Cycle Networks — Finding the Missing Links

Max Oblein
Supervised by Prof R. E. Wilson

March 13, 2021

Chapter 1

Introduction and background

Cycling is becoming increasingly more popular as a means of transport around large cities like Bristol [1]. This rise is due to a combination of factors, chiefly, increased environmental awareness, personal health, and the huge amount of road traffic meaning cycling is in some cases faster than driving for short journeys.

The National Travel Survey (NTS) [2] reports that cycling currently makes up 2% of all trips, where a trip is defined as a one-way course of travel with a single main purpose, and that the average length of these trips is 23 minutes. This seems a small proportion given that 38% of people surveyed own a bicycle. The third wave of the National Travel Attitude Survey (NTAS) [3] posed questions as to why people don't like to cycle. Their conclusions found that 60% of people agree that "it is too dangerous for me to cycle on the roads". Research suggests that an increased provision of cycling infrastructure is positively correlated with the proportion of people cycling within that city [4]. The key to increasing cycling rates in Bristol is thus to ensure cyclists feel safe and have segregated paths and lanes to use throughout their journeys.

Currently, when planning road networks, lots of money is spent with engineering consultants to decide where to build new roads [5]. Cycling is much less common as a mode of transport than driving, so planning the cycle network has a significantly lower budget. This means that selecting where new cycle paths are built or lanes are painted is often an arbitrary decision made by the local authority. Therefore, this project aims to develop a mathematical framework to inform these decisions.

There is some previous work in the area of planning cycle networks, and most studies focus on two areas: the effect infrastructure has on the number of cyclists and the evaluation and proposal of cycle schemes based on geographical factors. The first area motivates the need for cycling infrastructure [4]. The second is useful for the evaluation of routes suggested by this project based on urban factors such as the amount of retail or business land area at both origin and destination of the route [6]. The methods presented by Milakis et al. [7] give a comprehensive methodology for evaluating proposed cycle routes. They use a weighted scoring system to assign good scores to origin-destination (OD) pairs attached to important facilities e.g., universities. In the literature little work has been found in the area of assigning new cycle routes automatically using a mathematical approach.

The main concern when automatically proposing roads for new cycling infrastructure is deciding how best to load a road network with cycling demand. Two papers have been identified for their methods in loading street networks with cycling demand. One approach, formulated by Milakis et al. [7], approaches this problem with the use of large demand centres. The study is based in Athens and the idea is to draw centres around areas of large demand. The centres are chosen based on three main criteria: the land use (retail, offices, education or leisure), the OD matrices from the last two travel surveys in Athens, and the network centres detailed in the master plan of Athens. Milakis et al. identify eight such centres and they are used as the OD pairs from which to generate synthetic cycle journeys. This methodology is a good starting point for loading the network with demand, but a set of only eight nodes does not allow for realistic models of commuting in cities such as Bristol, where workers commute from and to a huge number of different locations. This methodology is also not automatic: it requires a survey of cyclists to decide which land use attributes give the largest demand to a centre, and creating and carrying out a comprehensive survey is costly. The survey used in this case is also only considering existing cyclist's behaviour and therefore cannot predict the impact of new infrastructure on cycling demand.

A second paper by Larsen [8] takes a more granular approach. The methodology begins by separating the study area (Montreal) into 300m grid squares; the demand within each of these squares is then derived using both observed cycling trips (OD data from Montreal travel survey) and potential cycling trips (car trips shorter than 2km). These two data sources give OD pairs and then standard shortest path algorithms are used to model the route taken between them. Flow intensity is then given by the proportion of routes passing through a grid square. This analytical methodology is then combined with a survey of cyclists as to which road they think should be prioritised for new infrastructure to yield a prioritisation index.

One part of the Montreal study [8] that could be of particular interest to this project is their study of "Dangling Nodes". These are defined as grid squares in which the cycle infrastructure ends. The study concluded that the presence of a dangling node in itself is not an indicator of a need for new infrastructure. Instead they recommend looking at the amount of infrastructure needed to connect the network and performing a case-by-case assessment. The automatic approach proposed in this project may allow for some dangling nodes to be eliminated from consideration, if connecting them to the rest of the network yields no step change in connectivity.

Research by Mauttone et al. [9] applies formal optimisation techniques to the cycle network in some case study cities. The research proposes optimisation of total user and construction costs where both are proportional to distance and user costs are increased on edges without cycling infrastructure. The paper uses a heuristic approach to optimise for user cost on large scale networks, and compares against exact solutions on smaller artificial network examples. However their algorithm does not naturally penalise disconnections in the resulting network, therefore a methodology for identifying links that join up the cycle network is needed.

One tool that has been recommended to us by sustainable transport planners [10] is the the Cycling Infrastructure Prioritisation Toolkit (CyIPT). This is a combination of smaller tools that aims to provide an interactive map of the UK, detailing areas for proposed new cycling infrastructure. The components of most interest are the propensity to cycle tool (PCT) [11] and the rapid cycleway prioritisation tool (RCPT) [12]. The PCT serves as the CyIPT's demand model, using 2011 census data to provide estimates of OD flows. The PCT also uses geographical factors such as 'hilliness' to reduce the propensity to cycle on any given route. Our key interest in the RCPT is its planned facility to compute cohesive networks, which represent more highly connected cycle networks. However, in the current RCPT documentation there is no formal mathematics presented as to how to find these cohesive networks.

Summary of studies

Study	Methodology
Athens [7]	Select centres throughout the city based on their land use e.g., university. The land uses have a cycling priority index based on an OD survey. Then load demand between centres.
Montreal [8]	Split city into 300m grid squares. Then load underlying road network using OD survey data. Prioritise grid squares based on OD demand and a cyclist survey.
Network Optimisation [9]	OD matrix derived from a 2009 household survey in Montevideo (Uruguay). The algorithm presented then tries to optimise for user cost given the length of edges in the road network.
CyIPT [11], [13]	Demand matrix comes from the PCT lower-level tool which uses 2011 census data for OD pairs along with geographical measures such as incline. These are used to assign a value to each edge as to how likely people are to cycle on it. The CyIPT then takes this demand and proposes schemes with high upside whist minimising estimated construction cost.

Chapter 2

Data and Initial Network Analysis

2.1 Data

To create the network needed for the analysis proposed in this project, OpenStreetMaps (OSM) has been identified as the primary data source. A python package OSMnx [14] has been used for easy conversion of OSM data to a network topology. To obtain the data from OSM we must query the OSM overpass API. OSMnx streamlines the query process, although in the case of this project a custom query is built to obtain all the data required to build the cycle network. The fist step is to convert the OSM map of Bristol, our selected bounding geography, to a network. This yields a set of 'ways' (all roads and paths) that can be cycled on within the bounding geography, hence excluding motorways etc. This can then be used to create a graph G whose edge set E represents the set of all ways and whose vertex set V represents junctions between ways and points at which ways gain or lose cycling infrastructure. The number of vertices in G will typically be much less than the number in the OSM data as it is not necessary to describe the curvature of each way in our study.

2.2 Prescribing the weights of edges

The first step in the analysis of Bristol's road network is to identify the edges corresponding to existing cyclepaths. The OSM tags of interest are 'Highway', 'Cycleway' and 'Bicycle', although not all of these are present for every way in the network. Our conditions for a way to be identified in the cycle network are: (a) its 'Highway' tag takes the value 'cycleway'; or (b) its tag set includes 'Cycleway'; or finally (c) its tag set includes 'bicycle' which takes the value 'designated'.

These criteria give rise to a binary indicator variable $\chi_{i,j}$ for the edge connecting nodes i and j which takes values,

$$\chi_{i,j} = \begin{cases} 1, & \text{if } (i,j) \text{ is a cycle path,} \\ 0, & \text{otherwise.} \end{cases}$$

The network with edges taking value $\chi_{i,j} = 1$ highlighted can be seen in Figure 2.1 with a zoomed in section showing the fine detail.

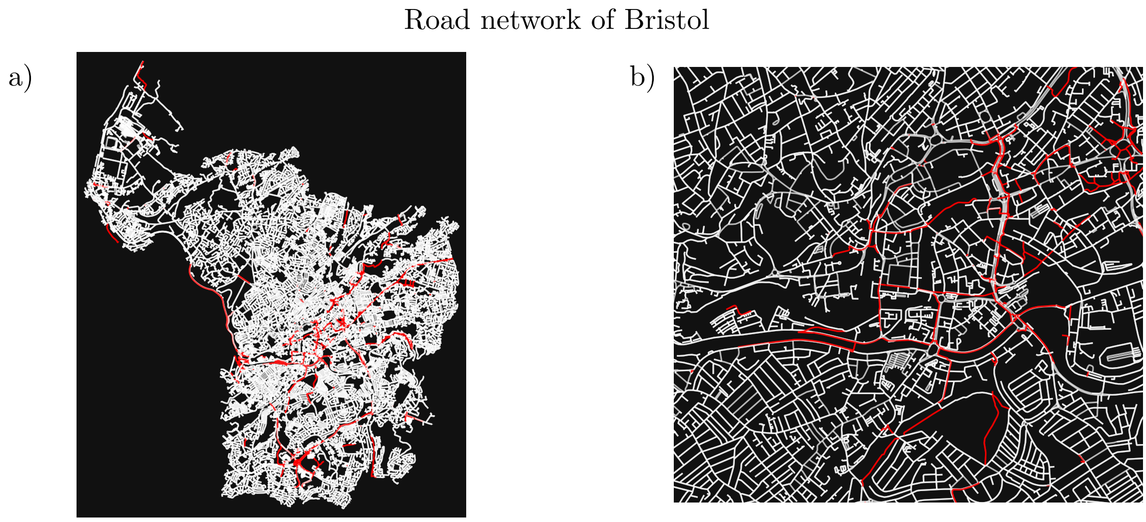

Figure 2.1: Road network of Bristol with edges (i, j) whose $\chi_{i,j} = 1$ highlighted in red. A section in the centre of a) is seen zoomed in b) to show the finer detail of the network.

The edges (i, j) of G are weighted to represent the cycled length $l_{i,j}$ of the edge and a parameter $p_{i,j}$ which represents the impedance to cycling. The value of $p_{i,j}$ can in principle take into account lots of real world factors such as: hills, speed limit of roads etc. The analysis presented here will take a very simplified approach by setting

$$p_{i,j} = 1 - \chi_{i,j}.$$

Each putative cyclist k should also have a personal parameter ω_k which describes their propensity to cycle on ways without cycling infrastructure.

From this the effective length of the edges in the graph G can be modelled as

$$\hat{l}_{i,j} = l_{i,j}(1 + \omega_k p_{i,j}), \tag{2.1}$$

meaning that an edge without cycling infrastructure has a larger perceived length than its true length according to the cyclists disposition. The idea then is that if a given cyclist tries to minimise the effective length of their route, they will prioritise routes that follow cycling infrastructure, provided the detour is not too large.

The literature suggests that the majority of cyclists prefer to cycle on designated cycling infrastructure [4]. This means that ω_k should be strictly positive. The larger the value given to ω_k, the larger the penalisation of routes with no cycling infrastructure. A learner cyclist k' should have a large value for $\omega_{k'}$, as it is highly unlikely that a learner would want to cycle without dedicated infrastructure.

2.3 Uniformly random OD pairs

Naturally the simplest way of loading the network with cycling demand is to sample OD pairs randomly from the set of all nodes within the network. The next step is to compute the shortest

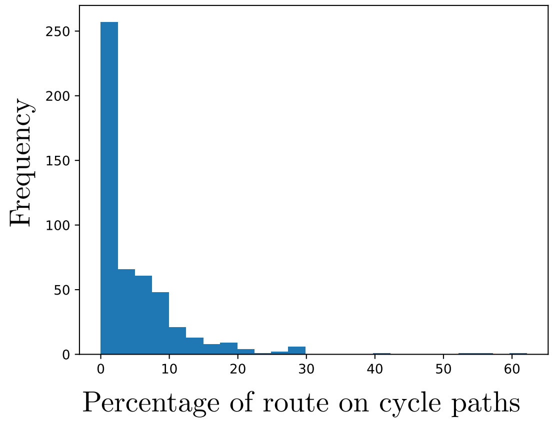

Figure 2.2: Histogram showing the percentage of routes taken on cycle lanes within shortest routes using uniform random OD selection, with parameter $\omega_k = 0$ so that streets with no cycling infrastructure are not penalised.

route between each OD pair. At this stage we suppose that $\omega_k = 0$ to model the case where cyclists have equal propensity to cycle on roads with or without cycling infrastructure. We compute 500 such shortest routes and the percentage of time spent on cycle infrastructure is found for each one, see the distribution in Figure 2.2.

Ignoring the huge peak at 0% the proportion of route length spent within the cycle network seems to decay exponentially, with the vast majority of routes spending less than 15% of their length on cycle infrastructure. This result makes sense given the unconnected nature of Bristol's cycle network, which does not allow for whole routes to use cycling infrastructure. The peak around 0% is due to the random selection of OD pairs giving rise to extremely short routes in areas of Bristol with no cycling infrastructure, so it is impossible for the shortest route to contain cycle lanes. In this computation the mean number of edges used in shortest paths is 120 which, when the network contains $56,000$ edges, is conducive to a large number of short routes. Another explanation for this peak is simply that the cycle network in Bristol is so sparse that most shortest routes cannot use cycling infrastructure. Some users may be able to increase their percentage of time on cycling infrastructure by allowing for not strictly shortest routes.

2.4 Comparison to second bounding geography

A second bounding geography, Amsterdam is also presented in order to compare summary statistics on a more highly connected cycle network. In order to control changes to the model, ω_k is left as 0 in the case of Amsterdam, to see whether connectivity or propensity to cycle on roads has a larger impact on the model result. Amsterdam is selected as a bounding geography not just for its much more comprehensive cycle network but also for its higher quality OSM data. Amsterdam's higher degree of connectivity should yield a higher percentage of time spent on cycle paths within shortest routes.

Figure 2.3 shows that by decreasing a cyclist's propensity to cycle on roads without infrastructure, we do in fact increase the percentage of length routes spend on cycle paths. However, this effect is small in comparison to using a different bounding geography with a more highly con-

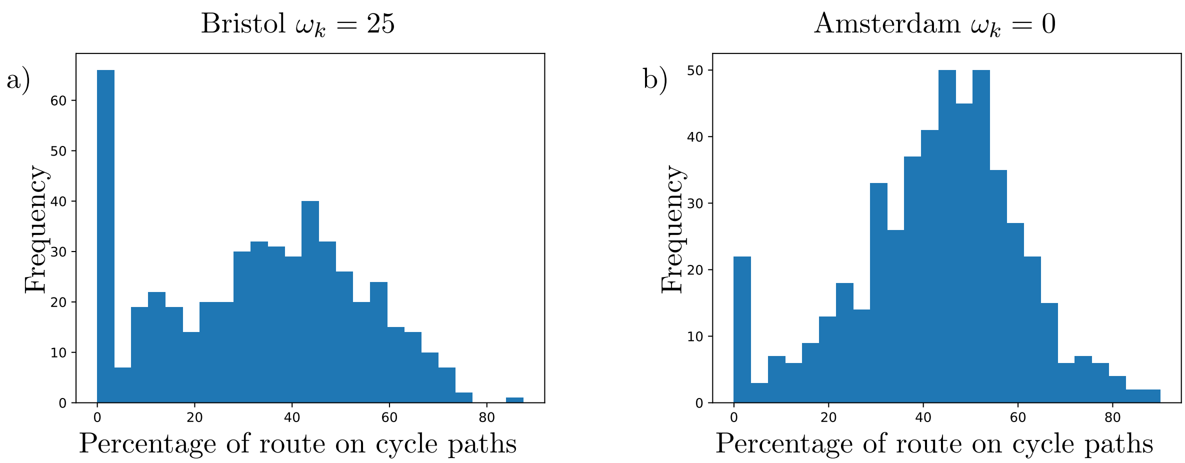

Figure 2.3: Histograms showing that both increasing ω_k in Bristol a) and the more connected city of Amsterdam (with $\omega_k = 0$) b) have a similar percentage of routes on cycle infrastructure.

nected cycle network. It is the case that in Amsterdam routes use more cycling infrastructure than Bristol even with $\omega_k = 0$. The issue with the outlier peak at 0% is still somewhat present, but in the case of Amsterdam almost certainly due to a small amount of very short routes that cannot feasibly use cycling infrastructure. Further computations (not presented here) have shown that as ω_k is increased for cyclists in Amsterdam, most routes are nearly 100% on cycle paths.

After restricting the model to choose shortest paths with at least 50 edges, the results obtained from the Amsterdam network with an $\omega_k = 0$ yield the same result but without a large peak at 0%. This confirms that this peak is caused by very short routes that simply cannot use cycling infrastructure. The problem is much larger in the Bristol network due to its low degree of connectivity, meaning paths would have to use almost the entire network to guarantee the use of cycling infrastructure.

Chapter 3

Upgrading the Network

Clearly Bristol's cycle network has a long way to go to reach the connectivity, and hence user satisfaction, of a city like Amsterdam. The goal of this project is to improve Bristol's cycle network, by building new cycle paths to ultimately yield a more connected network. In reality councils have a limited budget to build new cycling infrastructure, and wish to design a network that fits within that budget whilst maximising some factor of cyclist satisfaction.

Clearly the true optimisation problem is computationally infeasible on a road network the size of Bristol's. This chapter details a heuristic approach to approximate the optimal network design problem. The approach taken in this chapter is summarised in Figure 3.1. The proposed incremental heuristic has some key parameters detailed in Table 3.1.

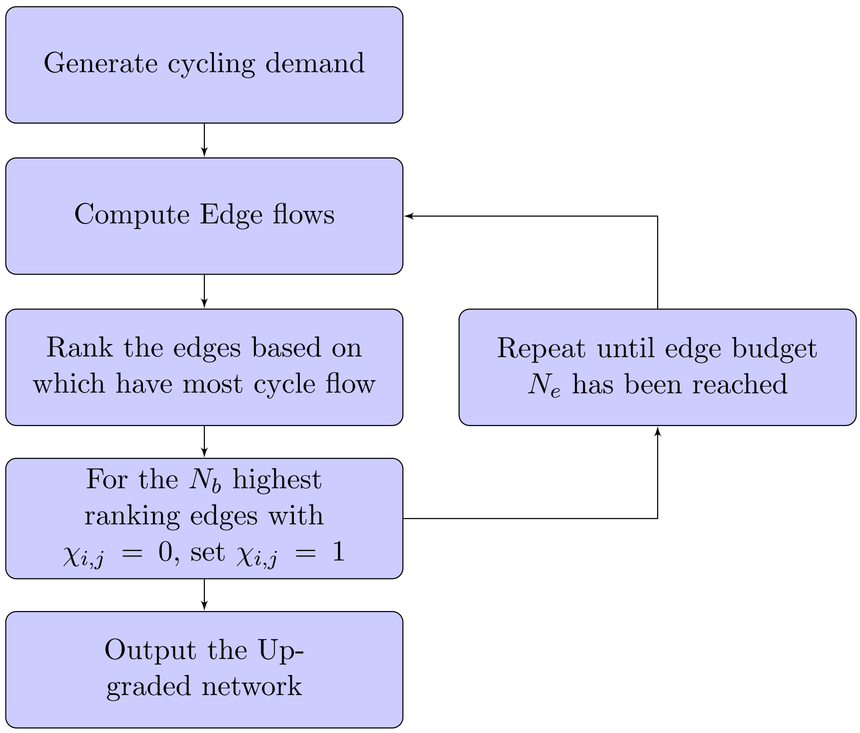

Figure 3.1: Summary of the heuristic upgrade approach proposed in chapter 3.

Key parameters	
Parameter	Definition
N_e	The total number of edges to be upgraded. In most cases 2000 to double the amount of infrastructure in Bristol.
N_t	The number of simulated cycling trips in a batch. This value should be large enough to avoid sampling bias but not too large as to focus demand on the centroids.
ω_d	The propensity to cycle factor for a 'design' cyclist. This is unique for the upgrade and need not be the same value in test cases.
B	The number of batches required to reach the edge budget with the current N_b.
N_b	The number of edges upgraded in a batch. This value is compute from N_e and B.

Table 3.1: Table of the key parameter for the heuristic proposed in Figure 3.1

This chapter will first detail the methodology for loading the Bristol network with cycling demand and ranking the network edges based on flow of this demand. We then move on to some computation specific to the Bristol network to show the impact of the proposed upgrade heuristic. Finally we discuss the importance and impact of a couple of the algorithms key parameters B and ω_t.

3.1 Generating cycling demand

To load the edges in Bristol's road network we need a realistic demand model. To find this empirical data is used. Bristol, like any large city, owes most of its traffic flow to commuters. If we can simulate commuter journeys and assume they can be made by bicycle we will get a more realistic model for the flows of cyclists through Bristol. This means that census commuting data can be used to predict cycle flows. Unfortunately the most recent census data available is from 2011 and therefore, is not entirely accurate to today's commuting trends. The data being used is at lower super output area (LSOA) level as this gives 263 zones in Bristol for flows to be generated between.

Each LSOA is modelled as its centroid, that is we assume all cycle flow comes from the centre of the LSOA. The data-set contains the longitude and latitude coordinates for each centroid, so the first step is to map these to nodes in the python network structure. The restulting centroid locations are plotted in Figure 3.2.

The commuting data is represented in the flow matrix

Figure 3.2: The map of Brsitol's road network with nodes representing LSOA centroids highlighted in red. WORKING ON THIS FIGURE NEED TO MAKE THE RED DOTS MORE VISIBLE.

$$F = \begin{pmatrix} f_{1,1} & f_{1,2} & \cdots & f_{1,263} \\ f_{2,1} & f_{2,2} & \cdots & f_{2,263} \\ \vdots & \vdots & \ddots & \vdots \\ f_{263,1} & f_{263,2} & \cdots & f_{263,263} \end{pmatrix}, \tag{3.1}$$

where elements represent the amount of directed commuting flow between the two centroids represented by the elements index.

It's worth noting at this point that by taking the assumption of all LSOA flow being generated at its centroid, there is no way to model flow within the LSOA so we set all diagonal elements

$$f_{i,i} = 0. \tag{3.2}$$

This matrix allows us to compute a probability vector

$$P(f_i) = \frac{1}{\sum_{i=1}^{263}} \begin{pmatrix} f_{i,1} & f_{i,2} & ... & f_{i,262} & f_{i,263} \end{pmatrix}, \tag{3.3}$$

corresponding to each centroid i to represent the probability of a journey that is initialised at i terminates at any other centroid j.

To load the Bristol network with cycle flow, journeys are initialised at a random origin centroid and the destination is chosen by sampling from all possible destination centroids using the probability vector $P(f_i)$. The shortest route between the O, D pair is then computed using the A* shortest route heuristic with edge lengths adjusted as detailed in section 2.2.

After simulating N_t cycling journeys on the Bristol road network we can form the edge flow matrix

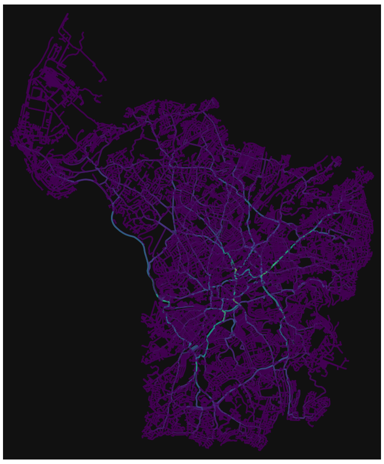

Figure 3.3: The map of Bristol's road network with edges coloured based on their $h_{i,j}$ flow value.

$$H = \begin{pmatrix} h_{1,1} & h_{1,2} & \cdots & h_{1,n} \\ h_{2,1} & h_{2,2} & \cdots & h_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ h_{n,1} & h_{n,2} & \cdots & h_{n,n} \end{pmatrix}, \tag{3.4}$$

where $h_{i,j} \in [0, N_t]$ is the number of simulated cycle trips that pass through the edge (i, j).

Figure 3.3 shows the Bristol road network with edges coloured corresponding to their flow in a trip simulation of size $N_t = 10,000$. It shows that under this LSOA model with a penalty factor of $\omega_k = 2$, cyclists tend to travel along major spokes with the highest flows being in the city centre which makes sense with a commuting model as most commuters travel to work in the city centre.

3.2 Experiments on the Bristol Network

To assess the user experience on an output upgraded network the scoring metric used in chapter 2 is used. The methodology is as shown in Figure 3.4

This distribution can then be compared with the distribution resulting from upgraded networks using a variety of different parameter sets.

The proposed upgrade algorithm in Figure 3.1 suggests an incremental approach by upgrading the edge budget in smaller batches. However, this is more computationally expensive than upgrading them all at once. This is because a new set of journeys must be generated for every batch. The idea is that with a batch process, cycling on paths generated in the previous batch is incentivised in the trip generation of the current batch, leading to more connected networks being suggested.

Figure 3.5 shows the output networks for the two extreme cases of this methodology. The one-shot network is computed by upgrading all 500 edges in one batch after simulating 500 trips. Whereas, the single-edge network is computed by upgrading only a single edge for each batch of 100 simulated trips. The one-shot uses a larger amount of simulated trips simply to account for sampling error which is not an issue for the single-edge method as trips accrue over batches.

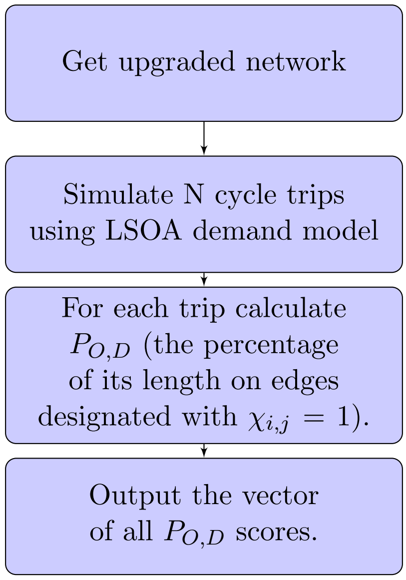

Figure 3.4: Computation of the scoring metric.

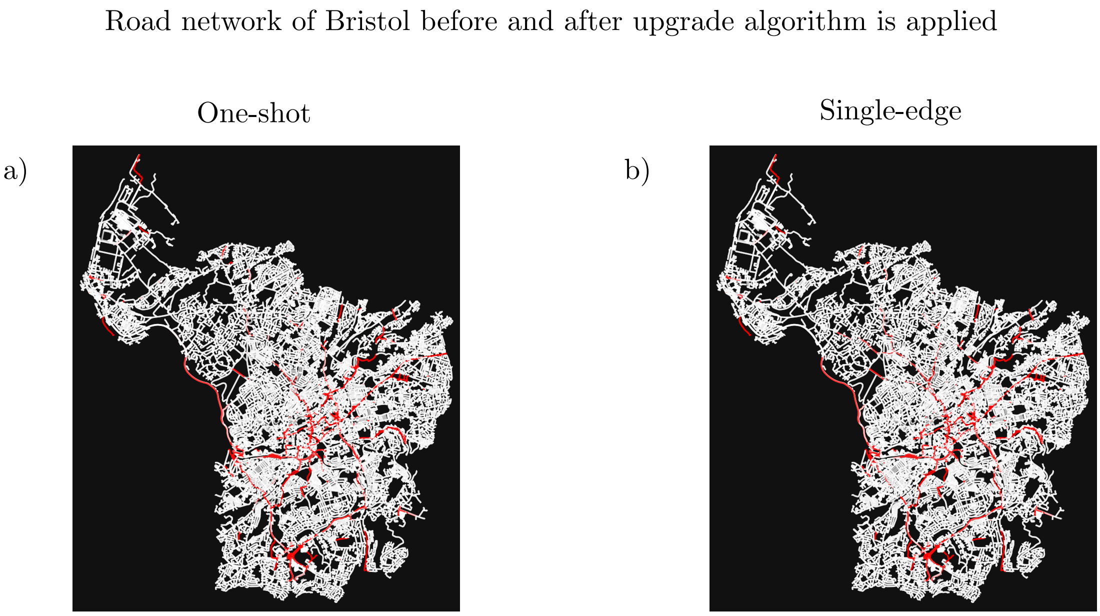

Figure 3.5: Road network of Bristol with edges (i, j) whose $\chi_{i,j} = 1$ highlighted in red. a) shows the network after a one-shot process has been applied to upgrade $N_e = 500$ edges and b) shows the result of a single-edge scheme upgrading $N_e = 500$ edges also.

Unfortunately, there is not much of a visual difference between the two extreme case networks. There is also very little difference seen in Figure 3.6 with the incremental process only having a slightly higher average $P_{O,D}$ score. This chapter goes on to experiment with difference batch sizes to see if anything more can be learnt about this parameter.

The batch process then, is a compromise between the computational efficiency of the one-shot approach and the ability for the single edge approach to model a test cyclists response to increased infrastructure. Hopefully in some network cases this will improve the connectivity of output networks. Figure 3.7 shows an upgraded network found by implementing a batch case

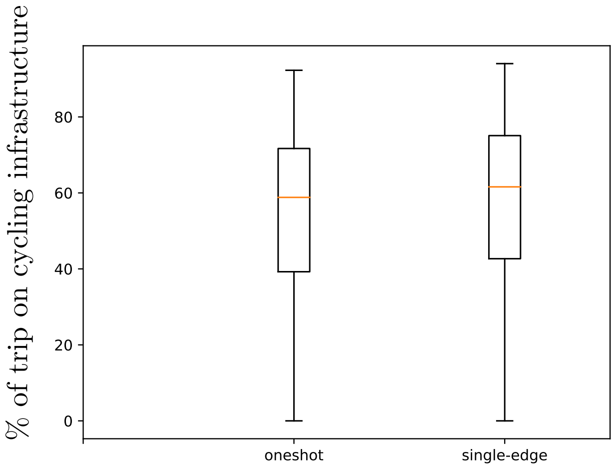

Figure 3.6: The difference in $P_{O,D}$ scores for the Bristol network upgraded in one batch and upgrading one edge per batch.

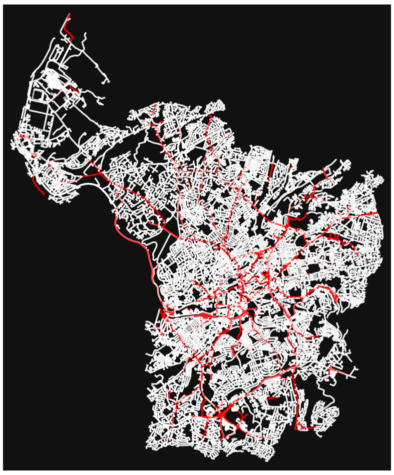

Figure 3.7: The output upgrade network for a batch process with parameters $N_e = 2000$, $N_t = 250$, $B = 20$ and $\omega_d = 2$.

of the heuristic.

In Figure 3.7 it is clear that the main spokes of the road network that are highlighted with high flow counts in Figure 3.3 are all upgraded with cycling infrastructure if they did not already have indicator $\chi_{i,j}$ this makes sense as these edges are the main thoroughfare for cyclists and will be ranked highest at the end of each batch's trip simulation. Then, once the main spokes have been upgraded the algorithm starts to fill in the lower demand areas that connect LSOA centroids to the key network spokes. One example of this effect is seen in the top of the figure. Cycling infrastructure is suggested into this area where demand is not particularly high but, as seen in Figure 3.2 there are LSOA centroids there.

When tested as detailed in Figure 3.4, Figure 3.8 is computed. It shows a large increase in the average percentage of their trip that cyclists spend on designated infrastructure. The current

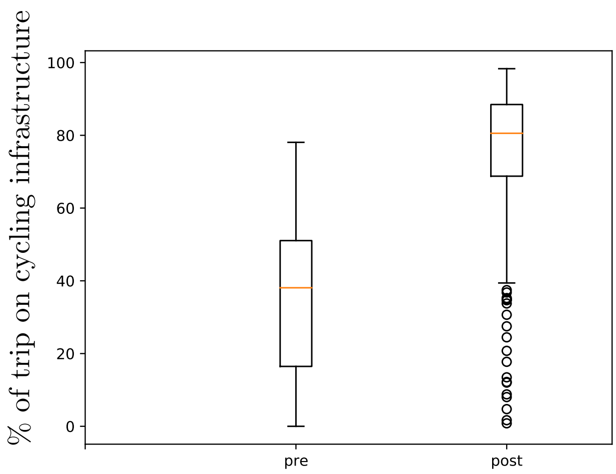

Figure 3.8: The difference in $P_{O,D}$ scores for the Bristol network now vs the proposed upgraded network using the proposed heuristic.

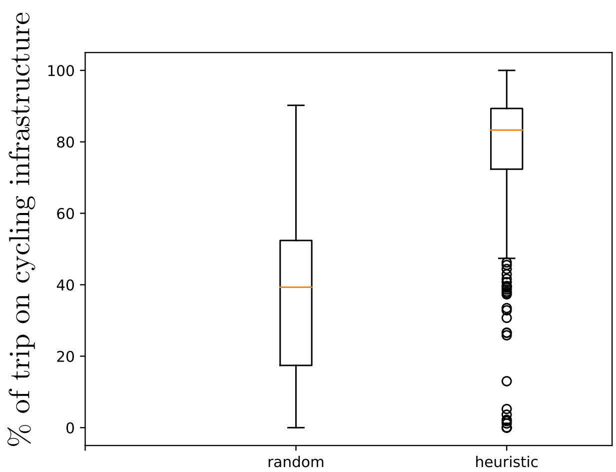

Figure 3.9: The difference in scores for two upgraded networks, both upgrade 2000 edges but one randomly selects edges whereas the other uses the heuristic proposed in this chapter.

Bristol cycle network achieves and average of 34% of journey length on cycling infrastructure. However, the proposed upgraded network achieves an average of 75% of journey length on cycling infrastructure. There are some outlier cyclists in the post-upgrade distribution, these trips are most likely due to test journeys being generated between two LSOA centroids that are very close together with low design demand between them and hence little cycling infrastructure has been generated along the shortest path from origin to destination.

The value N_e in this case is 2000 which means that after the upgrade has been performed there is more than twice the edges with indicator $\chi_{i,j}$ than in the original network. This will obviously yield an increase in the percentage of the length of cycling trips on designated infrastructure.

Figure 3.9 clearly shows that the heuristic proposed in this chapter vastly out performs a random process in terms of cyclist satisfaction. The random process, by doubling the amount of edges with indicator $\chi_{i,j}$ only achieves an average of 37% of the length of journeys on infrastructure

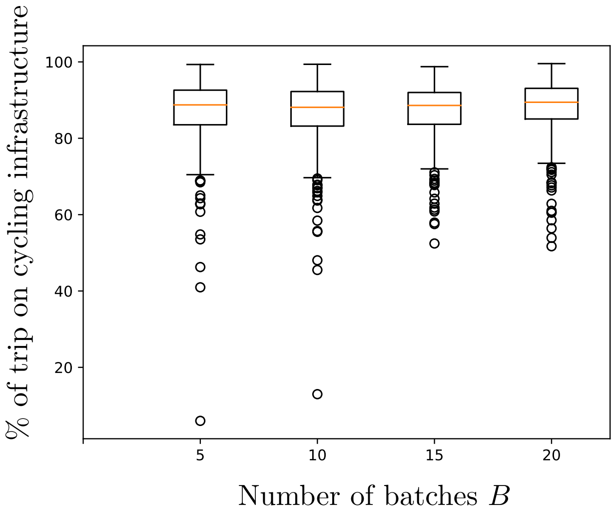

Figure 3.10: Figure showing that there is no change in $P_{O,D}$ scores for a variety of numbers of batches. This may be the case specifically for the Bristol network due to strong demand spokes being generated with relatively low demand.

a 3% gain over the original network. In contrast the upgrade algorithm proposed here yields a massive improvement with 78% of the length of journeys using designated infrastructure by upgrading the same number of edges.

3.3 Varying batch size

One of the key parameters of the upgrade algorithm proposed in Figure 3.1 is that of batch sizing N_b this parameter is not set but is interpreted from the parameter B, the number of batches and N_e, the number of edges to be upgraded overall. Figure 3.10 shows the resulting $P_{O,D}$ distributions from a sweep of values for B and hence N_b, in our test case of upgrading $N_e = 2000$ edges. It shows there is actually very little sensitivity to the number of batches in the case of the Bristol network, meaning that we can capitalise on the computational saving of reduced batch numbers. This low sensitivity is probably due to the well defined demand spokes seen in Figure 3.3 where these spokes have such heavy flow counts compared to other areas of the network they will always be prioritised for upgrade irrespective of batch sizing. It is possible that other network cases with less well defined demand spokes may have higher sensitivity to the batch size parameters, but this project focuses on the case study of Bristol so none are presented here. While it may seem that small batch numbers is the better choice from a computational stand point, later in this chapter we will discuss an alternative interpretation of the batches which will benefit from using a larger amount of batches.

3.4 varying ω

So far the value of ω has been fixed between upgrading and testing the network, that is to say $\omega_d = \omega_t$. However, it is worth exploring the impact that designing a network for a cyclist with a given propensity ω_d has on the resulting networks fitness for a cyclist with a different propensity

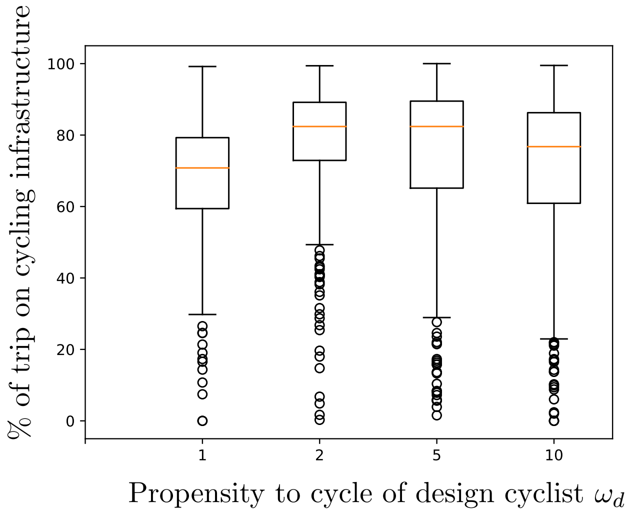

Figure 3.11: Comparing the $P_{O,D}$ distributions for a test cyclist with propensity to cycle $\omega_t = 2$ for a set of design cyclists.

ω_t. For example, if we design a network with cyclists who don't have much preference between cycling on and off designated infrastructure will this have an impact on the experience of a cyclist with a much higher penalty factor in the testing stage. Conversely, if a cyclist has a low penalty factor in the test stage will they prefer routes that are off the cycling infrastructure as they are more direct than the continuous cycle paths desired by more cautious riders.

Figure 3.11 gives the $P_{O,D}$ distributions for four different network designs with four different design cyclists $\omega_d \in [1, 2, 5, 10]$. The networks are then all tested with the same test cyclist with propensity $\omega_t = 2$. It can be seen that whilst all of the upgraded networks present a marked improvement over the original Bristol network, a cyclist who shares the same propensity as the design cyclist has the best experience. In this case the network with $\omega_d = 2$ has both the highest average $P_{O,D}$ score but also the least spread suggesting a better user experience.

Naturally the next step is to fix the design cyclist at $\omega_d = 2$ and see the effect that varying the propensity of the test cyclist in the same range $w_t \in [1, 2, 5, 10]$. Figure 3.12 demonstrates that the propensity to cycle of the test cyclist has a huge impact on the distribution of $P_{O,D}$ scores for the upgraded network. In the case $\omega_t = 1$ the test cyclist takes the true shortest route between their origin and destination which may not fall on suggested infrastructure that tends to be built onto existing infrastructure. We see also that as ω_t is increased the cyclist clearly is happy to take a longer detour to stay on designated infrastructure, with very large ω_t value cyclists spending almost 100% of their journey on cycle paths. This finding suggests that as long as upgraded networks have a connected path between origin destination pairs nervous cyclists will travel a longer distance to use them. This is not true in reality which is why physically a propensity to cycle of 10 is infeasible. the test cyclist with $\omega_t = 2$ seems the most realistic as most journeys do use a lot of designated infrastructure but when trips become too long by trying to sty within the cycle network the cyclist defaults to the shortest path.

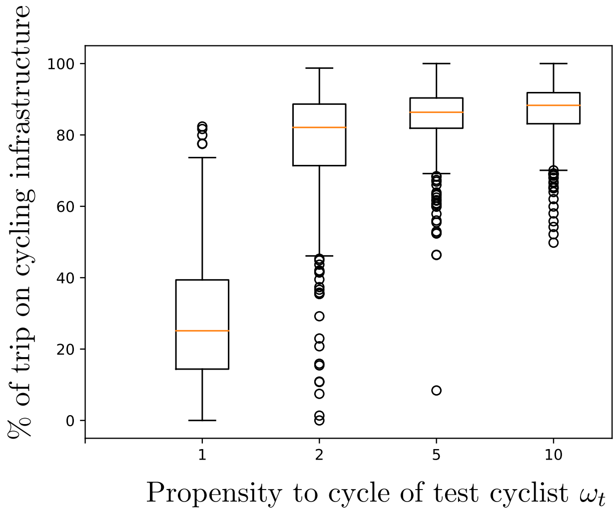

Figure 3.12: Comparing the $P_{O,D}$ distributions for a design cyclist with propensity to cycle $\omega_d = 2$ for a set of test cyclists.

3.5 Physical meaning of batches

The batches in the proposed upgrade algorithm are until now a synthetic time-step to encourage connected networks. However, there is a very nice parallel between batches and individual cycle schemes. In reality the council, or anyone else building new infrastructure, want to stagger the implementation of new cycle schemes to gauge the true impact that new infrastructure has on cycling demand. In this model it is appropriate to view each batch as a new scheme perhaps implemented year on year. The characteristic of growing connected cycle paths then follows naturally. If there is new infrastructure, naturally simulated cyclists with high penalty factor ω_d will want to use it to service their journeys. This would in reality then lead to higher flows on roads at the end of this new infrastructure, so long as it doesn't terminate in an LSOA centroid. Therefore the upgrade algorithm should and does prioritise these edges for upgrade in the next batch.

It is in the councils best interest to minimise their cost whilst maximising user experience. Therefore, it is worth exploring the improvement batch on batch for these incremental cycle schemes. If there is a huge improvement to the $P_{O,D}$ scores in the first couple of batches and only very small improvements after, it may not be worth investing in further improvement. This information is invaluable to someone implementing new cycle schemes as it could reduce costs drastically.

Figure 3.13 shows the score change of improvements made to the Bristol network throughout all 20 batches of the upgrade algorithm. The plot shows that most of the upgrade to the average $P_{O,D}$ scores occurs within the first few batches. It is worth investigating which edges get upgraded first to see if there is some strategy for prioritisation to be found.

Figure3.14 shows the order in which the upgrades are performed. Edges are coloured base on which batch they are upgraded in with darker edges being upgraded first. This plot indicates that the major infrastructure spokes seen built into the top and bottom of the network in Figure 3.7 are built within the early batches and suggests these are the streets the council should prioritise

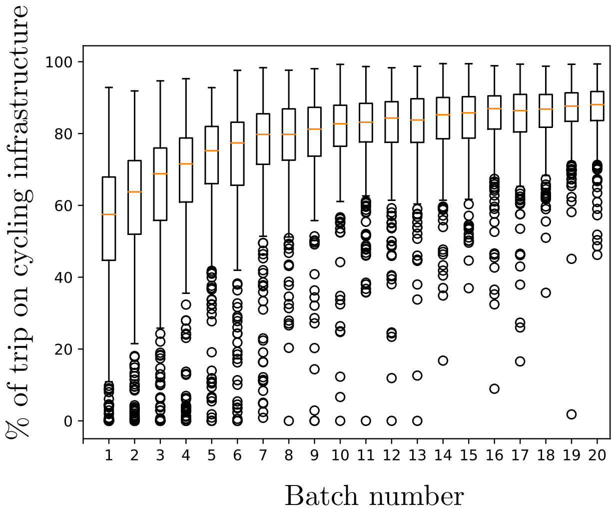

Figure 3.13: Graph showing the improvement in $P_{O,D}$ scores throughout the batches of a 20 batch upgrade of the Bristol network. It shows that most of the upgrades are made in the early batches.

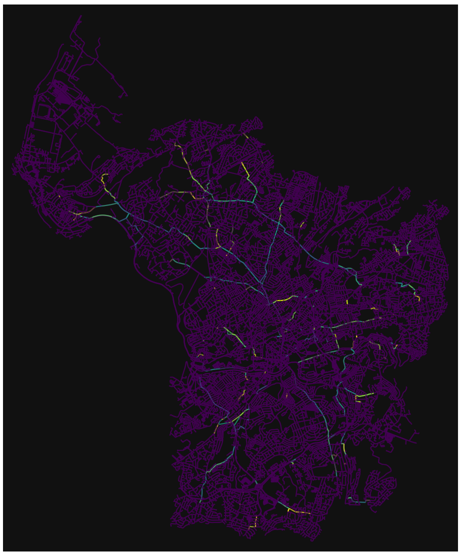

Figure 3.14: The map of Bristol's road network with upgraded edges coloured based on which batch they are upgraded in. The brighter the edge the later it is upgraded.

upgrading first. These upgrades correspond to the significant increase in average $P_{O,D}$ seen in Figure 3.13. The more yellow edges are upgraded in the last couple of batches and tend to be at the end of long spokes, indicating that the algorithm is upgrading in such add to existing infrastructure batch by batch. Finally some of the later batches begin to fill in small gaps near the locations of LSOA centroids which correspond to the small gains in $P_{O,D}$ spread seen in Figure 3.13.

Chapter 4

Optimal Networks

Bibliography

[1] C. Allan. Cycling UK's cycling statistics. `https://www.cyclinguk.org/statistics`. Accessed 2020-12-02.

[2] Conclusions from the National Travel Survey 2019. `https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/905887/national-travel-attitudes-study-wave-3.pdf`. Accessed 2020-11-09.

[3] Conclusions from the National Travel Attitude Survey wave 3 2019. `https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/905887/national-travel-attitudes-study-wave-3.pdf`. Accessed 2020-11-09.

[4] J. Dill and T. Carr. Bicycle commuting and facilities in major US cities: if you build them, commuters will use them. *Transp. Res. Rec.*, 1828(1):116–123, 2003.

[5] G. Topham. Chancellor announces £27bn for roadbuilding in budget. `https://www.theguardian.com/uk-news/2020/mar/11/chancellor-announces-27bn-for-roadbuilding-in-budget`. Accessed 2020-11-29.

[6] R. Cervero and M. Duncan. Walking, bicycling, and urban landscapes: evidence from the San Francisco bay area. *Am. J. Public Health*, 93(9):1478–1483, 2003.

[7] D. Milakis and K. Athanasopoulos. What about people in cycle network planning? Applying participative multicriteria GIS analysis in the case of the Athens metropolitan cycle network. *J. Transp. Geogr.*, 35:120–129, 2014.

[8] J. Larsen, Z. Patterson, and A. El-Geneidy. Build it. But where? The use of geographic information systems in identifying locations for new cycling infrastructure. *Int. J. Sustain. Transp.*, 7(4):299–317, 2013.

[9] A. Mauttone, G. Mercadante, M. Rabaza, and F. Toledo. Bicycle network design: model and solution algorithm. *Transp. Res. Rec.*, 27:969–976, 2017.

[10] P. Freestone. Personal communication. October 2020.

[11] R. Lovelace, A. Goodman, R. Aldred, N. Berkoff, A. Abbas, and J. Woodcock. The propensity to cycle tool: An open source online system for sustainable transport planning. *J. Transp. Land Use*, 10(1):505–528, 2017.

[12] R. Lovelace and J. Talbot. Rapid cycleway prioritisation tool: Technical specification. `https://www.cyipt.bike/rapid/`. Accessed 2020-12-08.

[13] DFT. The cycling infrastructure prioritisation toolkit. `https://www.cyipt.bike/report1/`, 2018.

[14] G. Boeing. Osmnx: New methods for acquiring, constructing, analyzing, and visualizing complex street networks. *Comput. Environ. Urban*, 65:126 – 139, 2017.