

Smarter Bus Information in Leeds

Thomas Forth

project demonstration url : www.tomforth.co.uk/dynamicbusmaps

email : thomas.forth@gmail.com

twitter : @thomasforth

Executive summary:

Leeds, an English city region with a population of 2.4 million, is the largest in Europe to have no metro or tramway. Since deregulation in 1984, bus ridership has fallen by over a third compared to a doubling in London in the same period. Some of the underlying causes of this decline can be solved by using data better; bus stops in Leeds do not show route maps, or explain fares and very few provide live timetables. I have prototyped and built a location-aware mobile website that combines three sources of information to automatically generate route maps for bus services within a 5 minute walk of the user. Directions to stops, as well as live timetable data for the next hour, are provided automatically.

Visit www.tomforth.co.uk/dynamicbusmaps in any modern browser to try it.

Principal sources of data:

- NaPTAN (National Public Transport Access Nodes) — <http://data.gov.uk/dataset/naptan>
- WY Metro live bus information — <http://wypte.acislive.com/>
- Yorkshire Travelnet — <http://www.yorkshiretravel.net/>

Processed data:

- The project is hosted at www.tomforth.co.uk/dynamicbusmaps. A demonstration, source code and processed data is available at that site.
- An excel spreadsheet for the automatic generation of the required raw data from Yorkshire Travelnet route descriptions is also provided at that site and relies on several macros which are safe to run.

Tools used:

- heatmap.py — jguy.com/heatmap/ — used to create heatmaps of NaPTAN data combined with WY Metro live bus information as KML overlays for google earth.
- Raphael — [raphaeljs.com](http://dmitrybaranovskiy.github.io/raphael/) — a javascript library used to draw bus route maps.

About Leeds:

The city of Leeds has a population of around 800 thousand. For simplicity and comparability I use the name to describe the larger urban zone (LUZ) of 2.4 million people as defined by Eurostat. This definition largely overlaps with the county of West Yorkshire. For comparison, London's LUZ has a population of 11.9 million.

The problem.

In 1984, the year I was born, buses were deregulated in all but one UK city, with private operators in a free market replacing local government subsidiaries as the providers of services. Figure 1 shows the resulting trend in bus ridership in Leeds compared with London, where buses were kept under the centralised control of Transport for London (TfL). The decline in ridership in Leeds is despite a growing population, a wider trend of increased public transport use across Europe and the introduction of free bus passes to all UK citizens over 60 in 2008.

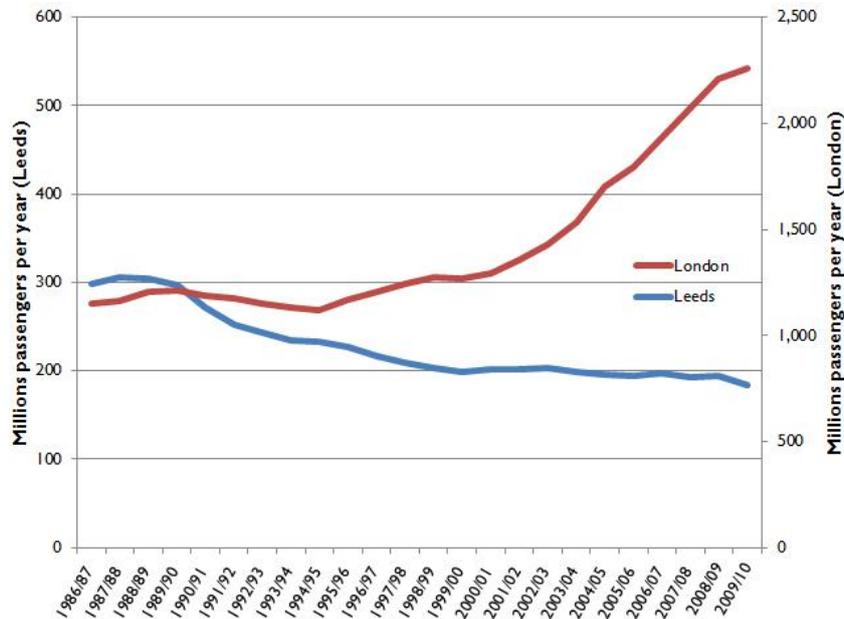


Figure 1 : Annual bus ridership in West Yorkshire (blue line, left scale) compared with bus ridership in London (red line, right scale) since deregulation in 1986. Raw data from WY Metro and TfL.

How can data help?

My idea helps because currently available information is not connected and provided to users in a useful way. There are 18754 bus stops in Leeds, see figure 7, and yet a typical bus stop contains only the information shown in figure 2. Timetables are provided for only some of the routes, maps are provided for none and a live timetable can be found only by scanning the QR code found on a few selected stops. This means you're already waiting for the bus by the time you learn how long you'll have to wait. If you don't know what bus you want to take, or which stop you need to be at, the current system provides no information. It's no wonder that people in Leeds increasingly don't use the bus; current systems are not combining the data that exists with widely available technology well enough to make the system understandable.

Combining and providing relevant information based on location.

Drawing bus maps takes time and costs more money than competing bus companies want to spend. This is especially true if competitors can quickly steal customers one company has earned on a particular route. Leeds' regional public transport authority, WY Metro, are able to provide some services but they do not have the resources to lead growth in public transport. They have however been able to provide the real-time information that, when combined with other open data, might let others do that.

My solution creates a personalised route map for every possible location in the city as shown in figure 3. It shows users how to get to nearby stops as shown in figure 4 and it gives them real-time timetables and service updates for available stops and services as shown in figure 5. I have already written the algorithms required to create simplified bus maps — figure 6 — for implementation in the future. Information on the names of each bus stop and the areas of the cities they're in are ready for future use and I have done some interesting experiments with this data.

Currently my system works for both directions of 8 bus routes in North-West Leeds, an area chosen for the high density of bus services serving a large student community (figure 10), but it scales without any significant issues. Currently information on routes needs to be parsed one route at a time from Yorkshire Travelnet — <http://www.yorkshiretravel.net/> — into the Excel spreadsheet that prepares the raw data for my service. The system I have developed for this process makes it possible to add a new route in around ten minutes so this is not a major bottleneck to further expansion and it is conceivable that by moving the data from Excel to a shared google docs spreadsheet route data could be updated collaboratively.

The best way to see the value my service adds to raw data, and to imagine the ways it could be built upon and improved, is to try it at www.tomforth.co.uk/dynamicbusmaps .

An interesting extra: explaining why George Galloway is the MP for Bradford West.

I don't need to tell a story with data to explain why my system is needed but when I started using the data to achieve my task some interesting things emerged.

The NaPTAN dataset lists 18754 unique bus stops in Leeds so to see where they were I projected them as a heatmap using heatmap.py — jguy.com/heatmap/ — to create figure 7. Timetable information is available via wypte.acislive.com/ for 13820 of these stops which changes the heatmap a bit to create figure 8. These figures suggest that Leeds has a lot of bus-stops serving most of the population but this is actually misleading. When we ignore those stops where there are fewer than 5 buses an hour we get figure 9 and we see that large parts of Leeds are neglected by high-frequency services and that the key urban centres in the region are disconnected.

Much has been made in the UK of George Galloway's recent victory in the 2012 Bradford West by-election. My own reading is most similar to that in the Guardian piece entitled, "Bradford West result was symptom of UK's brutal north-south divide". The article bemoans the lack of infrastructure spending in the North and sure enough if you overlap high-frequency bus services with the constituency boundary — figure 11 — the results are quite astounding.

Correlation is not causation, but it's often a good place to start.



Figure 2 : Total information provided at my nearest bus stop.

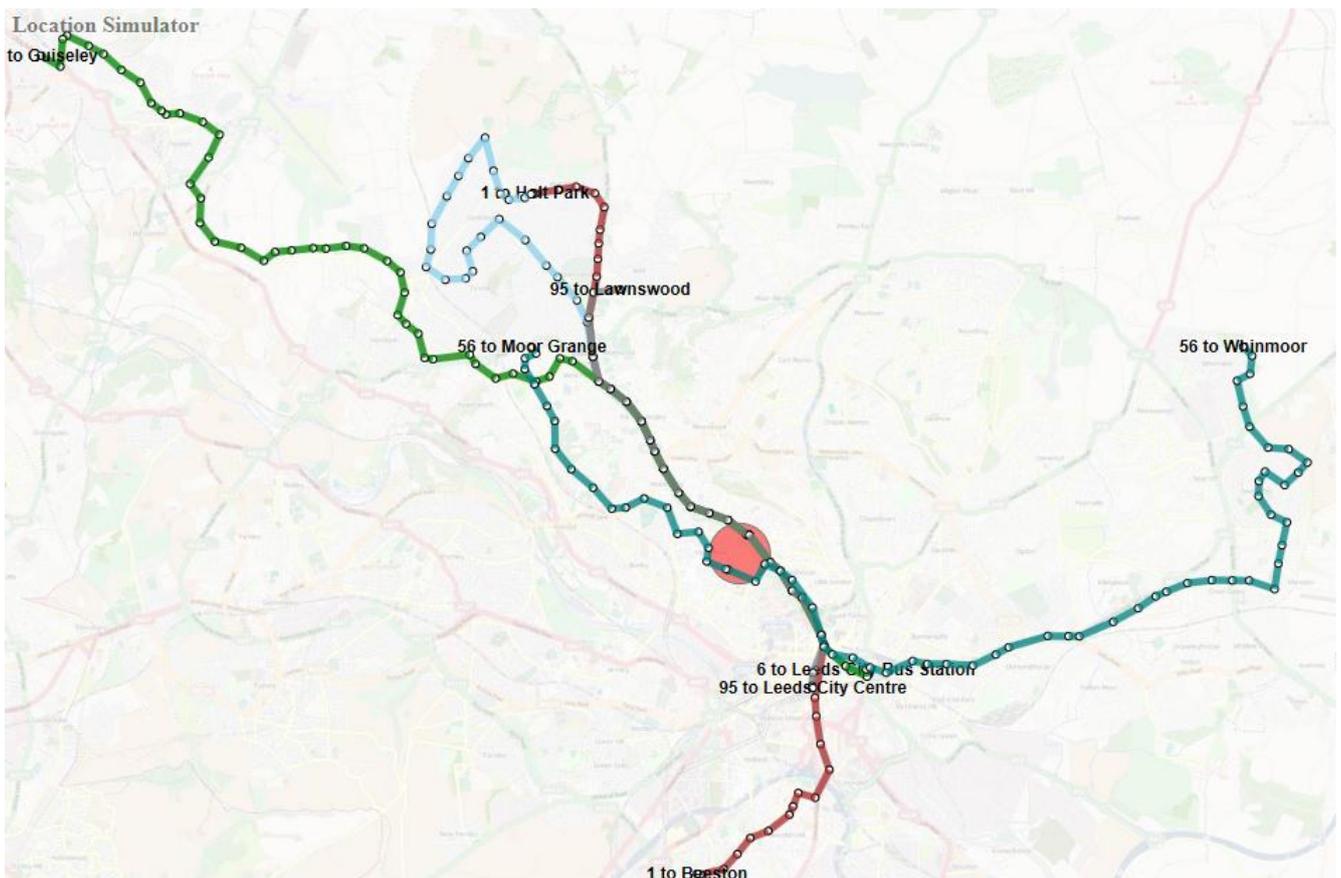


Figure 3 : Automatically generated map of all bus routes available within a 5 minute walk of the user's location (red circle).

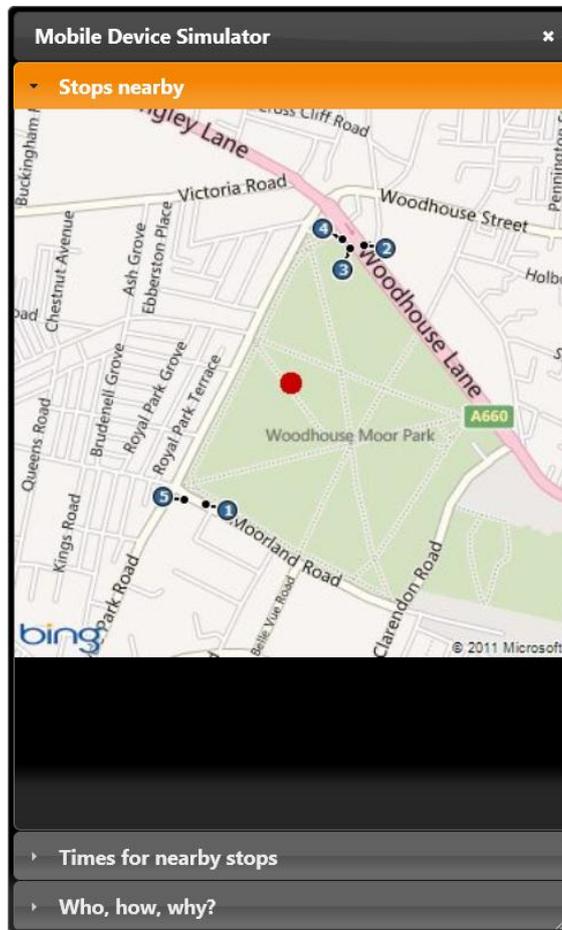


Figure 4 : Local map displaying current location (red) and location of the closest bus stop for each available service in figure 3

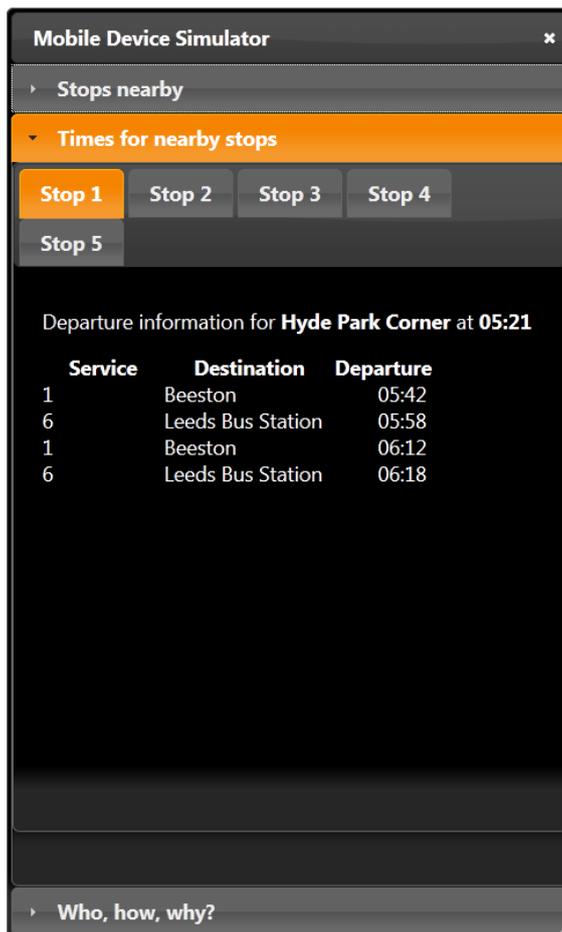


Figure 5 : Timetable information is available instantly for all the stops shown in figure 4.

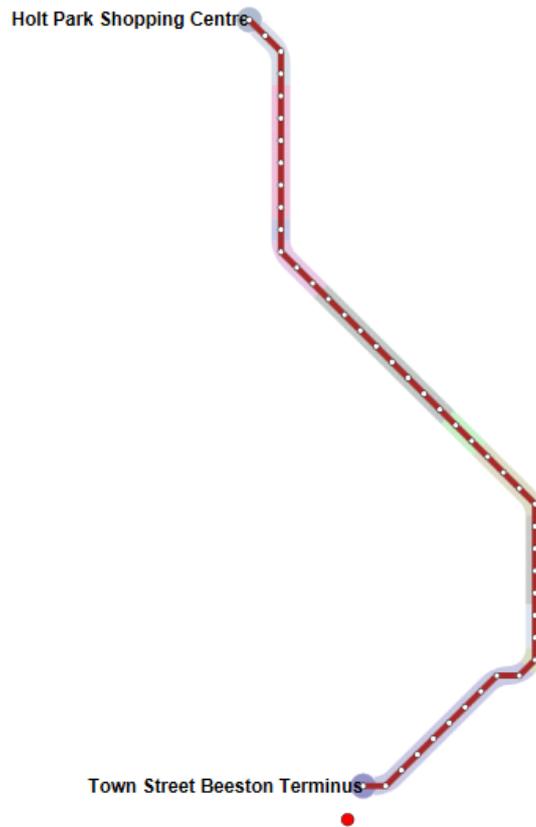


Figure 6 : Simplified maps, such as this version of Route 1, also in figure 3, may be easier than a full map to view on a mobile device and can be generated automatically.

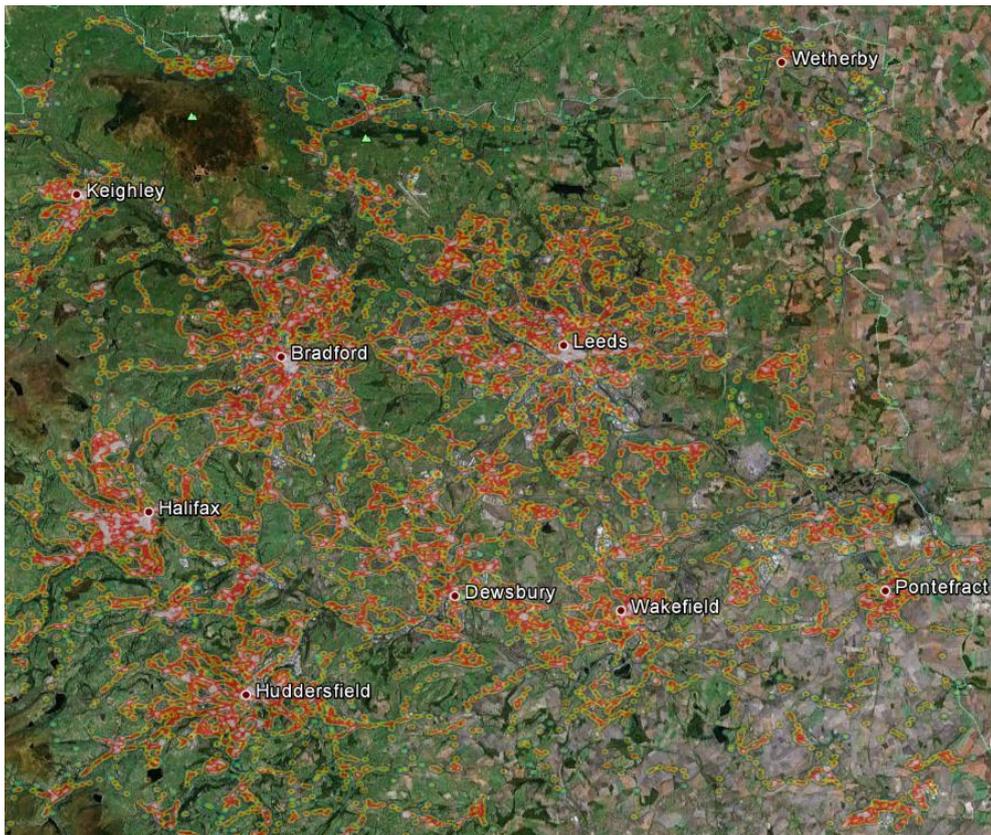


Figure 7 : Heatmap of all bus stops in Leeds.



Figure 8: Heatmap of bus stops in Leeds with live data, mostly just static timetables.

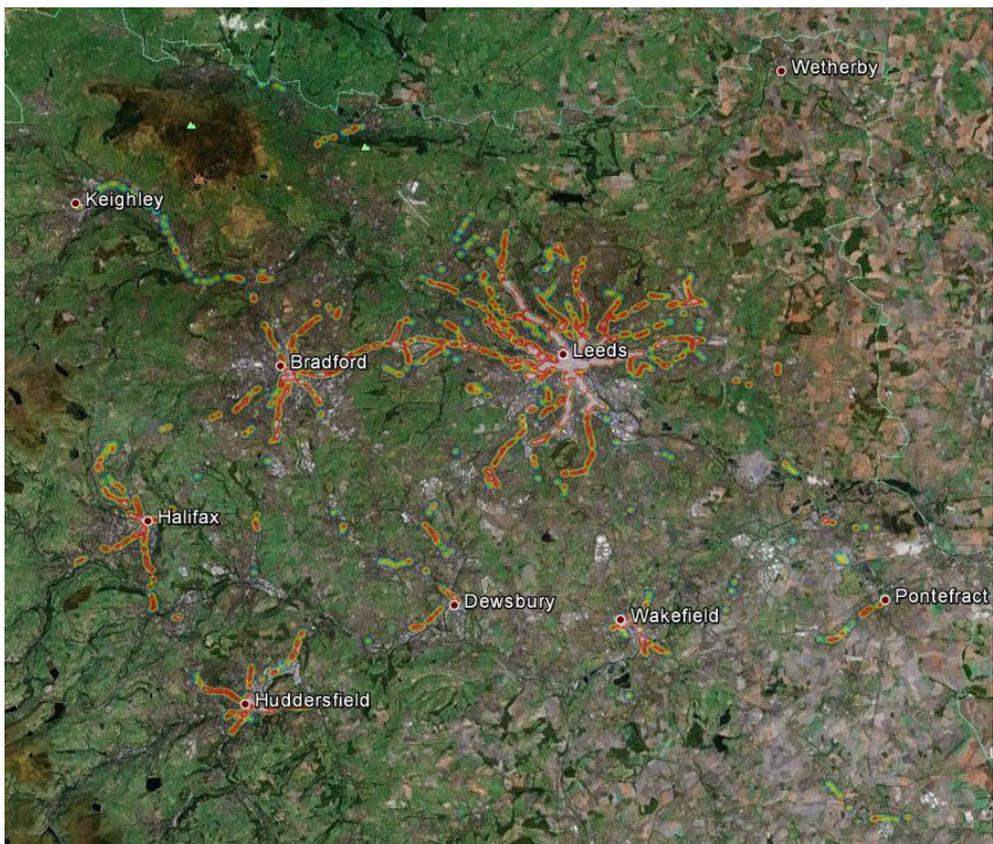


Figure 9: Heatmap of high frequency (>6 buses per hour) bus stops shows substantially less coverage than figure 8.

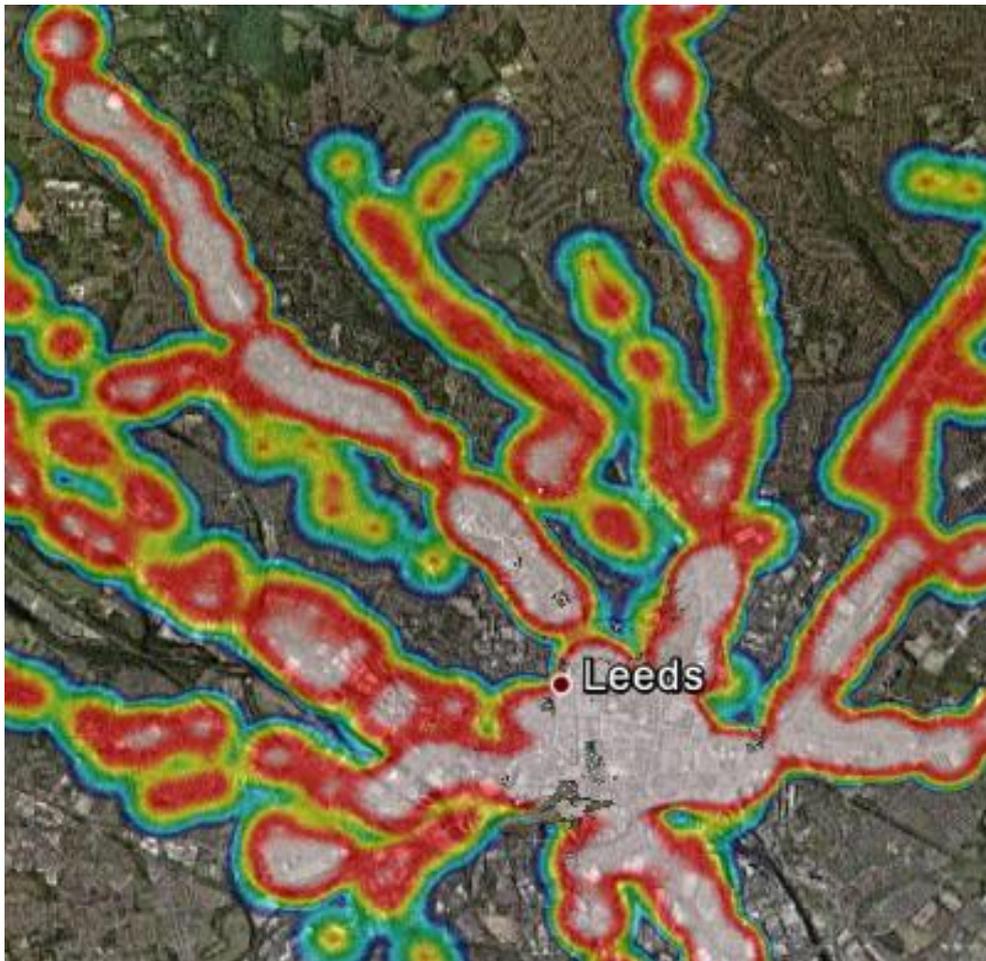


Figure 10 : Heatmap of high-frequency (>6 buses per hours) stops in central and North-West Leeds shows good coverage.

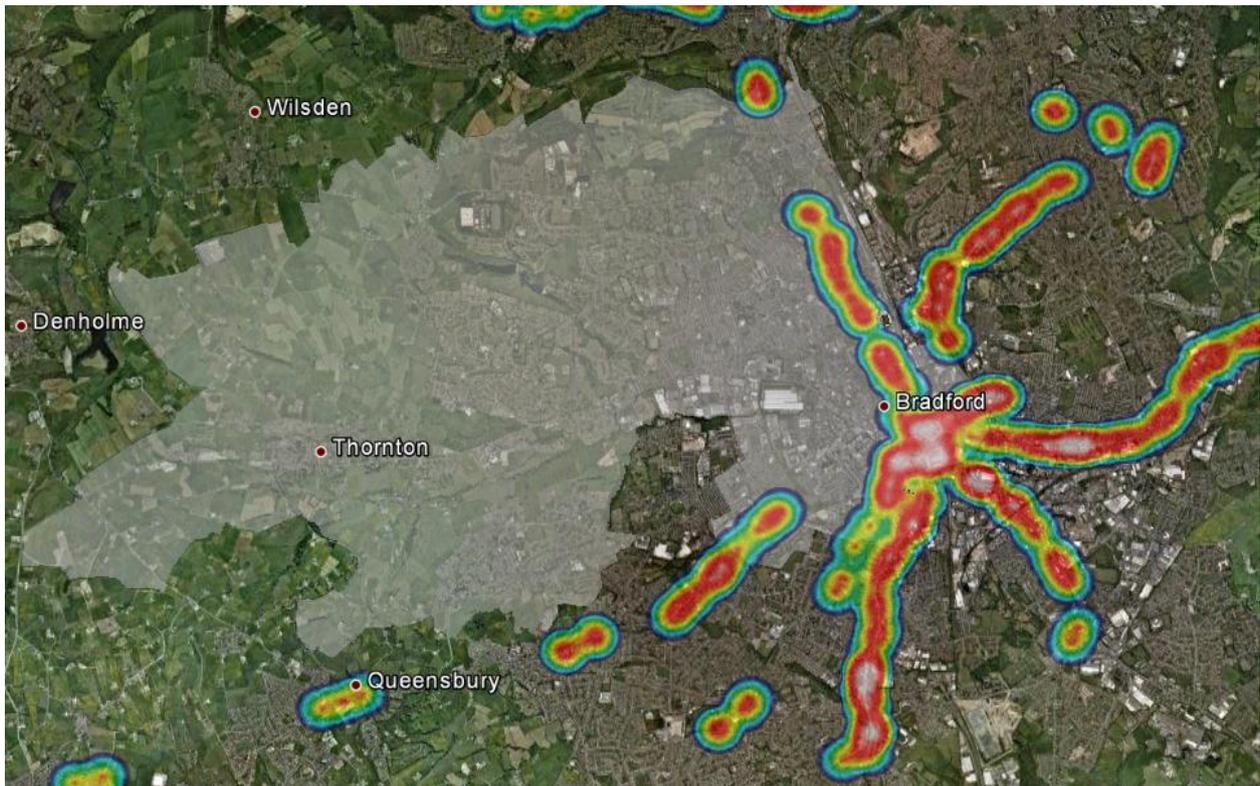


Figure 11 : Bradford West constituency in white with high frequency services from figure 9 overlaid. Note the almost complete lack of high frequency services within this mostly urban constituency now famous for rejecting the three major parties and electing George Galloway.