



Smart Cities

Session III: Networks, Flows and Big Data

Lecture 5: *Material and Electronic Networks 2: People Movements*

Michael Batty

m.batty@ucl.ac.uk

 @jmichaelbatty

<http://www.spatialcomplexcity.info/>

<http://www.casa.ucl.ac.uk/>

Outline of the Lecture

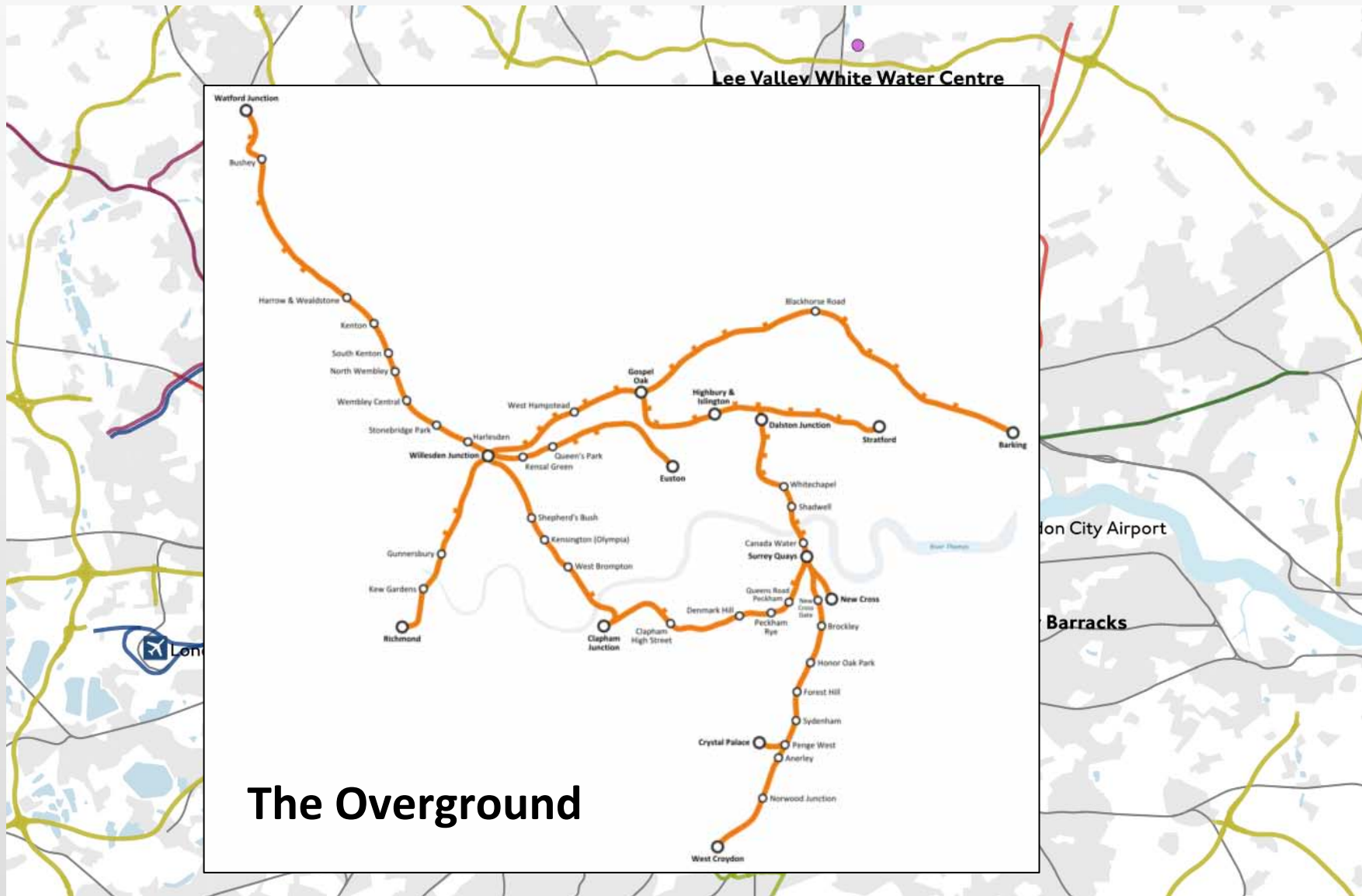
We have dealt with the material that is greyed out

1. Context: New Data for and from the Smart City
2. Multimodal Flows: An Analytical Gap
3. Why London?
4. A Partial View of London's Network
5. Three Approaches, Three Problems –
6. Problem 1: Flows on the tube lines – by trains
7. Problem 2: Representing Networks
8. Problem 3: The 'Shortest'-Path Problem

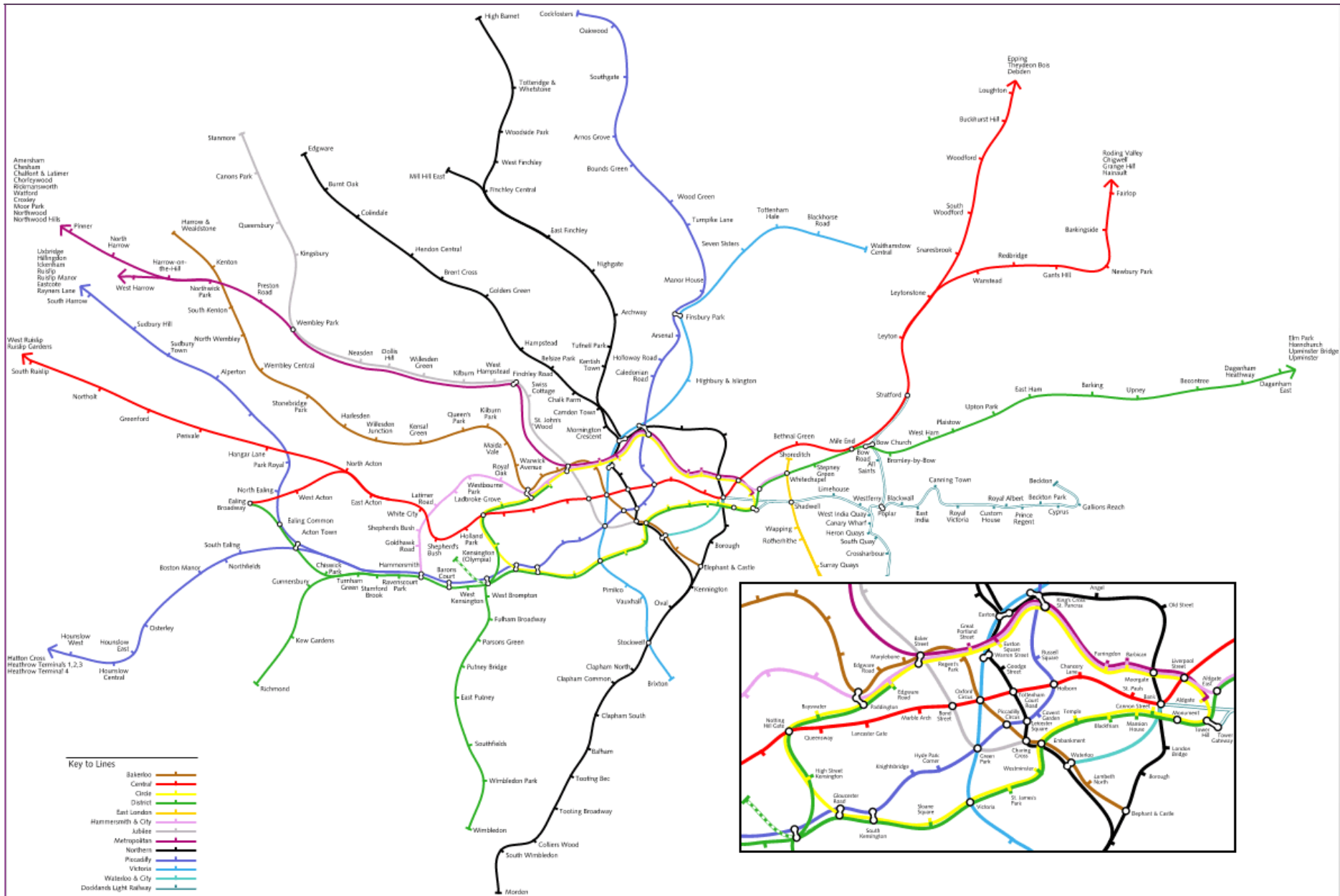
I want to introduce a little more background on the automated scheme in London that we are discussing – some history first and then a map of the network

It is quite confusing because so many networks overlap so we should be clear about the core network – the underground which has 270 stations; the Docklands Light railway has 45; the overground has 83 but this extends outside of Greater London: then national rail is more complicated because it dovetails with overground and underground stations also coincide but are not the same as national rail stations.

We are mainly going to deal here with the greater London networks but the touch card data is for a much bigger area and also for the bus system and it can also be related to national rail.

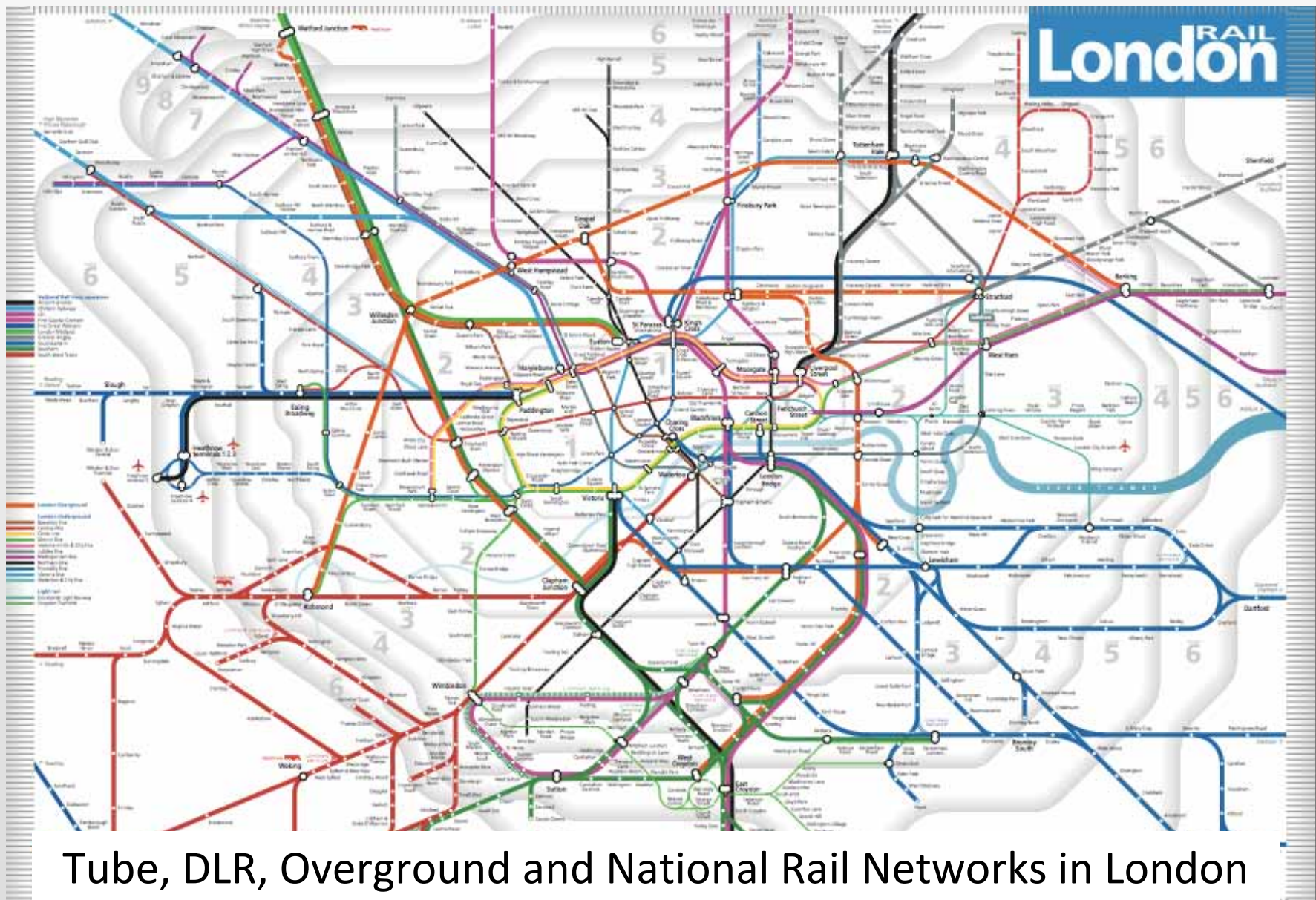


The Overground



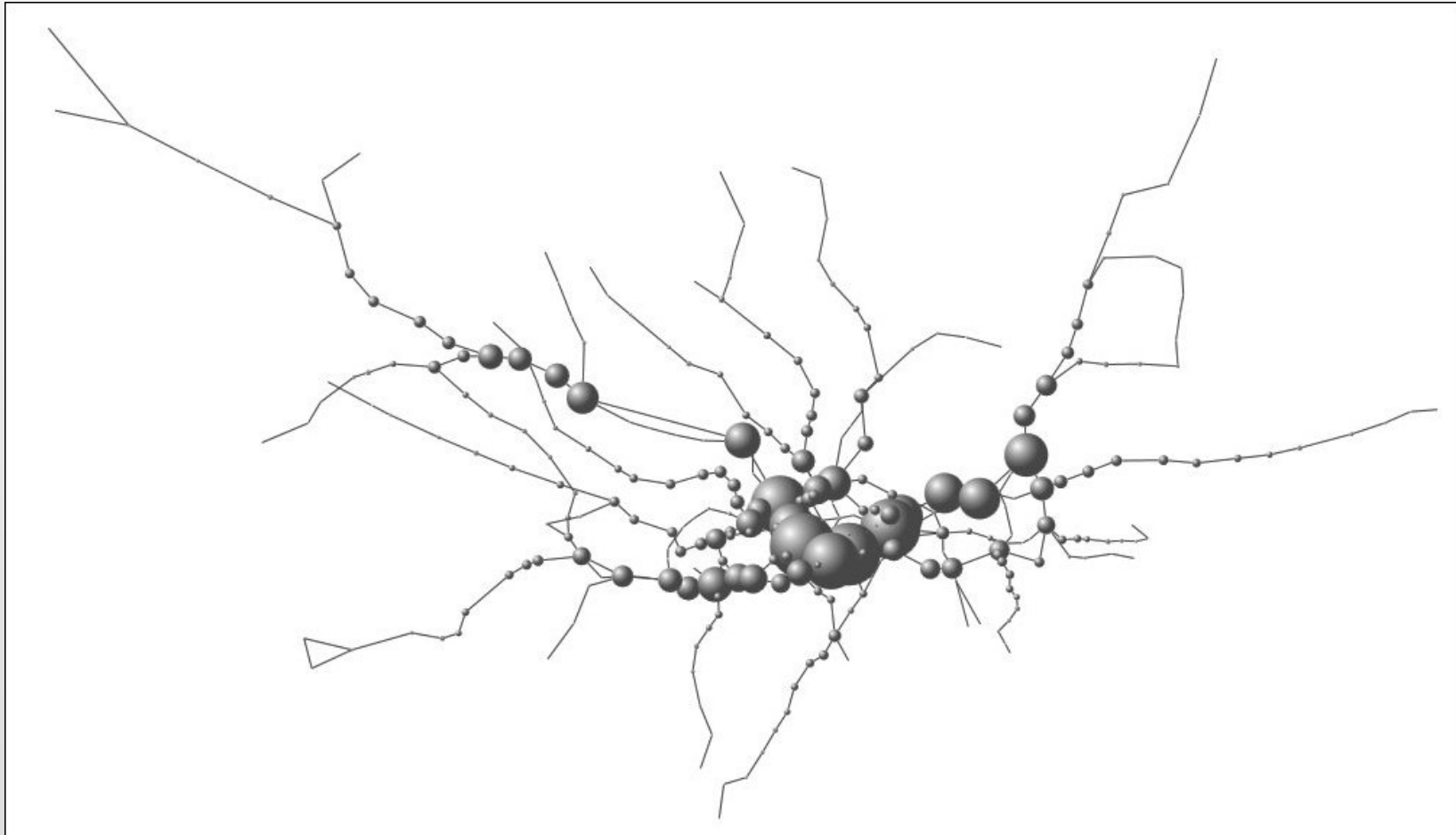
Centre for Advanced Spatial Analysis





Tube, DLR, Overground and National Rail Networks in London

The simplest network for the first problem, based on an analysis of the network, not the flows: shown below are the degrees



Problem 2: Representing Networks

We use standard graph algebra to represent the network where we define three indices of centrality

Degrees of the graph

$$\left. \begin{aligned} \sigma_i &= \sum_j a_{ij} \\ \sigma_j &= \sum_i a_{ij} \end{aligned} \right\} \sigma = \sum_i \sigma_i = \sum_j \sigma_j = \sum_i \sum_j a_{ij}$$

Betweenness Centrality

$$C_k = \sum_i \sum_j \frac{\sigma_{ikj}}{\sigma_{ij}}$$

Closeness Centrality

$$L_i = KD_i^{-1} = K \left(\sum_j d_{ij} \right)^{-1}$$

Representing Flows

Trip Volume

Entries and Exits

$$\left. \begin{aligned} T_i &= \sum_j T_{ij} \\ T_j &= \sum_i T_{ij} \end{aligned} \right\} T = \sum_i T_i = \sum_j T_j = \sum_i \sum_j T_{ij}$$

Changes in

Trip Volumes

$$\left. \begin{aligned} \Delta_i &= T_i - T'_i \\ \Delta_j &= T_j - T'_j \end{aligned} \right\} \sum_i \Delta_i = \sum_j \Delta_j = 0$$

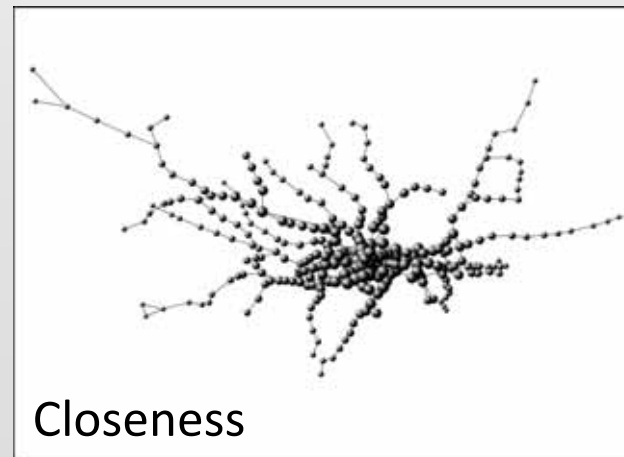
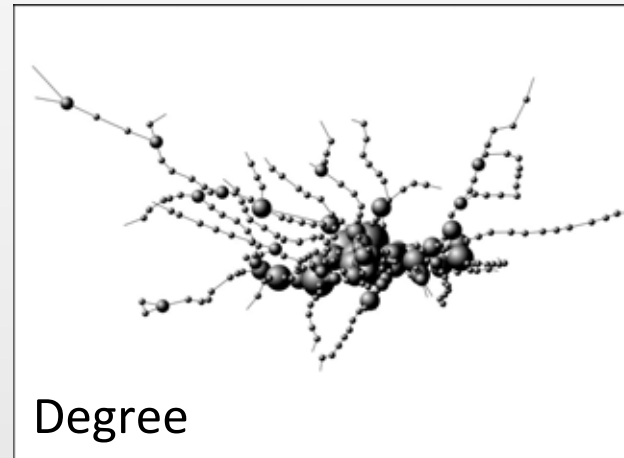
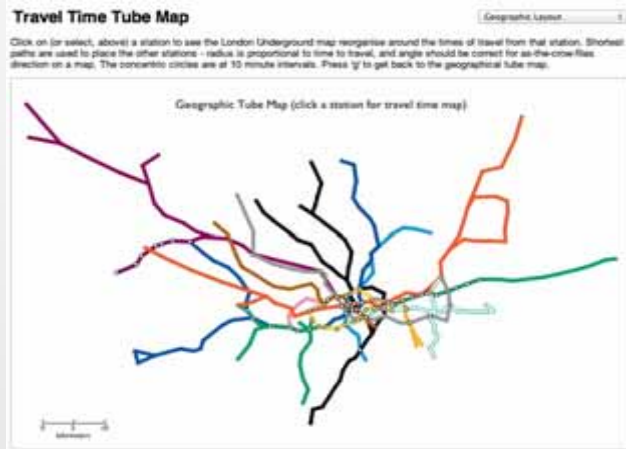
Weighted Betweenness
Centrality

$$p_{ijk} = \frac{\sigma_{ikj}}{\sigma_{ij}} = \frac{\sigma_{ikj}}{\sum_{\ell} \sigma_{i\ell j}}, \quad \sum_k p_{ijk} = 1$$

$$\tilde{C}_k = \sum_i \sum_j T_{ij} p_{ikj} = \sum_i \sum_j T_{ij} \frac{\sigma_{ikj}}{\sigma_{ij}}$$

A Preliminary Analysis (1)

The Minimal Tube Network and the Three Centrality Indices



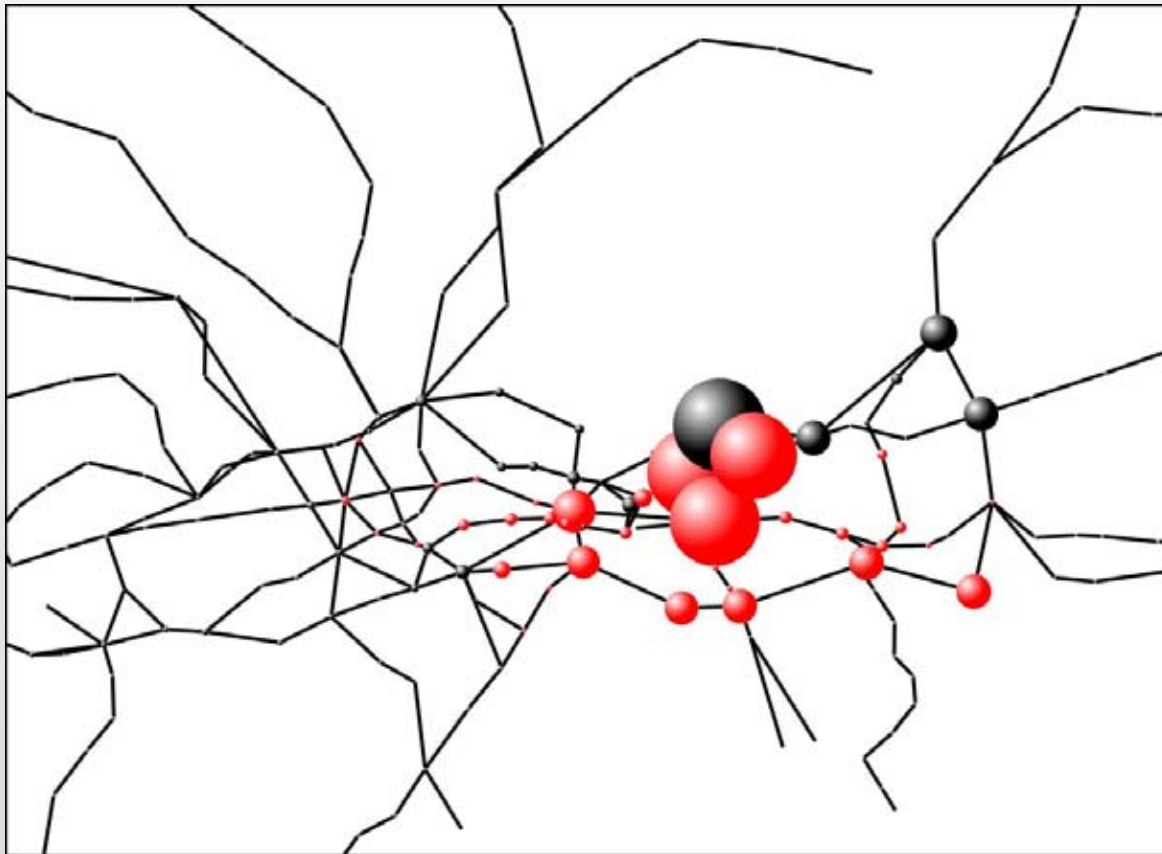
A Preliminary Analysis (2)

- Top Stations
- By Centrality

Station	d_i	Station	\hat{C}_i	Station	\hat{L}_i
Baker Street	7	Green Park	16399	Green Park	2.137
King's Cross	7	Waterloo	15644	Westminster	2.107
Bank	6	Bank	15008	Bond Street	2.101
Earl's Court	6	Baker Street	14441	Oxford Circus	2.089
Green Park	6	Westminster	14139	Waterloo	2.089
Oxford Circus	6	Bond Street	11429	Bank	2.074
Waterloo	6	Liverpool Street	11186	Baker Street	2.071
Canning Town	5	Stratford	10814	Victoria	2.065
Liverpool Street	5	Mile End	10302	Hyde Pk Corner	2.053
Paddington	5	Bethnal Green	10017	Embankment	2.041
Shadwell	5	Finchley Road	8905	Piccadilly Circus	2.041
Turnham Green	5	Earl's Court	8706	St. James's Park	2.035
Acton Town	4	King's Cross	8679	Regent's Park	2.032
Bond Street	4	Wembley Park	7968	King's Cross	2.029
Camden Town	4	South Ken	7182	Liverpool Street	2.026
Canada Water	4	Euston	7156	Marble Arch	2.026
Canary Wharf	4	Gloucester Rd	7042	Tottenham Ct Rd	2.026
Embankment	4	Paddington	7028	Moorgate	2.020
Euston	4	Victoria	6558	Charing Cross	2.017
Finchley Road	4	Harrow-o-t-Hill	6253	Great Portland St	2.017

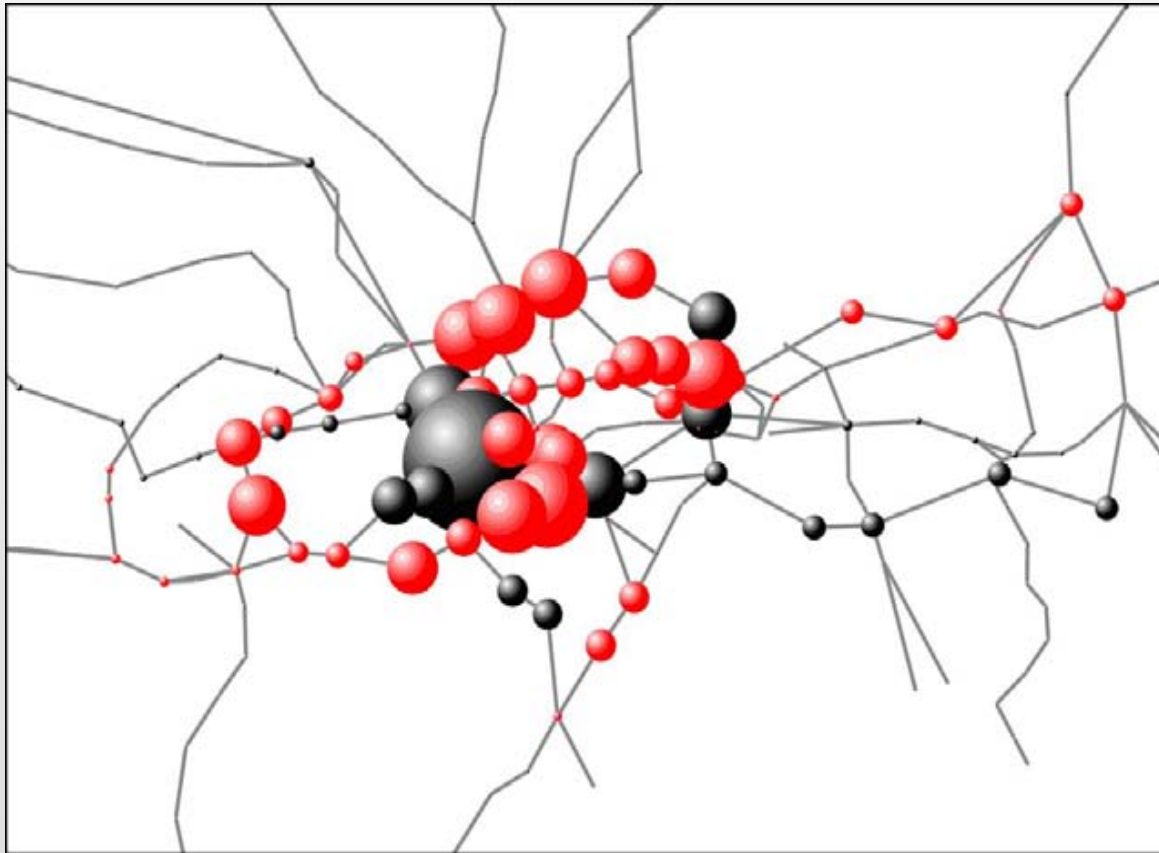
A Preliminary Analysis (3)

Closing Liverpool Street



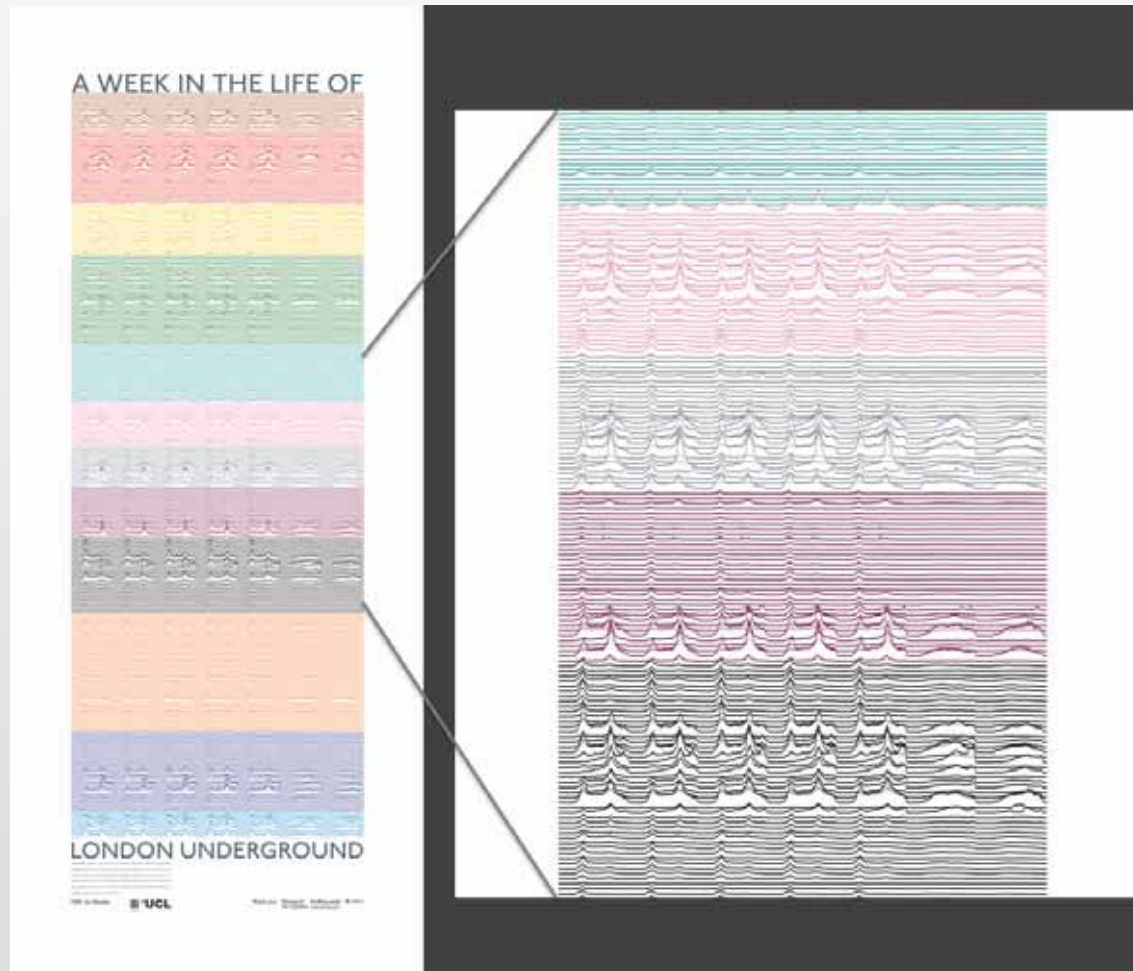
A Preliminary Analysis (3)

Closing Green Park

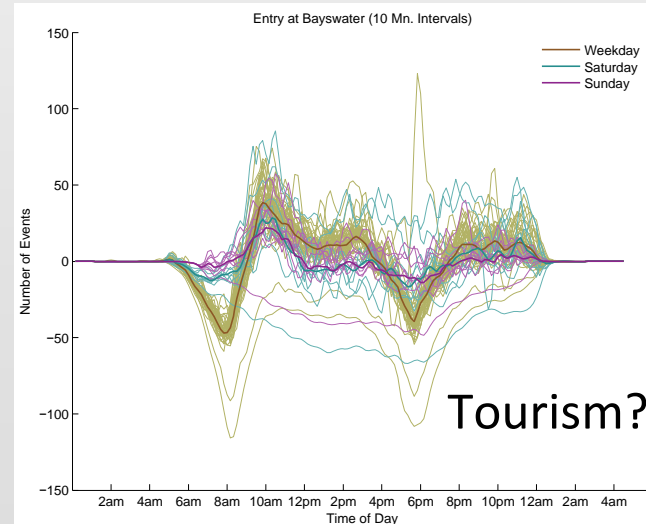
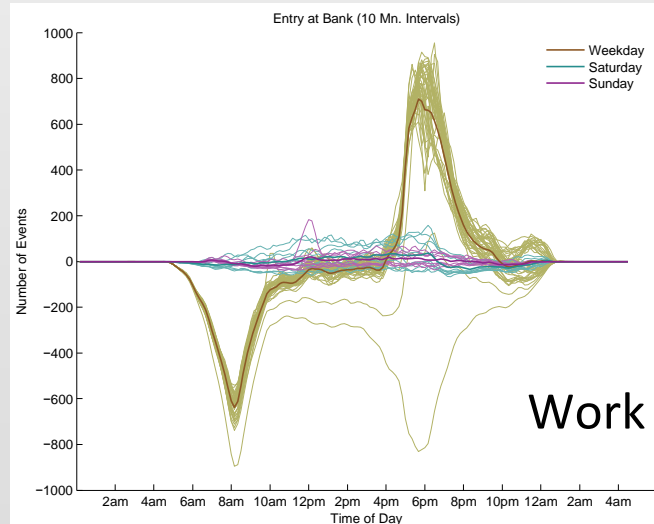
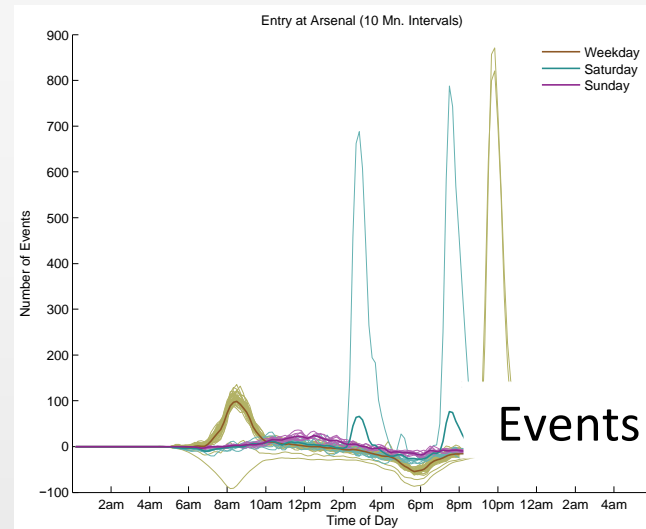
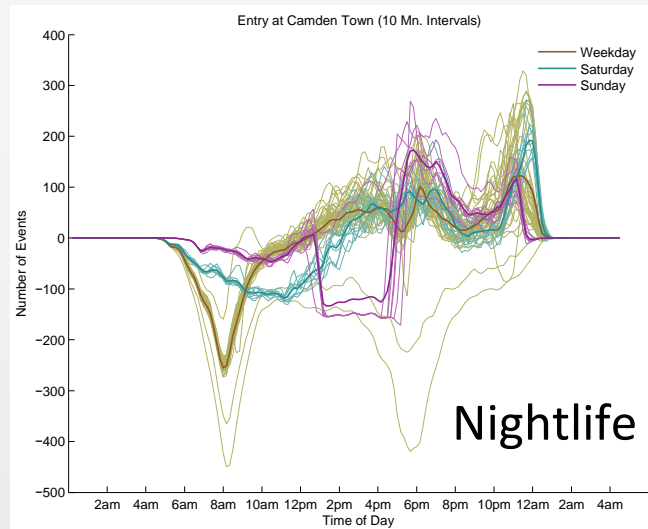


Problem 3: The Shortest Path

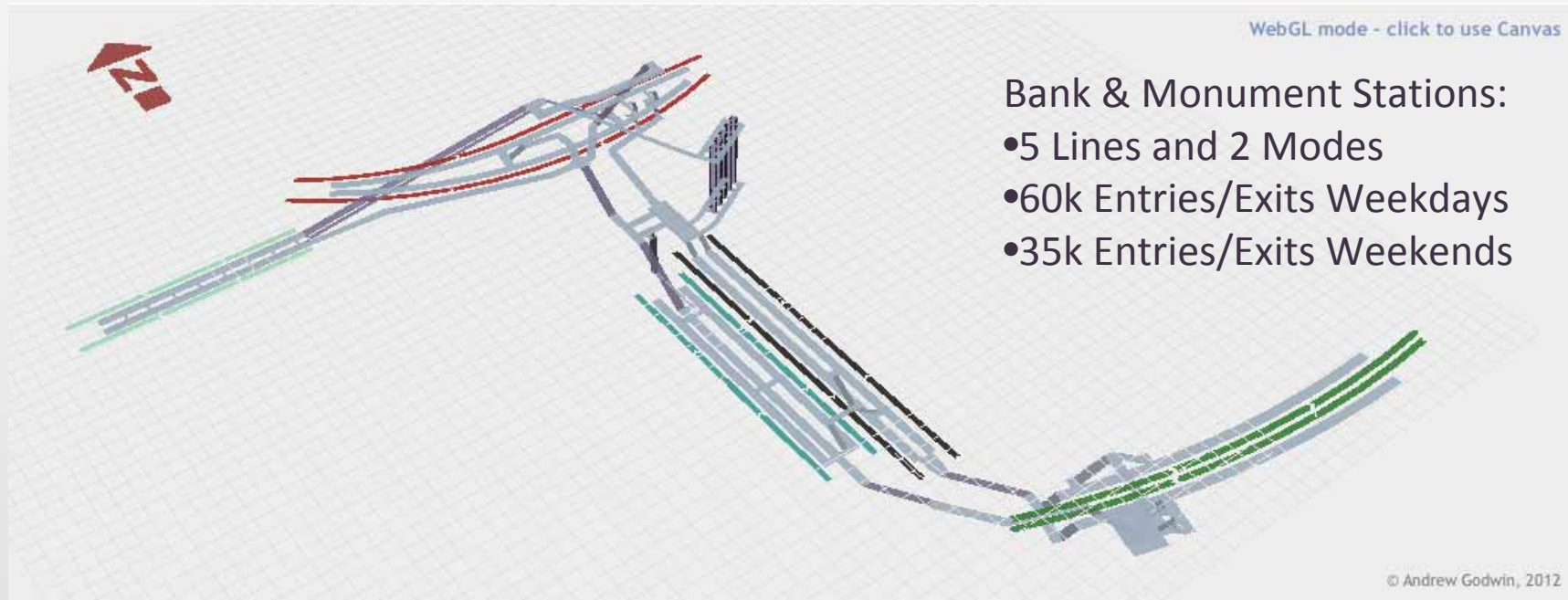
but to remind you, here we are dealing with flow data with the usual rush hour peaks and troughs



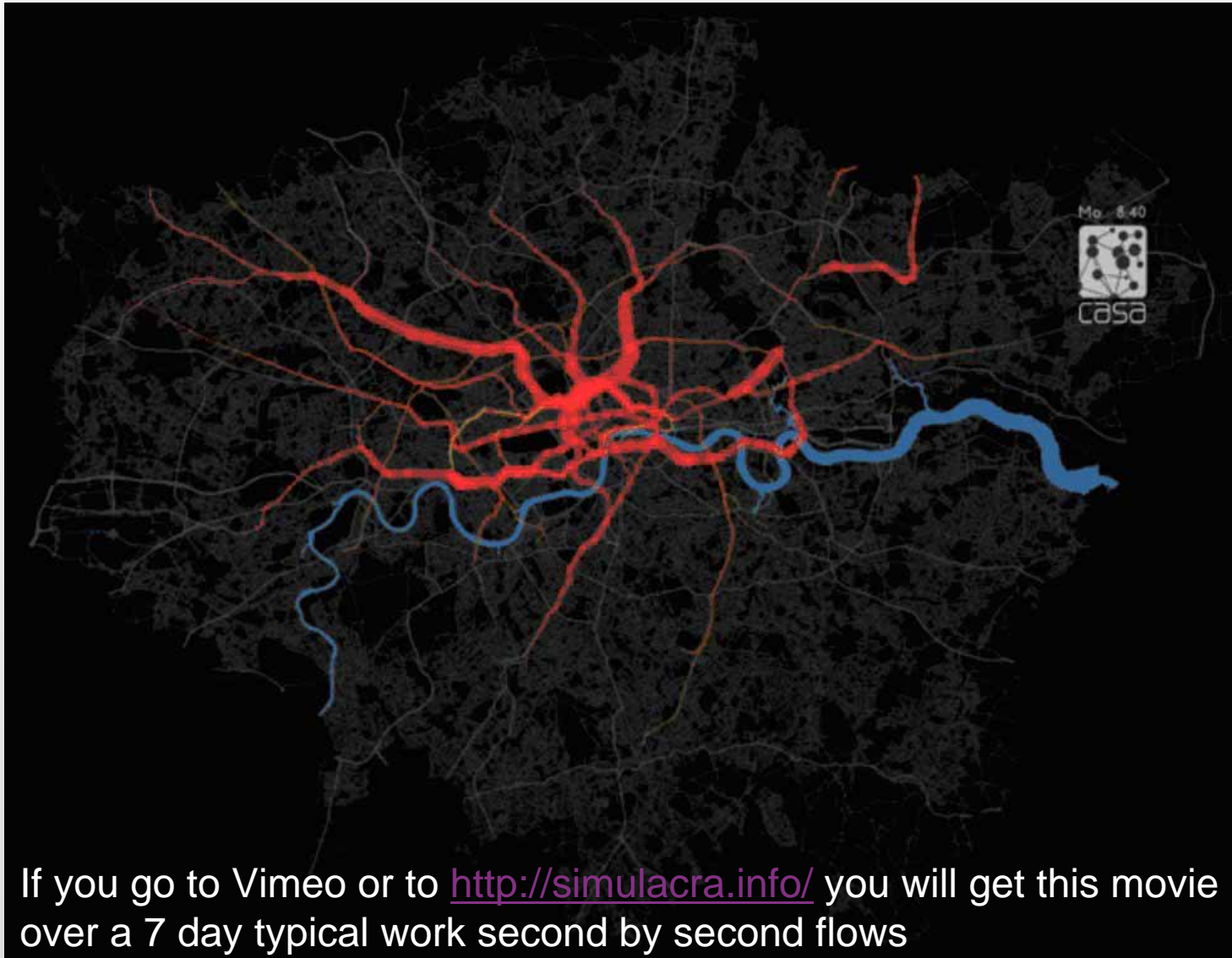
Station Time: Weekdays, Saturdays and Sundays



Problem 3: The 'Shortest'-Path Problem



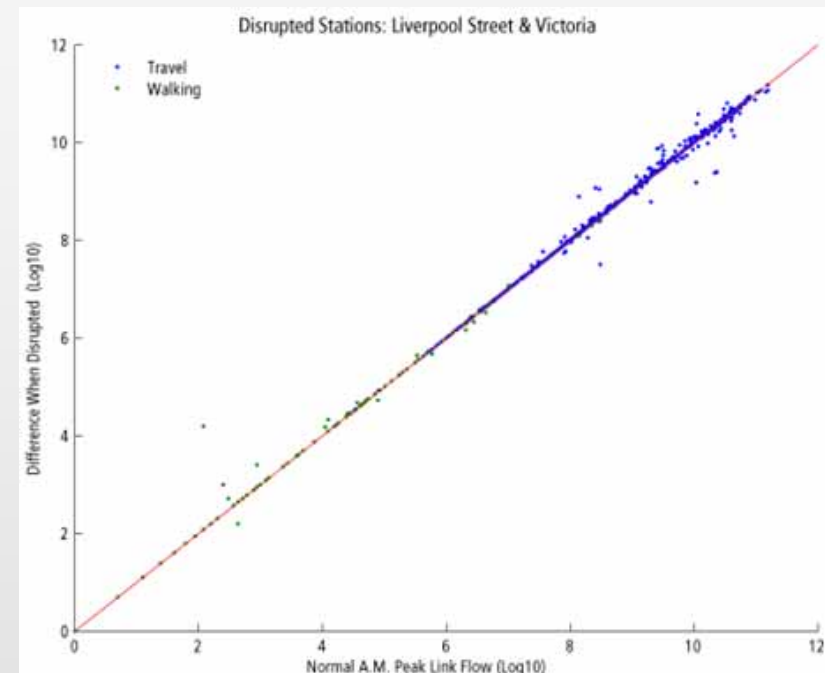
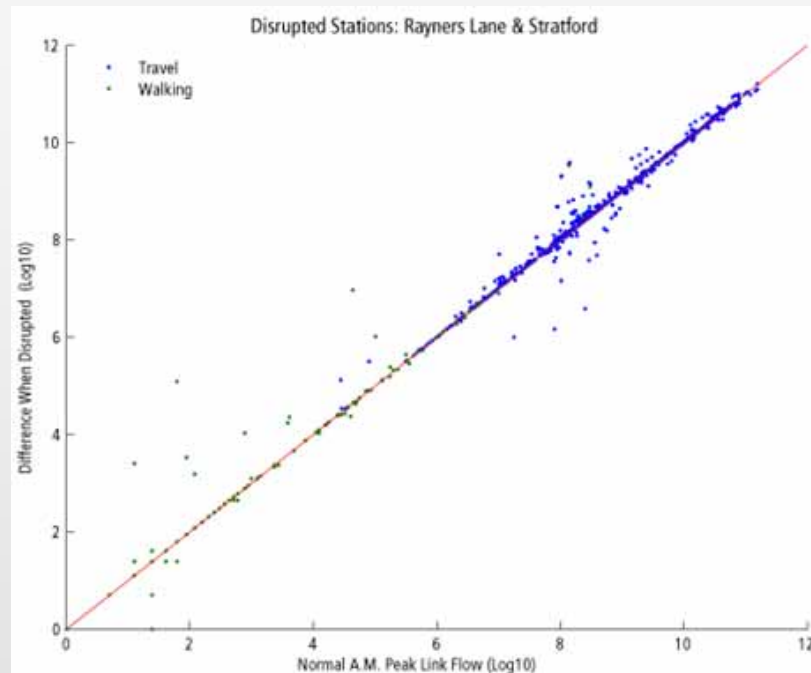
Although a simple station/line network may be sufficient for small cities, for 'Mega-Cities' such as London, New York, or Tokyo a much more detailed network is needed with interchanges measured down to the platform level. The 'penalties' for changing lines (and permitted Out-of-Station Interchanges) can be severe and should be included in a schematic network representation.



Methodology

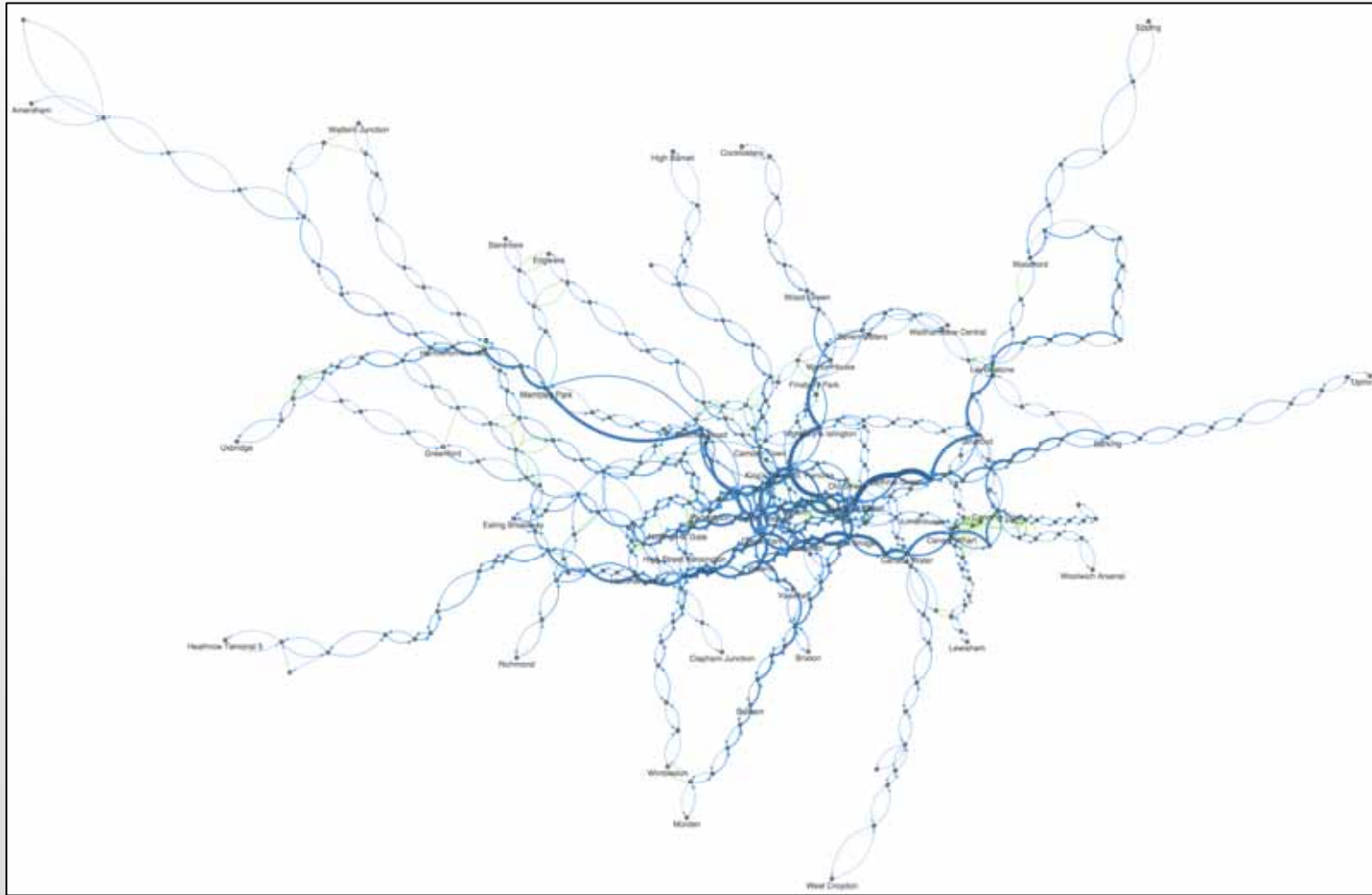
1. Build average O/D flows between all Under- and Over-ground stations
 - 33 days of activity with 100% coverage of pseudonymous Oyster cards
 - More than 300 million unique trip segments (of which 120 million by some form of rail)
2. Build walking network between all stations within 5km of each other
 - Routing on OSM network provided by routino using realistic preferences for walk speed and intensity of road usage
3. Build integrated travel-time network representation of both modalities
 - End-to-end travel time extracted from routino routing
 - Physical layout of stations inferred from real-time platform data
4. Simulate simultaneous disruption for 1 or 2 stations
 - Use real O/D matrix and remap disrupted trip segments
 - Realistic disruption on basis of entry/exit/interchange breakdown
 - Measure changes in volumes and 'lost' travel times across segments

Link-Level Disruption

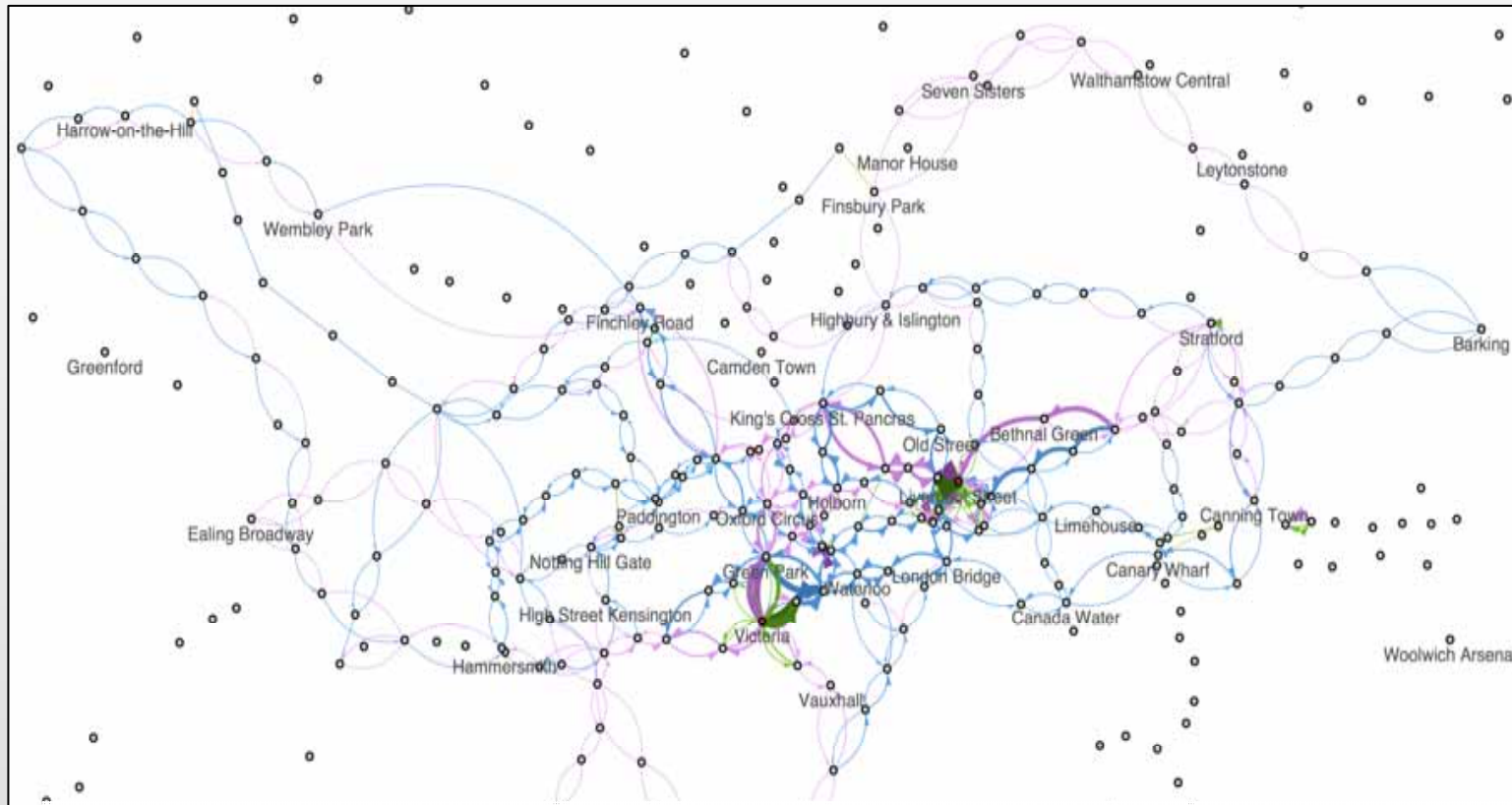


Single- and dual-station disruptions produce unexpected link-level interactions: changes in shortest-path typically cause some links to lose passengers, and gains are often less than expected. Moreover, it is not the biggest and most central stations that cause the largest shifts!

The Undisturbed Network

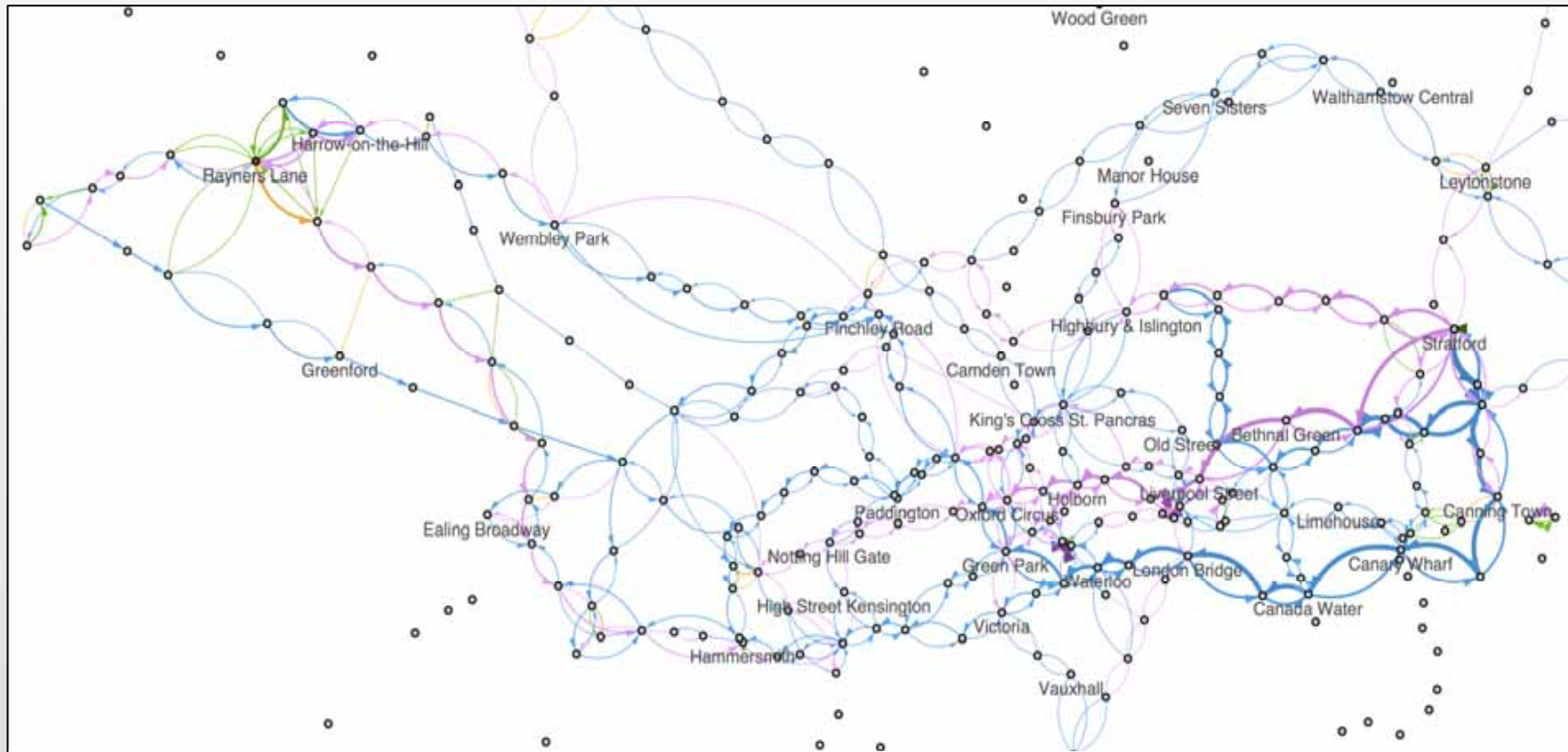


Liverpool Street & Victoria



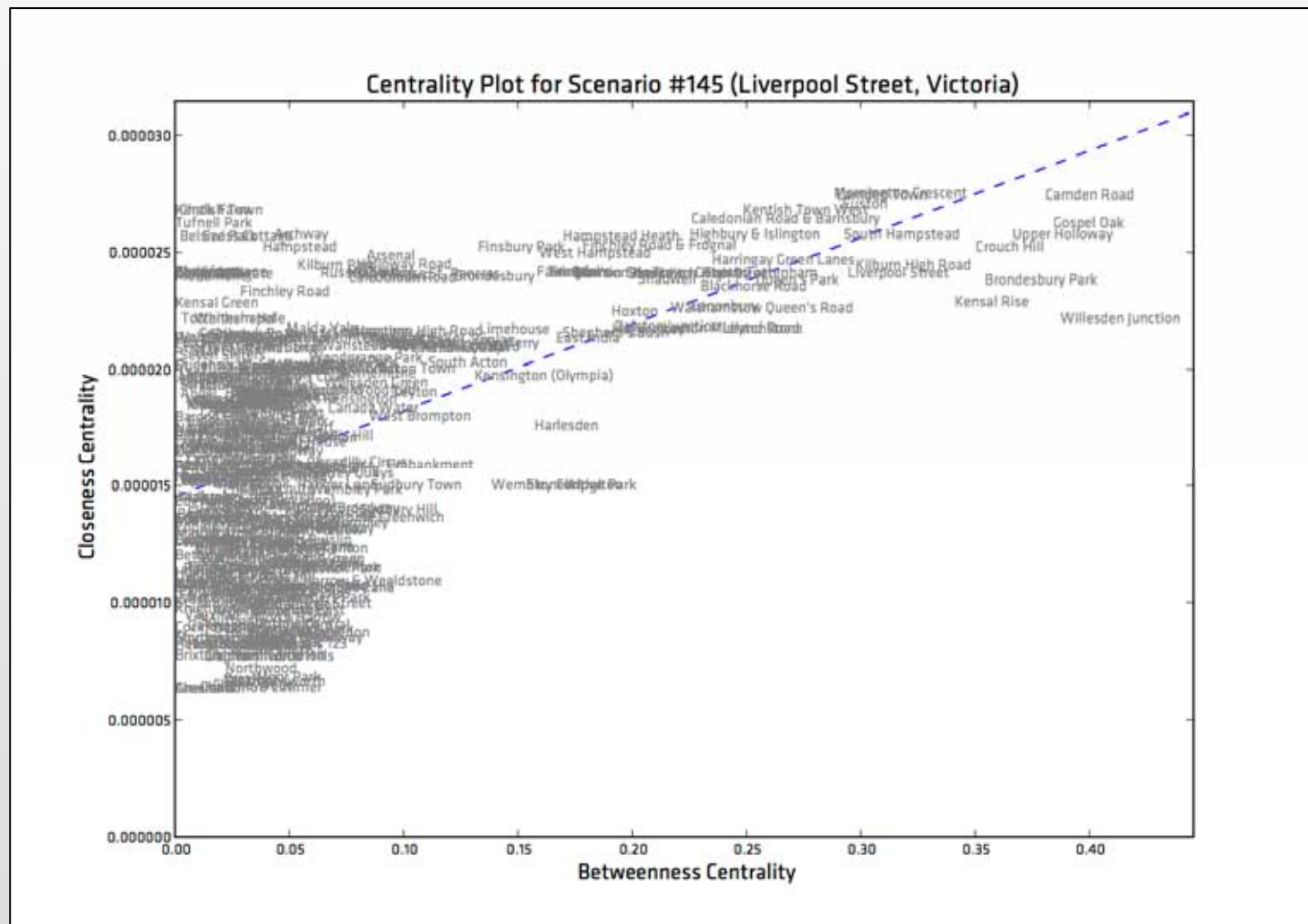
Two of London's busiest stations – because of connections to mainline rail – but if disruptions are localised to the Tube *alone* then there are many more local substitutes.

Rayners Lane & Stratford

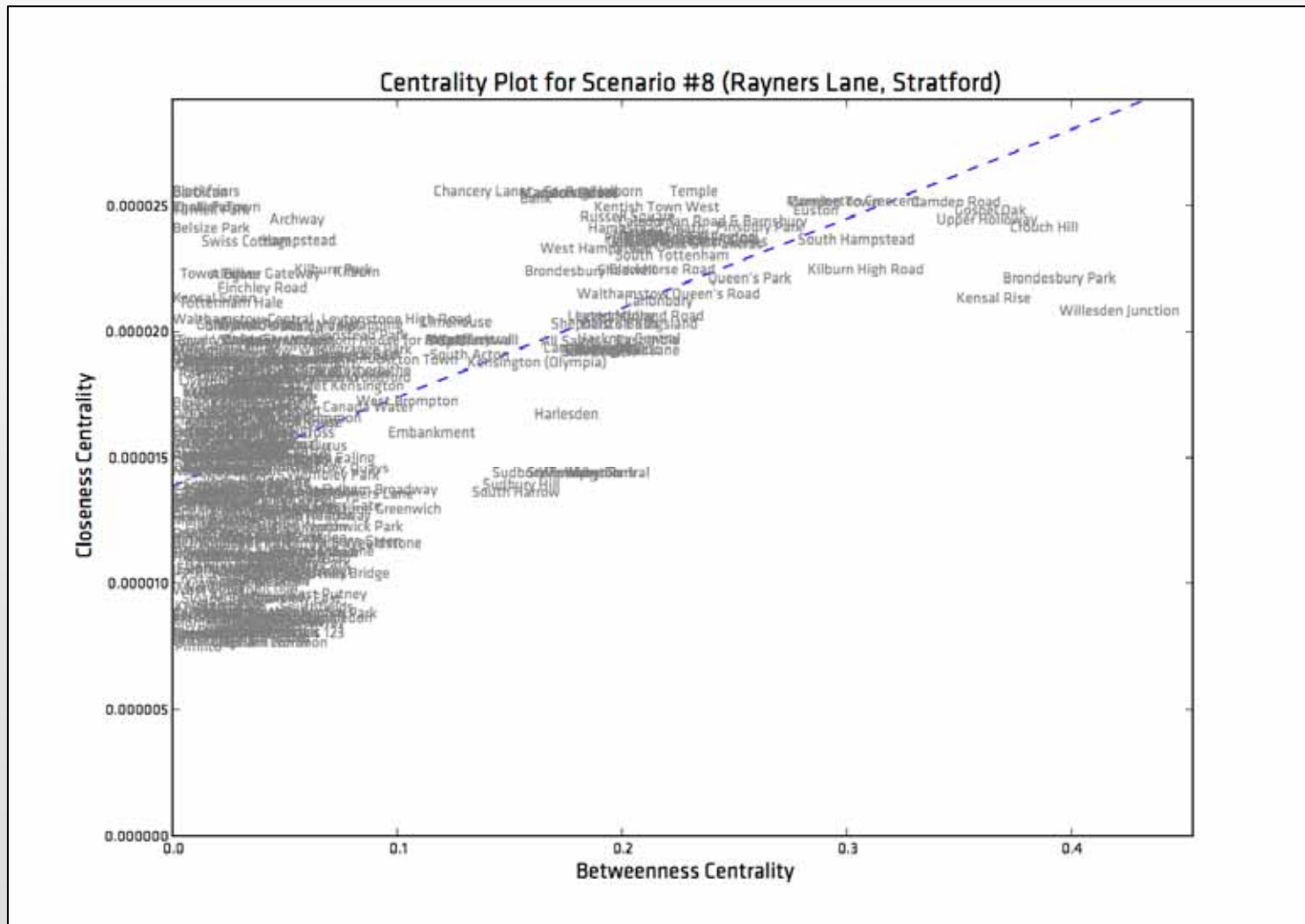


Secondary interchanges outside the core seem to cause greater disruption. Major re-routing required to complete journey, and time lost to walking long distances or travelling via more circuitous routes is much greater.

Betweenness & Closeness: Liverpool Street & Victoria



Betweenness & Closeness: Rayners Lane & Stratford



The Impact of the London Olympics

Strategy #1: Demand Distribution

The distribution of Olympic venues helped to create spectator contra-flows to much of the normal rush hour traffic; however, commuters also adjusted their behaviour in small, but important ways: increasing their use of the routes least affected by Games (e.g. the Victoria line, some NR) and moderating their use of heavily trafficked routes (e.g. the Central line).

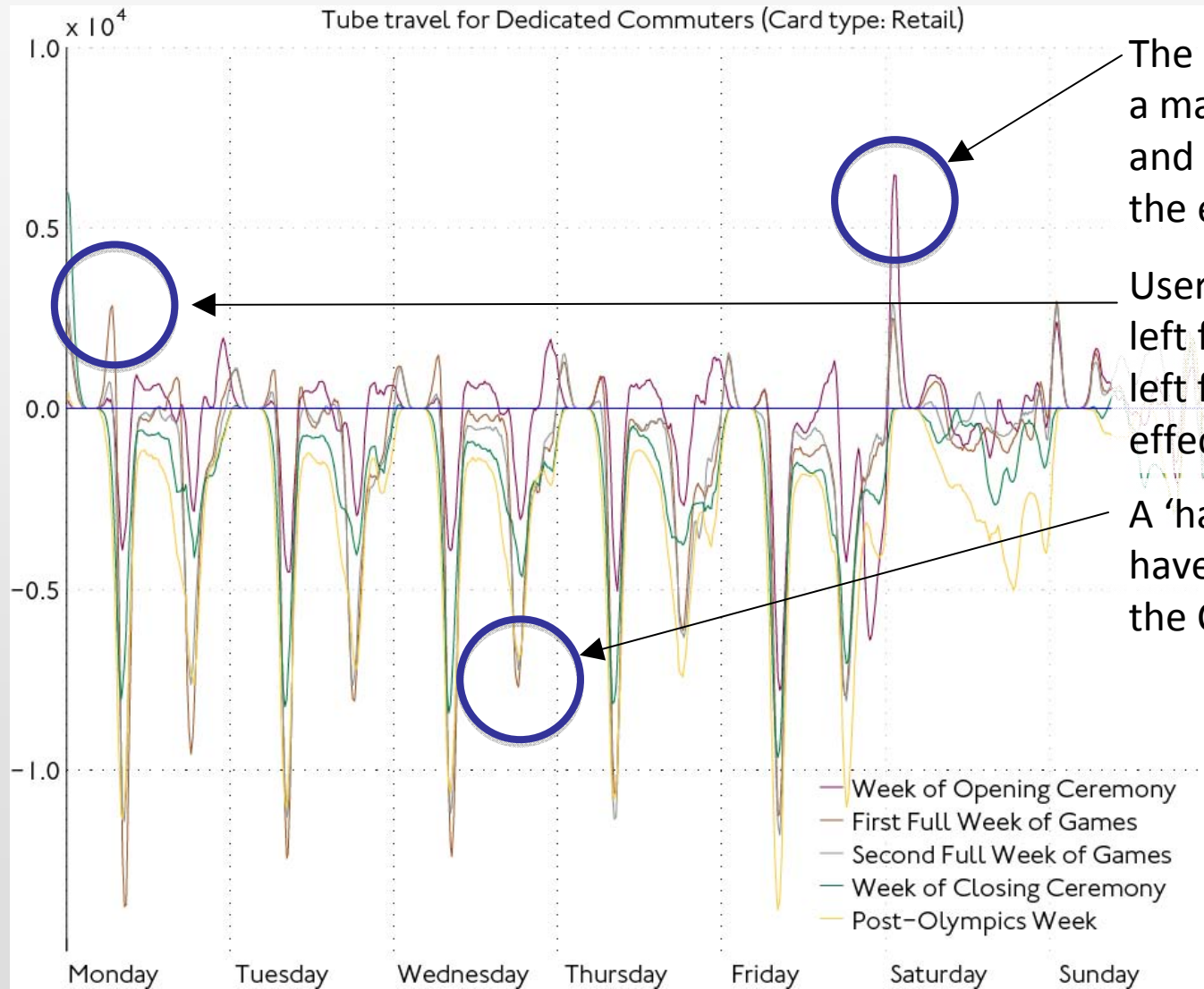
Strategy #1: Demand Distribution

Londoners with greater discretion over their travel clearly responded more actively to suggestions that they avoid intensive travel during the Games. However, the single largest group of users – ‘Commuters’ who make use of bus, rail and Tube network on a daily basis – made proportionally fewer adjustments to their average weekly activity.

Before the Games	Games Week #1	Games Week #2	Count	Pct.
Commuter	Commuter	Commuter	2,250,000	87%
Commuter	Commuter	Infrequent	122,000	5%
Commuter	Infrequent	Commuter	146,000	5%
1-3 Times/Week	1-3 Times/Week	1-3 Times/Week	782,000	48%
1-3 Times/Week	Infrequent	Infrequent	274,000	17%
1-3 Times/Week	Commuter	Commuter	75,000	5%

For Commuters, n ≈ 2,580,000; For 1-3 Times/Week, n ≈ 1,640,000

Strategy #3: Demand Shifting

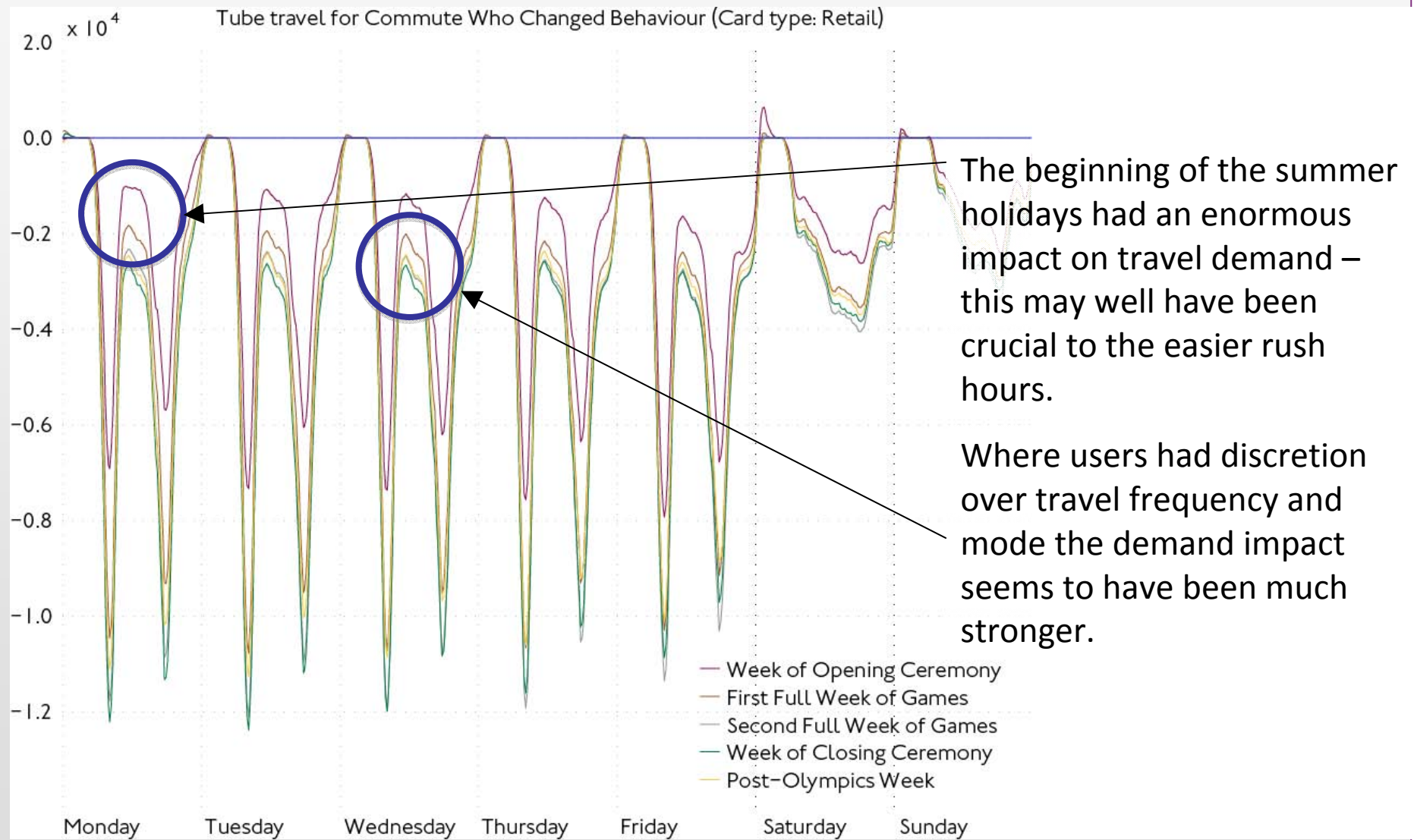


The Opening Ceremonies had a massive impact on activity, and users adjusted quickly to the extra hour.

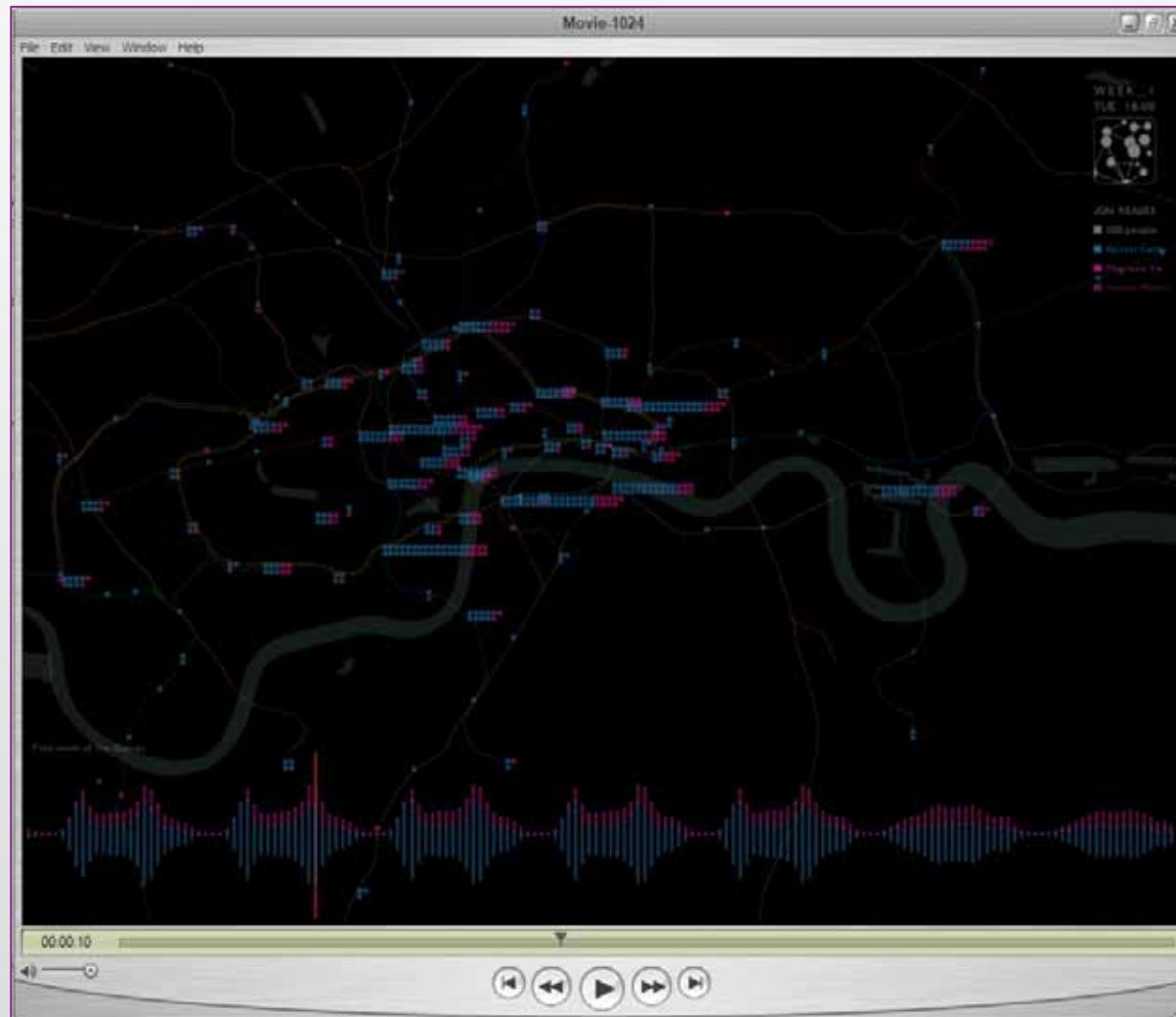
Users left for work earlier, and left for home later. And some left for home earlier! The effect fell off rapidly.

A 'hangover' effect seems to have come into play following the Closing Ceremonies.

Strategy #3: Demand Shifting



To end, an interesting movie of the flows during the Olympics



Next Steps

1. Identify meaningful measures enabling comparison *between* scenarios:
2. Need to capture both individual station and cumulative network impacts
3. As well, the network model could be improved in several ways:
4. Better-respect known route-choice preferences using RODS survey data and, potentially, Space Syntax-like 'cognitive complexity' (e.g. compare difference between route time and map complexity)
5. Improve modelling of interchange penalties by taking re-entry and ticket cost into account for different passenger groups
6. Improve modelling of interchange times at particularly large and complex stations (e.g. King's Cross St. Pancras)
7. Incorporate National Rail system flows since this will change many network measures substantially (although non-Oyster users would make analysis of potential hotspots much, much more complicated).

Credits

Jon Reades did the work on the Oyster Card data and Richard Milton on the tube, bus and rail counts of vehicles from the APIs

Ollie O'Brien is in charge of the bikes project

All did their work in CASA and Jon Reades is now at Kings (KCL)