once a spatial interaction model for a city or region has been established, it can be used for simulation and prediction. Doubly constrained models always have the best results however it is not always easy to get reliable observed data. Therefore, single constrained models are also needed and helpful.

References

1. An, W. & Liu, Y.-H. keyplayer: An R Package for Locating Key Players in Social Networks. *R Journal* **8** (2016).
2. Batty, M. Cellular automata and urban form: a primer. *Journal of the American Planning Association* **63**, 266–274 (1997).
3. Borgatti, S. P. Identifying sets of key players in a social network. *Computational & Mathematical Organization Theory* **12**, 21–34 (2006).
4. Dennett, A. Estimating flows between geographical locations: ‘get me started in’ spatial interaction modelling. *Technical report* (2012).
5. Fotheringham, A. S. & O’Kelly, M. E. *Spatial interaction models: formulations and applications* (Kluwer Academic Publishers Dordrecht, 1989).
6. Freeman, L. C. Centrality in social networks conceptual clarification. *Social networks* **1**, 215–239 (1978).
7. Jordan, F. Predicting target selection by terrorists: A network analysis of the 2005 London underground attacks. *International journal of critical infrastructures* **4**, 206–214 (2008).
8. Lee, J.-S. *et al.* The complexities of agent-based modeling output analysis. *Journal of Artificial Societies and Social Simulation* **18**, 4 (2015).
9. Oshan, T. M. A primer for working with the Spatial Interaction modeling (SpInt) module in the python spatial analysis library (PySAL). *Region* **3**, 11–23 (2016).
10. Rodrigue, J.-P., Comtois, C. & Slack, B. *The geography of transport systems* (Routledge, 2016).
11. Wilson, A. G. A family of spatial interaction models, and associated developments. *Environment and Planning A* **3**, 1–32 (1971).