Department of Computer Science

Submitted in part fulfilment for the degree of MEng.

How important is a city street?

Harry Kelly

1st May 2022

Supervisor: Richard Wilson

Acknowledgements

I would like to thank my supervisor for all the help he gave me throughout this project.

Contents

List of Figures

Executive Summary

This project defines and evaluates an approach to identify streets within an urban network that are best suited to pedestrianisation. Reduced car traffic within a city centre is a desirable goal due to the reduced impact on both the environment and public health, as well as the benefits of more freedom for pedestrians and cyclists available in quieter and pedestrianised streets. By pedestrianising streets, car travel can be discouraged or re-routed away from the city centre, as has already been done by cities such as York [1]. However excessive road closures can cause congestion, so it is important to strike a balance when closing roads, and to select streets that would most benefit whilst having the least negative impact.

By making use of the eigendata centrality [2], nodes within an urban network can be assigned an importance based on both network topology and spatial data. This can be used to find nodes, and therefore edges, that are both central to the network and adjacent to businesses and workplaces - identifying streets that act as hubs of economic and social activity. By ranking streets based on their eigendata centrality, a list of candidate streets for pedestrianisation can be created.

This list can be further filtered by the application of a simple travel time model of traffic flow used to estimate the impact of closing each street to traffic. By applying a shortest path algorithm for a large number of generated journeys, using an estimate of time to travel down a street as the weight, an estimate of time to travel a journey on the network can be found. This algorithm can be applied to many variations of the network to measure the impact of closing a road to traffic, in order to identify those with a small impact. The process can also be used to assess the impact closing all of the identified streets at once would have on the network, and decide if it makes for an appropriate pedestrianisation plan.

By applying this method to the case study of York, I found that the eigendata centrality works well for identifying hubs of activity in an urban network, and can be used to identify important streets. Additionally, I have found that application of the method on a network containing streets pedestrianised in the real world can result in the selection of those pedestrianised streets. However the overall results of this process are not ideal, as outside of a city centre they tend towards selecting any roads on the outskirts without

a clear need to be pedestrianised. This method can only be used as an additional tool alongside an experienced person able to provide intuition and insight - it is not appropriate as a standalone method of identifying candidates currently.

Throughout this project I made use of the open data source OpenStreetMap. I followed all relevant policies when using data from and contributing to OSM, which is important as it is used in many important works and projects. Furthermore, as this project could be used by town or city councils as part of city planning, I have made the weaknesses of this method clear throughout the project, and recommend that should this ever be applied it is used as an additional tool by those with knowledge in the field, as it is in no way a substitute for an experienced professional.

1 Introduction

Pedestrianised streets and car-free areas are of great importance in a city centre, providing a safer and more pleasant environment for pedestrians and cyclists. With a shift away from cars and towards public transport being a necessary step to reduce a city's impact on the environment, as well as walking & public transport often being more desirable ways of travelling into and within a city centre, creating more pedestrian spaces is an action city councils should be considering.

Selecting roads to pedestrianise can be difficult: you must consider the need / appeal for the street to be pedestrianised, the impact this would have on its neighbouring streets, and how it would affect the road network as a whole. These problems can be addressed thanks to the availability of extensive data on a city's highways, their classifications (pedestrianised, primary road, etc) and speed limits available from online maps. Details on a city's buildings, businesses, and residential areas are all also available via similar providers. In some cases cities also provide traffic survey information, which can be used in complex traffic models.

In this project I intend to lay out and assess one possible approach to the problem of pedestrianising streets, which can be used to better inform those making these decisions. This can be achieved using publicly available map data, a centrality measure to identify high-importance streets, and a means of analysing journey time distribution on an urban network based on a shortest path algorithm. The results of this process will then be analysed in order to judge if this approach is an appropriate means of assessing a road's eligibility for pedestrianisation.

2 Literature Review

2.1 Digital Maps of Towns & Cities

The highways of towns and cities form an urban network made of highways (roads, footpaths, cycle paths, etc) and intersections. These urban networks can be modelled as two types of graphs: primal and dual. Primal graphs use edges for highways and nodes for intersections, and are well suited for spatial relationships such as maps. Dual graphs represent highways as nodes and their intersections as edges, and are not confined by geographic constraints, as such they are better used for understanding abstract relationships such as social activity within an urban network. As this project will be focusing on spatial relationships, I will make use of primal graphs throughout.

A graph is planar if when drawn on a 2-dimensional plane its edges only intersect at their endpoints. Most maps of large areas are non-planar due to tunnels and bridges, and most maps of city centres are planar as these features are not typically present. However, as I intend to focus on the City of York, which has a number of highways crossing without an intersection due to the city walls and bridges across the rivers (and therefore the highways following the river), my graphs throughout this project will be non-planar as to maintain these properties.

By representing urban networks as graphs, we can perform many kinds of analysis on a town or city centre.

2.2 Centrality Measures

Formalised by Freeman [3], centrality measures are ways to define the most important / central nodes in a network.

2.2.1 Degree Centrality

One example of a simple centrality measure is Degree Centrality. The principle behind degree centrality is that the most important nodes are those with the most connections. For each node the centrality is equal to the number of edges the node is connected to - the node's degree. This can help find highly connected nodes, such as a complex junction. If you wished to express the degree centrality of a node i, $DC(i)$, as a formula it would be expressed as follows:

$$DC(i) = \text{degree}(i)$$

2.2.2 Closeness Centrality

Another centrality measure is the closeness centrality, in which the principle is that the most important nodes are those that are closest to all other nodes. The centrality for each node is equal to the average of the lengths of the shortest paths from that node to all other nodes. The closeness centrality for node i, $CC(i)$, can be expressed as follows:

$$CC(i) = [\sum_{j \in N} d(i,j)]^{-1}$$

Where N is the set of all nodes in the network, and $d(i,j)$ is the distance between nodes i and j. This centrality measure is close to how people most commonly think of a city centre, where the most central intersection is one that is on average the closest to all others, which may often be some form of town square / geographically central point.

2.2.3 The Eigenvector Centrality

Proposed by Bonacich [4], the eigenvector centrality builds on the principles of degree centrality. Whilst degree centrality considers a node important if it is well connected, it can be argued that not every connection is equally important (consider a collection of well-connected nodes that are almost entirely isolated from the rest of the graph). The eigenvector centrality is built around the idea that a node is more important if it is connected to more nodes that are also important. The eigenvector centrality of node i, $EC(i)$, can be calculated by either of the following expressions:

$$EC(i) = \frac{1}{\lambda} \sum_{j \in M(i)} EC(j) = \frac{1}{\lambda} \sum_{j \in N} a_{i,j} EC(j)$$

Where λ is the eigenvalue, $M(i)$ is the set of neighbouring nodes of i, N is the set of all nodes in the graph, and $\boldsymbol{A} = (a_{i,j})$ is the adjacency matrix of the graph ($a_{i,j} = 1$ if node i shares an edge with node j, and $a_{i,j} = 0$ otherwise).

2.2.4 The Eigendata Centrality

The eigendata centrality was developed by Agryzkov et al. [2] as an improvement upon the eigenvalue centrality that incorporates the use of geo-referenced data to its calculation. The core principle of the centrality is that a node is important if it is linked to another important node by a road that has a large amount of data associated with it. This helps avoid a pitfall of eigenvalue centrality, where more connected but less developed areas of a city could be considered more important than more developed lesser-connected areas.

Given the following inputs, the eigendata centrality measure for nodes of a graph ($\vec{c}$) can be computed by the following process [2]:

Inputs:

- Adjacency matrix A - representation of the graph being analysed, elements indicate whether pairs of nodes are connected by an edge
- Data matrix D - Stores the data values corresponding to each node in the network. Entry d_{ij} contains the value of category j assigned to node i.
- Weight vector $\vec{v_0}$ - Length is equal to the number of categories of data included in analysis. Values are in range [0,1], to indicate magnitude when calculating importance values.

Process:

1. Construct data vector $\vec{v}$
$$\vec{v} = D \cdot \vec{v_0}$$
2. Normalise data vector
$$\tilde{v} = \frac{1}{\max_i \ v_i} \vec{v}$$

3. Construct weight matrix $\boldsymbol{W}$ as

$$w_{ij} = \tilde{v}(i) + \tilde{v}(j)$$

4. Compute base importance α

$$\alpha = \min(w_{ij})_{w_{ij} \neq 0}$$

5. Decide smoothing factor ϵ, where

$$\epsilon < \frac{1}{10}\alpha$$

6. Construct weighted adjacency matrix $\boldsymbol{A}^*$

$$\boldsymbol{A}^* = \boldsymbol{A} \circ (\boldsymbol{W} + \alpha \boldsymbol{J}) + \epsilon \boldsymbol{J}$$

   where $\boldsymbol{J}$ is the matrix of all ones, and $\circ$ is the element-wise multiplication operation

7. Compute dominant eigenpair of $\boldsymbol{A}^*$, $(\lambda_1, \vec{x_1})$, as per the eigenvector centrality

8. Calculate eigendata centralities

$$\boldsymbol{c} = \frac{1}{\lambda_1}[\boldsymbol{A}\vec{x_1} + \vec{x_1}]$$

2.2.5 Applying the Eigendata Centrality

A primal graph can easily be used to represent the layout of a city, and with access to geo-referenced data, a collection of tagged coordinates can be assembled to represent the cities locations and businesses of various categories. This is used to create a data matrix, where one axis contains the various categories of data (shops, restaurants, etc) and the other axis contains the ID of each node making up the graph of the city. The value for each cell of the matrix represents the number of data points of each category to have each node as their closest node. This data matrix can be used to create a weight matrix for the edges, such that an edge's weight is the total number of data points associated with the two nodes that the edge connects. Additionally, any edges with no data points associated with them are given a small base importance, as even if a street is empty it still has a small importance in its ability to connect places.

The centrality makes use of a smoothing factor to avoid rapid variations in results in order to better focus on diffuse centres of activity. However, the smoothing factor can be sent to 0 in order to identify areas that act as hubs/attractors. Agryzkov et al. recommended the centrality be used for modelling areas of commercial activity. Others have recommended it for assessing the impact of changes on a network. [5]

2.3 Macroscopic Traffic Flow Theories

Macroscopic traffic flow theories focus on the big picture: traffic intensity, road area, average speed. These theories are good for evaluating system-wide control strategies in urban networks. There are a variety of macroscopic traffic flow theories. Almost all of them are based on three core variables: speed, flow / volume, and concentration.

2.3.1 Speed and Flow Relation

It is generally understood that a greater flow of vehicles in a network results in a slower average speed of vehicles. [6] However, work by Godfrey [7] showed that after a certain maximum capacity of flow in an urban network, the average speed of vehicles remain mostly the same - instead of slower vehicles and a greater flow when adding traffic, we see similar values for both and a backlog of vehicles queuing at entrances to the network (such as car parks and homes). This means that attempting to assess a network based on average vehicle speed can be deceptive. It may appear making a change to a network does not affect it, and that the vehicles remain at the same average speed, however there may be a greater backlog before or after the change. A more sophisticated form of measurement is required.

2.3.2 Travel Time Models

Travel time models provide an overview of how a network is operating at specific times. [6] Vehicles are dispatched from a specified location in the network, and each vehicle's time and position can be noted at desired intervals. Whilst the idea of simulating a lot of traffic on individual journeys sounds like a good approach to assessing a network, these methods are limited by the fact they only look at one point, and therefore require a lot of repetition and resources to find any significant results.

2.3.3 Network Capacity Models

Smeed [8] wished to find how many cars could 'usefully' enter a city centre, and took the approach of focusing on both the capacity and the proportion of area in a city dedicated to roads. This found that a large amount of road is often unused due to uneven distribution, and that the amount of vehicles that can travel an urban network depends on their speed and is proportional to the area of usable roadway. This approach of finding network capacity may be prone to demanding the most possible road area, which would not be helpful for my purposes.

An alternate approach to measuring network capacity proposed by Loder et al.[9] makes use of Macroscopic Fundamental Diagrams. This approach requires that the MFDs are estimated using data from traffic sensors at points around the city. From there the network's critical point between free flowing vehicles and congestion can be estimated.

2.3.4 Two-Fluid Theory

Two-Fluid theory takes the approach of modelling vehicles as two types of fluid: vehicles that are moving, and vehicles that are stopped due to traffic conditions. The average speed of all vehicles is proportional to the fraction of vehicles moving, and the trip time per unit distance is linearly related to the stop time per unit distance. Two fluid theory has shown that an increase in concentration of traffic results in a decrease in average vehicle speed, and that there is a maximum possible traffic flow for both concentration of traffic and average speed of traffic - any increase or decrease in either variable results in only a decrease in traffic flow. [6]

2.3.5 Other Approaches to Traffic Flow

The Revised Monograph on Traffic Flow Theory [10] details many more approaches to traffic flow theory. Continuum flow models, much like two-fluid theory, models traffic in terms of fluid behaviour, and focuses on the overall statistical behaviour instead of interactions between individual particles / vehicles. Another approach detailed is in-depth traffic simulation, such that every vehicle is modelled as an independent entity traversing an urban network. These simulations can be complex, modelling lanes and intersections, and very computationally expensive. However, simulations on a bigger and less detailed scale do exist, and can share common features with macroscopic traffic flow theories.

2.3.6 Summary

Centrality measures are a means of assigning importance to nodes in a network. The eigendata centrality is one such centrality measure, in which both the topology of the network and the amount of spatial data assigned to each node affects the value; this makes it particularly useful for use within urban networks. This project will make use of the eigendata centrality to assign an importance to the nodes, and through that edges, of a network in order to identify central and economically active streets.

There exist a variety of ways to analyse traffic flow in an urban network, though most methods do not seem appropriate to apply in this project. Some require a large amount of computational power - simulating a large number of vehicles and their interactions or a fluid in a network are both intensive, and would require more resources than I have available in order to run these models over potentially hundreds of variations of a network. One simpler method uses techniques that directly oppose the goals of this project, by using the proportional area dedicated to roads as a key measure of a network's capacity. Though they fall into the category of too intensive, the idea behind travel time models - to measure network capacity based on a vehicles progress over a period of time - has the most potential to intuitively simplify to a method that I can apply. This project will use a simple travel time model that considers the time taken to complete a number of journeys on a network.

3 Methodology

3.1 Overview

I intend to model the City of York as an urban network and use data about business and other non-residential buildings to apply the eigendata centrality measure to identify which streets are the most important. From there I will explore the effect of pedestrianising some of the more important streets on the traffic capacity of the City. This will be done through a simple analysis using Dijkstra's Shortest Path Algorithm to estimate a total travel time for a number of journeys on both the original and reduced road networks.

I have chosen York as it is a smaller city with a distinct city centre area that already has lots of pedestrianised and semi-pedestrianised streets. This should lend itself well to fast computation times, and includes examples of streets already deemed important enough to be pedestrianised.

3.2 Representing a City as a Network

I will model the urban network of York as a non-planar graph, as there are a small number of overlapping non-intersecting highways (bridges) in the city. The edges of the graph will represent highways such as roads, footpaths, and pedestrianised streets. Nodes of the graph will represent the intersections of these highways. The coordinates of these elements will be equal to their latitude and longitude, to allow for the matching of building data to the streets of the graph.

The city will be modelled as a multigraph to preserve features such as streets that terminate at the same intersection (self-loops on nodes), and sets of streets that terminate at the same pairs of intersections (parallel edges). These features can be common in urban networks, such as self-loops in cul-de-sacs, or separate pedestrian, bike, or car routes connecting a pair of intersections. The graph will also be directional, to preserve features such as one-way streets, which are also common in York.

Figure 3.1: Map of York acquired

Blue: residential, Orange: commercial / workplace

3.3 Data Sources

My data source is OpenStreetMap (OSM). OSM was used successfully by O. Goldsmith to model the city of Portsmouth in his project "How important is a city street?" [5]. Goldsmith detailed that OSM has relatively good coverage of the UK, with varying levels of completeness and quality in different areas of the country. He also emphasised the importance of OSM's large number of high-quality developer tools and libraries, compared to alternatives such as Ordnance Survey. I intend to follow Goldsmith's lead and make use of OSM and the OSMnx Python package. As the administrative boundary of the city of York is incredibly large, and includes a large amount of rural areas, data will instead be imported using OSMnx for the area within 2km of the coordinates (53.9586012, -1.0807963). This acquired area will include the city centre and an appropriate amount of suburbs without becoming too large to process in a reasonable amount of time. The recovered map, along with the data for buildings and homes, can be seen in figure 3.1.

3.4 Data Processing

The graph of the city will be stored as a MultiDiGraph from the NetworkX python library, the format used by OSMnx. Isolated nodes will be removed, and the remaining nodes and edges will be re-indexed from their OSM IDs to a simple incremental index starting at 0.

Building data for businesses & public amenities will be converted from shapely points, polygons and multipolygons to lists of points and areas, by finding the centroid and area of the polygon & multipolygons. As centrality measures consider nodes, not edges, each building's point will be assigned to its nearest intersection, and inserted into a data map detailing the total area of buildings assigned to each intersection, in order to be used for the eigendata centrality. Similar data for residential buildings will be processed in the same manner, to be used in the traffic capacity analysis. The data map will be constructed as a python dictionary, with keys equal to the ID of each node, and values equal to the total area of all buildings for which the node is their nearest node.

3.5 Applying Eigendata Centrality

Using the processed data, I will apply the Eigendata centrality as detailed by Agryzkov et al. [2]. This will be achieved by using the data map to construct the normalised data vector (python list) required for the eigendata centrality. The adjacency matrix of the network will then be found using NetworkX [11]. The weight matrix will be represented as a 2-dimensional NumPy NDArray [12], where the value at index (i, j) is equal to the sum of the data vector values for indexes i and j. The weighted adjacency matrix will then be calculated by performing element-wise multiplication on the adjacency matrix of the graph and the matrix found by adding the base importance to every element of the weight matrix. The weighted adjacency matrix will then have the smoothing factor added to each of its elements. The dominant eigenvalue and eigenvector will be found using SciPy's eigendata function on the weighted adjacency matrix [13]. The eigendata centrality will then be found by performing matrix multiplication on the adjacency matrix and the dominant eigenvector, adding the dominant eigenvector to the result, and then dividing that by the dominant eigenvalue.

This will result in a centrality value for every node representing the importance of every intersection, where important nodes have both a high number of buildings and connections to other important nodes. These values can then be used to find an eigendata value for every highway, by taking the average eigendata value of the highway's pair of nodes - as an important edge is one that connects two important nodes.

3.6 Calculating Travel Time

With an importance value for the streets of York, I will then examine the effect of closing various roads belonging to the most important 2% of eligible roads (available to car traffic) on the time taken to travel various car journeys within the city, using a shortest path algorithm to estimate travel times. Service roads will be ineligible as they are either car parks or exist to provide vehicle access to a location, both of which should not be closed.

Ideally I would have liked to make use of more sophisticated or complex methods, but most approaches I have found are either only tangentially related to my aim of evaluating the change in capacity for an entire network when edges are removed, or seem beyond my capability to implement. The work on network capacity by Smeed [8] would be good for evaluating a network's capacity, however as the method evaluates based on the proportion of area dedicated to roads, it would judge any removal of roads harshly. On the other hand, the work by Loder et al [9] may yield good results, but as it is a very recent paper it makes use of complex methods I do not believe I could replicate with my current knowledge and abilities.

First, a copy of the highway graph will be created, containing only roads and pedestrianised streets (routes that cars could in theory use). As York has a lot of ginnels, alleyways, and streets a car could not use, it is important that these are removed when considering motor vehicles. Pedestrianised streets will be kept so that we can see if these streets are considered for closure / pedestrianisation.

1,000 journeys will be generated, by choosing weighted random origin and destination nodes. Origin choice will be weighted by the area of homes assigned to each node, and destination choice will be weighted by the area of commercial & public buildings assigned to each node.

Each edge on the road graph will be given a travel time estimation based on their length and speed limit. Each journey will then be evaluated based on Dijkstra's Shortest Path Algorithm using the travel times of the edges, and any journey that can't be completed will be noted. This process will be applied on an unchanged graph of the City, and then on a number of variants in which a single road from the top 2% of eligible roads has been removed, for every journey in every run of the process data will be collected on whether the journey could be completed and the journey's travel time. For the applications on the reduced networks every journey will also have the difference between its travel time and its travel time on the unchanged network recorded.

As it can be assumed that if it takes longer for vehicles to complete a journey there will be more vehicles in the network at any given time, we can use this impact on travel time as a proxy for congestion. The numerical value of the travel times under this approach will not be realistic, as they do not account for stopping at intersections or the snowballing impact of congestion, but as I intend to compare the difference in distribution of these values, I believe the numerical value itself does not matter so long as I can evaluate a difference in distribution.

Using the data collected about the impact of removing each road a set of roads will be found that, when removed, have impact on neither the proportion of completable journeys nor the travel times of any journey. A map will be created highlighting the roads belonging to this set of roads considered perfect for removal.

A second set of roads will then be created including those that, when removed, have no impact on either the proportion of completable journeys or the travel times of any journey, but not necessarily both. The journey travel time estimation process will be applied on a network with these roads removed. The percentage decrease in completable journeys will be calculated, and a histogram showing the cumulative travel time of all journeys on the new network compared to the original, and the increase in travel time to each journey caused by the change. A map will also be created of these roads. A final application will be performed on a network with all of the top 2% of roads removed, as a comparison point of the worst possible impact to the metrics.

This approach is still rather simple - it only considers the time it takes a series of single vehicles to travel the optimal route through an empty network, and cannot account for the increase in congestion that may be created by removing a road. However the advantage of being able to run this analysis relatively fast compared to more complex approaches, whilst still gaining an understanding of the magnitude of the impact on journey times is important, as I intend to run this analysis a large number of times.

I hope that some of the more important streets with a low traffic impact identified will be included in the list of existing pedestrianised streets in York.

If so, this analysis could be performed again with those streets removed, to identify if there are more streets in York that could be pedestrianised with little impact to car traffic.

4 Results

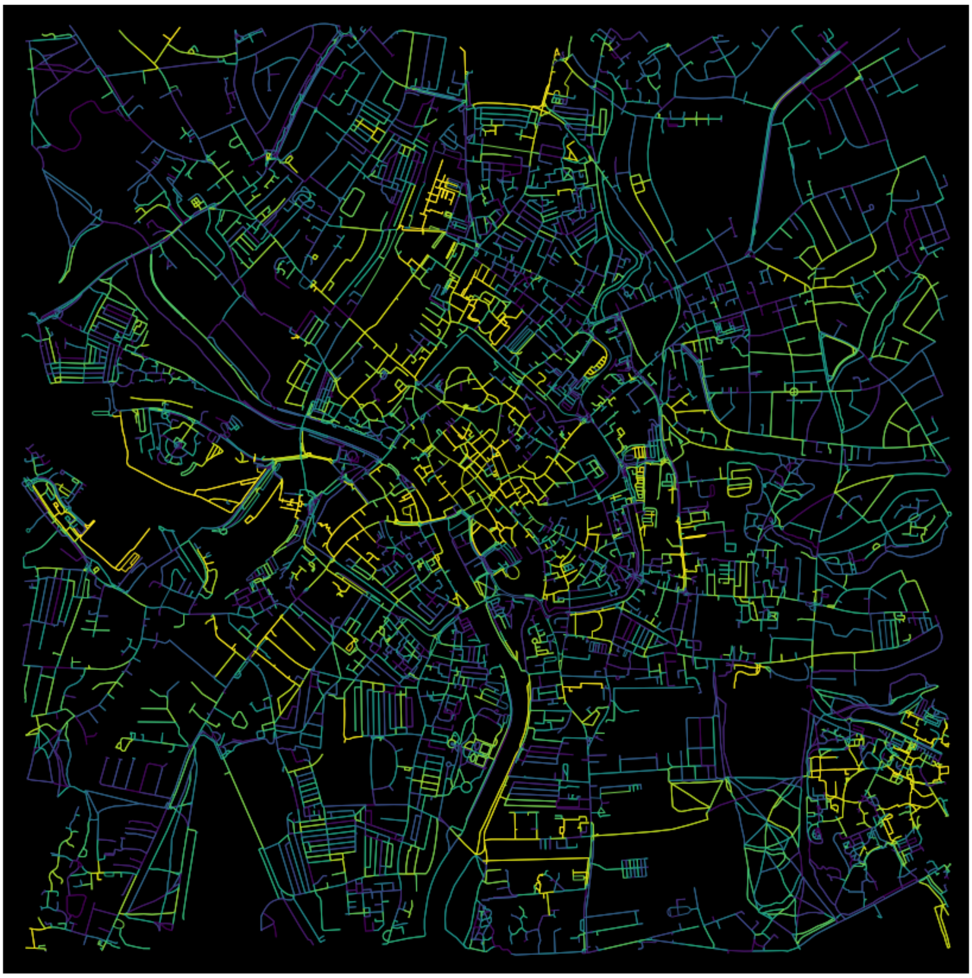
After importing and processing the data as laid out in the methodology, the two processes were applied.

Figure 4.1: Eigendata value of all highways (linear)

4.1 Eigendata Centrality

First, the eigendata centrality was applied for commercial buildings and workplaces. The results when plotted linearly as a heat map can be seen in figure 4.1, which shows a very small cluster of very high scoring edges, with the rest being uniformly low. This cluster consists of pedestrian routes within the hospital, meaning they are all very close to a very large workplace.

Figure 4.2: Eigendata value of all highways (binned)



To get a better sense of which highways are considered important the data can be plotted with each edge being placed into one of 100 equally sized bins based on their eigendata value, as seen in figure 4.2. Under this plotting, we can see that the vast majority of the city centre is considered important, along with the university (bottom right), retail park car parks (small tightly-packed collections of parallel highways), and areas of residential streets near schools (larger more spread-out collections of parallel highways next to large empty spaces).

Figure 4.3 shows the binned eigendata values when considering only highways cars could use, with all other highways shown in a dark grey, essentially the same data though a little easier to read.

Figure 4.3: Eigendata value of roads (binned)

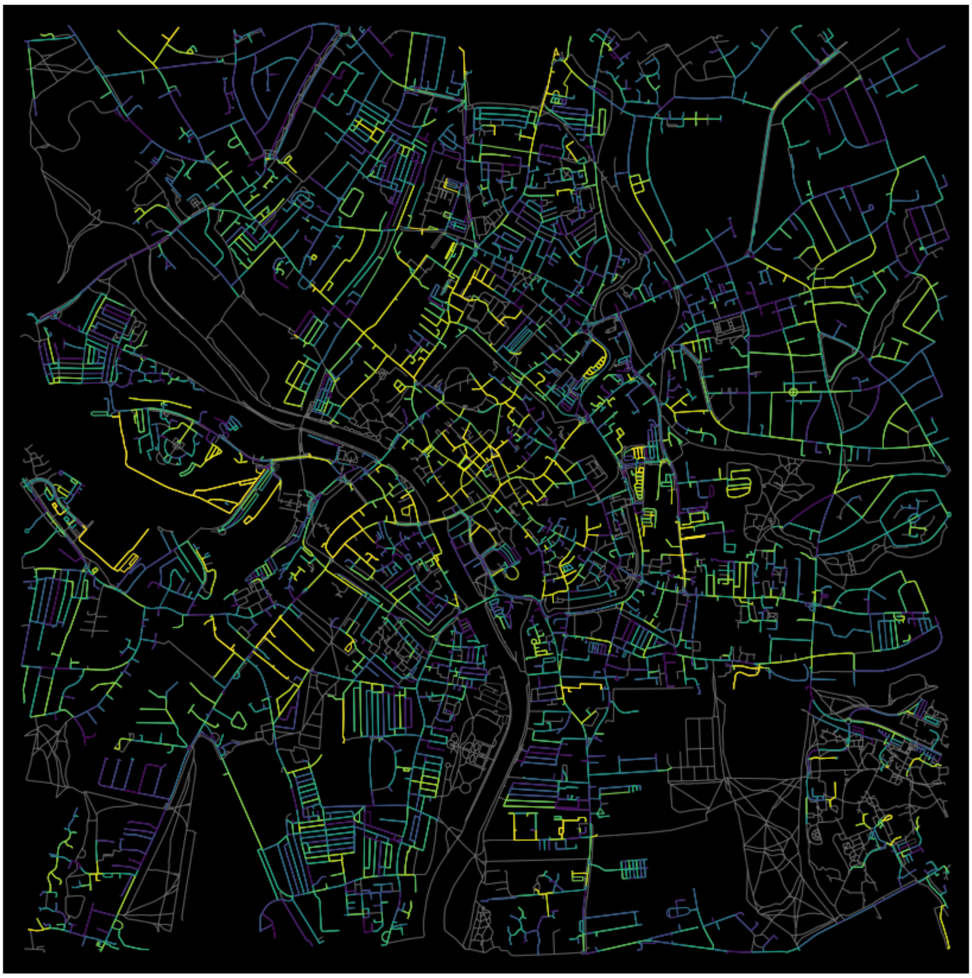


Figure 4.4 shows the graph of York's roads, with those that are in the top 2% of eligible roads in red. As expected, most of the roads selected are within the city centre, though a fair amount were selected from the outer regions (mostly near schools or business parks). A fair amount of the roads selected in the very centre of the city are roads that are pedestrianised or semi-pedestrianised [1].

4.2 Travel Time Calculation

Before the removal of the roads, 99.5% of journeys were completable. Those that could not be completed are likely due to minor errors in the connections within the dataset causing a small subsection of the graph to not be connected to the rest of the graph. Regardless, the number is

Figure 4.4: Top 2% of roads eligible for removal

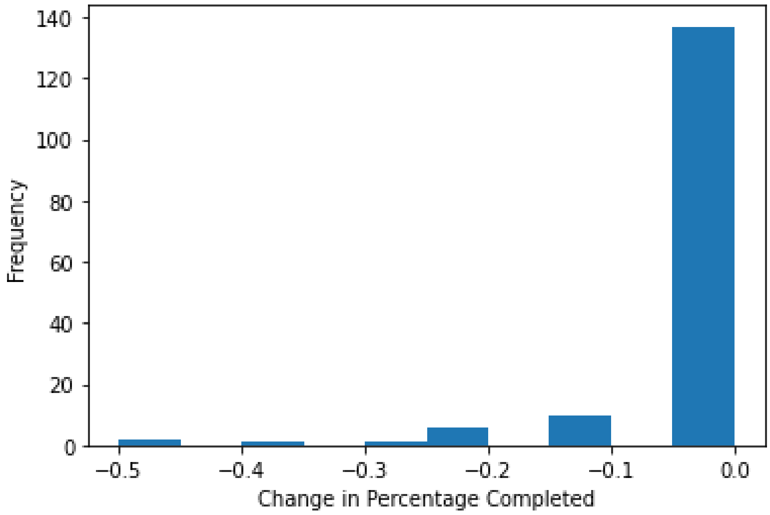
statistically insignificant.

4.2.1 Removing Each Road Individually

157 roads were included in the top 2% of eligible roads. Performing the travel time analysis on 157 networks, each with one of the roads removed, led to the production of two key graphs.

Figure 4.5 shows the change in the percentage of journeys that could be completed on each network, and shows that removing a single road has very little effect on this, with the vast majority of networks having a decrease of 0-0.05%, or 0-0.5 journeys - and since the number of journeys must be an integer, naturally these are all a decrease of 0. All of the networks had a decrease of, or less than, 0.5% - 5 journeys. This makes sense, as the only

Figure 4.5: Change in journeys completed for each road removed



way a node could not still be connected to the network when a single road is removed is in the relatively uncommon case that the node is at the end of a dead-end street. Selecting to only remove roads that have a decrease of 0 in this metric would be a sensible decision when applying this process.

Figure 4.6: Increase in average travel time for each road removed

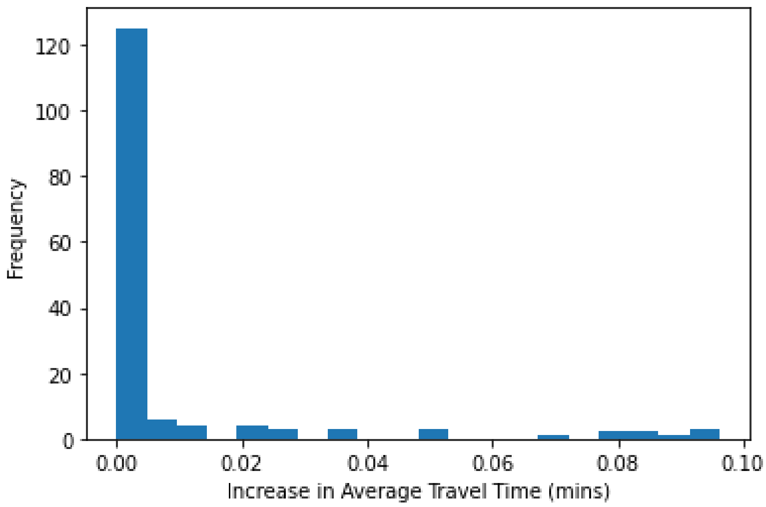


Figure 4.6 shows the increase in the average travel time of all completable journeys on each network. This shows that selecting most roads as the single road to remove suffers little to no difference in average travel time. This shows that most of our candidate roads do not lie on the shortest paths for many (if any) journeys, and that those that do lie on a shortest path have an equivalent path or an alternative with only a small increase

in journey time. It would make sense to select only roads with a small increase in this metric when applying this process, and the distribution here suggests we should start at 0 and only increase our permitted difference should we desire more roads to be removed.

Figure 4.7: Roads with an impact on neither completed journeys nor travel times

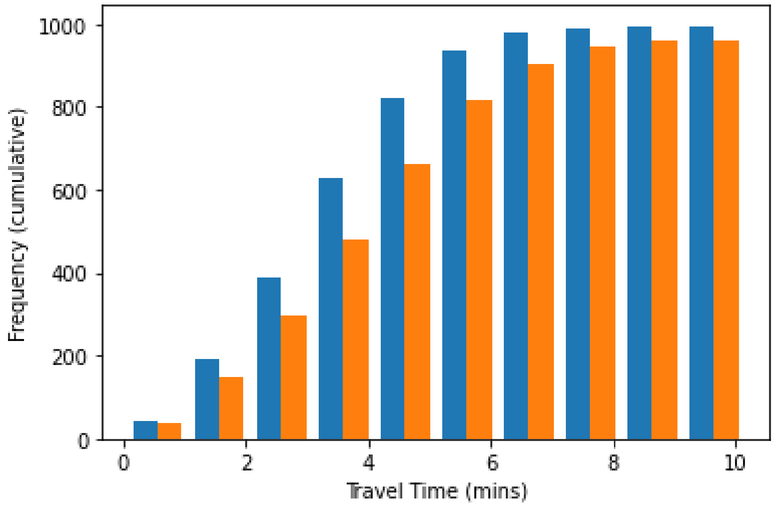
4.2.2 Removing Roads With No Impact

Applying the requirements that any road selected for removal must have had no impact on the proportion of completable journeys or their travel times when removed individually results in 58 roads being removed, shown in red in Figure 4.7. Applying the evaluation process on this network shows no change in either completable journeys or travel time. This 58 can be considered 'perfect' to remove under this system, and the roads selected in the city centre are in reality pedestrianised highways [1], though those selected in the suburbs are residential streets with many houses on them that would not be practical to close.

Figure 4.8: Roads with no impact on either completed journeys or travel times

Widening our criteria to any candidate road that either has no impact on the proportion of completable journeys or travel times, but not necessarily both, results in 137 roads being removed from the network, shown in Figure 4.8. Once again the roads selected within the town centre include pedestrianised streets [1], but now also include quieter residential / retail roads, and in one case a piece of the A1036, the A-road that encloses the city centre - a rather daring choice of street to pedestrianise. Most of the new removed roads outside of the city centre are segments of main road on the outskirts of the map, most likely selected because their position on the outskirts means they are unlikely to affect many journeys whilst also being considered more important in the eigendata centrality as their nodes are

Figure 4.9: Travel time distribution across original and reduced network



assigned buildings that may really belong to a node (and therefore edges) beyond the area considered.

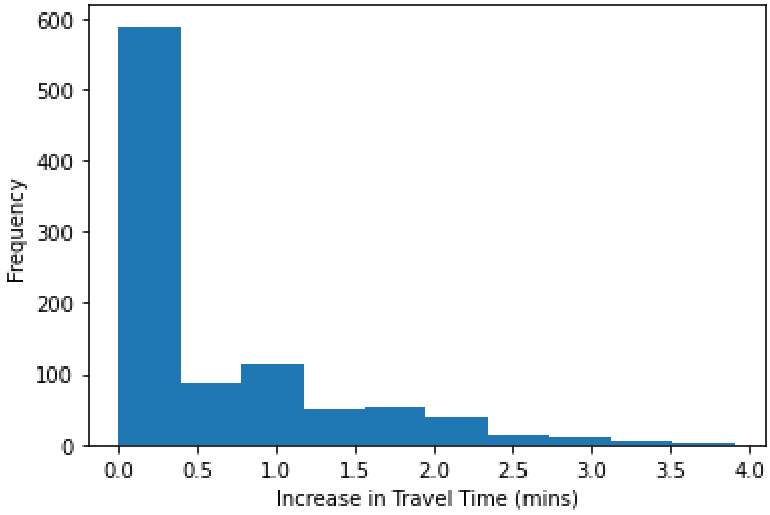
This new network sees a decrease of 3.3% in completed journeys, resulting in a 96.2% completion rate. Figure 4.9 shows the distribution of travel times across the original (blue) and reduced (orange) networks. The overall distribution is similar, with the new network having slightly longer travel times, and not achieving the same number of journeys (as expected).

The difference in travel times for journeys only completed on both networks is shown in Figure 4.10, where it can be seen that the majority of journeys experience a delay of less than half a minute. Though this is not an accurate number in terms of real-world time, it does show that removal of the 137 roads has only a small impact on most of its journeys, with only a few experiencing delays on the higher end.

4.2.3 Removing All Candidate Roads

Removing the entire 2% of (157) candidate roads shown in Figure 4.4 results in a 5.7% decrease in completed journeys - much more than even the rather generous removal of 137 roads previously. The travel time distribution follows the same pattern as previously shown (with a slightly lower total journeys), and the increase in travel time follows the same distribution over the same values as previously shown in Figure 4.9. The larger decrease in completed journeys leads me to believe that pedestrianising all 2% of the candidate roads would not be viable in a real-world scenario.

Figure 4.10: Increase in travel time between original and reduced network



4.3 Analysis of Results

The roads selected for pedestrianisation through this process are for the most part good selections. When using the strictest criteria for completed journeys and travel times, a small selection of choices are created, of which those within the city centre are good examples of existing pedestrianised streets, and those outside of it are less ideal (though in some cases arguably an interesting experiment). This pattern is amplified with the more lenient criteria, with some potential future pedestrianisation candidate streets being selected, but also some pieces of main road that would not initially come to mind (though again, closing a piece of A-road that could be diverted may make for an interesting experiment). Closing the entire 2% does not result in any more particularly useful candidates compared to the filtered choices, showing that applying the journey time calculations can assist in filtering choices. Whilst I don't believe pedestrianising all of the candidates selected by any of the processes would result in an improved city, I believe the method could be used to highlight the best candidates for a human to then perform further investigation into.

5 Conclusion

Within this project I have laid out an approach for identifying candidate roads to be pedestrianised in urban networks, making use of the eigendata centrality to do so, and a means of analysing the impact the closure of each road would have on vehicle travel times in a straightforward, non-computationally-intensive way by making use of a shortest path algorithm. The combination of these tools enables the creation of a candidate list of roads to pedestrianise based on the amount of commercial activity on the streets, which can then be filtered to only include candidates with a minimal impact on the travel times of a vehicle. This approach is easily adaptable should the specific aim be changed.

This method of selection when applied to the City of York has led to the selection of streets in the city centre that are in fact fully or partially pedestrianised, as well as selection of streets on the outskirts of the city centre which are not obvious candidates for pedestrianisation. This leads me to believe this method of selection, when considering a city centre, could be used to identify roads to be pedestrianised, after some further improvements or alongside a human hand.

If this work was to be furthered, I would recommend incorporating modern techniques for measuring vehicle travel time and network capacity. Performing a simulation of the network which can account for vehicle interactions and congestion could provide a much more robust understanding of the effect pedestrianising a street would have on the wider city. Having more time or computational power would allow for consideration of a larger pool of candidate roads, and better selections. With a more efficient travel time calculation method, a lot more computational power or significantly more time, an iterative method could be constructed in each step of which a single road is removed from the network, and the travel time analysis is performed on the resultant network in order to determine which single road would have the least impact upon removal.

Bibliography

[1] 'Footstreets - City of York Council.' (n.d.), [Online]. Available: https://www.york.gov.uk/york-city-centre/footstreets. [Accessed: 3 May 2022].

[2] T. Agryzkov, L. Tortosa, J. F. Vicent and R. Wilson, 'A centrality measure for urban networks based on the eigenvector centrality concept,' *Environment and Planning B: Urban Analytics and City Science*, vol. 46, no. 4, pp. 668–689, 2019. DOI: 10.1177/2399808317724444. eprint: https://doi.org/10.1177/2399808317724444. [Online]. Available: https://doi.org/10.1177/2399808317724444.

[3] L. C. Freeman, 'Centrality in social networks conceptual clarification,' *Social Networks*, vol. 1, no. 3, pp. 215–239, 1978, ISSN: 0378-8733. DOI: https://doi.org/10.1016/0378-8733(78)90021-7. [Online]. Available: https://www.sciencedirect.com/science/article/pii/0378873378900217.

[4] P. Bonacich, 'Factoring and weighting approaches to status scores and clique identification,' *The Journal of Mathematical Sociology*, vol. 2, no. 1, pp. 113–120, 1972. DOI: 10.1080/0022250X.1972.9989806. eprint: https://doi.org/10.1080/0022250X.1972.9989806. [Online]. Available: https://doi.org/10.1080/0022250X.1972.9989806.

[5] O. Goldsmith, 'How important is a city street?' Undergraduate project report, Department Of Computer Science, The University of York, York, UK, May 2021.

[6] J. C. Williams, 'Revised monograph on traffic flow theory, chapter 6,' [Online]. Available: https://www.fhwa.dot.gov/publications/research/operations/tft/.

[7] J. W. Godfrey, 'The mechanism of a road network,' *Traffic Engineering and Control*, vol. 11, no. 7, 1969.

[8] R. J. Smeed, 'Road capacity of city centers,' *Traffic Engineering and Control*, vol. 8, no. 7, 1966.

[9] A. Loder, L. Ambühl, M. Menendez and K. W. Axhausen, 'Understanding traffic capacity of urban networks,' *Scientific Reports*, vol. 9, no. 1, p. 16 283, Nov. 2019, ISSN: 2045-2322. DOI: 10.1038/s41598-019-51539-5. [Online]. Available: https://doi.org/10.1038/s41598-019-51539-5.

[10] 'Revised monograph on traffic flow theory,' [Online]. Available: https://www.fhwa.dot.gov/publications/research/operations/tft/.

[11] NetworkX Developers. 'adjacency_matrix - NetworkX 2.8 documentation.' (n.d.), [Online]. Available: https://networkx.org/documentation/stable/reference/generated/networkx.linalg.graphmatrix.adjacency_matrix.html. [Accessed: 3 May 2022].

[12] NumPy Developers. 'numpy.empty - Numpy v1.22 manual.' (n.d.), [Online]. Available: https://numpy.org/doc/stable/reference/generated/numpy.empty.html. [Accessed: 3 May 2022].

[13] The SciPy Community. 'scipy.sparse.linalg.eigs - SciPy v1.8.0 manual.' (n.d.), [Online]. Available: https://docs.scipy.org/doc/scipy/reference/generated/scipy.sparse.linalg.eigs.html. [Accessed: 3 May 2022].