A FRAMEWORK TO AUTOMATE THE DESIGNING OF CYCLING NETWORKS IN NODE-BASED GENERATION APPROACHES FOR SAFETY AND EQUITY, IN THE CASE OF LONDON.

Jin Jiang

Date: 04/09/2023

MSc Urban Spatial Science - CASA0010

Supervisors:

Ioannis Toumpalidis; Hannah Fry

GitHub repository: https://gitlab.com/bryden-wood-ct-beyond-spaced/beyond-spaced

Word count: 9420 words

This dissertation is submitted in part requirement for the MSc in the Centre for Advanced Spatial Analysis, Bartlett Faculty of the Build Environment, UCL.

Abstract

The important role of developing cycling infrastructures extends beyond transport, but also for a sustainable life for all the residents and a net-zero environment. Cycling in the UK and Inner London has had a revival in recent decades, yet the existing cycle tracks in areas such as Outer London boroughs are limited and low-quality. Moreover, there is a varying need for cycling network plans for different areas with different construction and investment levels, however, the main part of the research looks at filling the gaps and missing links in the existing network links.

Therefore, this study aims to develop the methodology for generating new more equitable, safe, and service-efficient cycling networks, with consideration for the potential demand of all populations and road safety. When doing target node selection, the Facility Location Problem is used to capture potential cycling demand represented by demographic data. Network Analysis methods generate and grow networks, including the Greedy Triangulation algorithm and betweenness. Then, the performance of the network results was analysed in terms of connectivity, directness, accessibility and population coverage rate to validate the methodology.

Declaration

I hereby declare that this dissertation is my original work and that all sources have been acknowledged. It is 9420 words in length.

Table of Contents

List of figures

List of tables

List of acronyms and abbreviations

Term	Abbreviation
Minimum Spanning Tree	MST
Minimum Weighted Triangulation	MWT
Greedy Triangulation	GT
Location Set Cover Problem	LSCP

Acknowledgements:

I would like to sincerely appreciate two supervisors for their guidance. Ioannis Toumpalidis gave me much help on the whole method and technical problems. Prof. Hannah Fry provided help with the overall direction and structure of the dissertation.

I would also like to acknowledge the help and accompany of Wangbing Tan as my study buddy.

1. Introduction

1.1 Background

The increasing demand for cycling, both before and during the pandemic, has become a crucial aspect of urban transportation systems worldwide. The drive for decarbonization and the need to address traffic congestion problems in cities have elevated the importance of cycling as an environmentally friendly and efficient mode of transportation. The United Kingdom, recognizing this trend, introduced the Cycling and Walking Investment Strategy (CWIS) in 2017, with a subsequent update in 2021, outlining ambitious plans to promote cycling and walking as natural choices for short journeys by 2040 (Cycling and Walking Investment Strategy: Report to Parliament, 2020). These plans involve substantial investments aimed at improving cycling infrastructure and accessibility.

During the pandemic, temporary cycling lanes were established in many cities to facilitate social distancing and reduce the reliance on public transport. While these initiatives improved the convenience of commuting for some, they also sparked debates regarding their long-term impact on traffic congestion, given the reduced road width and motor vehicle capacity (Quinn, 2020). As we transition into the post-COVID era, it is imperative to assess whether these temporary changes should be retained to benefit residents. Comprehensive measures are needed to determine the value of preserving certain cycling lanes.

An essential component of encouraging cycling is the provision of robust cycling infrastructure. Some researches are underscoring the pivotal role of infrastructure in promoting cycling behaviour (Nelson and Allen, 1997). Additionally, cyclists' perception of safety significantly influences their choice to cycle within their neighborhoods (Hull and O'Holleran, 2014). Segregated cycling lanes, differences in travel speeds between cyclists and vehicles, and other safety measures contribute to enhancing this perception. Therefore, the development of safe, reliable, and sustainable cycling infrastructures can potentially foster greater cycling adoption.

Despite progress in cycling infrastructure within Inner London during and after the pandemic, research indicates an uneven distribution of infrastructure between Inner and Outer London (Tait et al., 2022). Cycling infrastructure, including lanes and routes, remains scarce in Outer London, hampering efforts to encourage cycling as a mode of transportation. This issue is particularly evident in our selected case study area, a London borough evaluated as underperforming in terms of cycling infrastructure(Tait et al., 2022).

In light of these challenges and disparities, there is a pressing need for a broader vision for planning cycling infrastructure across the entirety of London, recognizing that while central areas are pivotal to travel plans, a large number of journeys commence outside the city centre. A holistic approach to designing cycling infrastructure network is essential for establishing cycling as a sustainable and accessible mode of transportation for all of London's residents.

put a map showing existing cycling infrastructures in London?

1.2 Research question

The main question would be: Where should cycling infrastructures be constructed so that the network is safe and equally connected around the area? The objective of this research is to develop a framework for designing cycling network with a wide vision for safety and equity.

This study will also look at:

1) The optimisation of selecting nodes from road intersections. Which method could be applied in selecting target nodes when concerned with serving more population?

2) The application of the Greedy Triangulation algorithm for growing new networks from nodes, focusing on safety and equity. The growing process of networks.

3) Validating the proposed cycling network. What performance indicator do we use to measure the network results? To what extent do they improve the cycling network's performance?

2. Literature reviews

Research on cycle lanes and networks has emerged rapidly over the last twenty years or so, particularly since 2010. Most research in cycling level and bicycle networks are quantification research, focusing on physical characteristics and trip behavior statistics of the cycling network. A review paper on effects of cycling divided relevant research into three evolvements: targeting the links which are paths and lanes, to the nodes which are intersections and finally to the whole network of cycling network (Buehler and Dill, 2016).

When looking at the factors impacting cycling choice, most research did a correlation analysis between bikeway infrastructure and cycling level. Using big data such as mobile data and transport flow data, a more data-driven approach emerged in recent years. Combined with network analysis which rose speedy development in a decade ago, many network-level analysis applying graph theory and algorithms emerged. Compared to the evaluation and correlation analysis before head, there is more research in designing cycling network.

2.1 Factors that impact people's cycling choices.

The decision to choose cycling as a mode of transport is influenced by a variety of factors. Nelson and Allen argue that 'if you build them, commuters will use them', emphasising the importance of infrastructure development in facilitating cycling journeys (Nelson and Allen, 1997). Buehler and Dill noted that safety is the primary factor influencing people's willingness to cycle (Buehler and Dill, 2016). CROW by the Netherlands is a commonly used credible reference for design guidance in this area. More recently, the UK Department for Transport's published Cycle Infrastructure Design in 2020, sets out five basic designing principles, including cohesion, orientation, safety, attractiveness and comfort, and associated measurement factors (GREAT BRITAIN: DEPARTMENT FOR TRANSPORT, 2020). In addition, research by Schoner and Levinson emphasised that the connectivity and directness of cycle networks play a key role in influencing cycle commuting choices (Schoner and Levinson, 2014). Together, these factors determine whether an individual uses cycling as a viable mode of transport.

2.2 Analysing street network

2.2.1 Complex Network and its application in street network and transportation.

Complex network and graph theory have been widely applied in the real world, study fields which include social science, biological, communication technology and so on. When looking at street networks and transportation systems, *spatial networks*, a specific kind of complex network, are introduced to study the topological properties of networks and the spatial characteristics of physical space and transport connections (Cardillo *et al.*, 2006). In network-level analysis, street networks are planar graphs with the nodes as the intersections and the edges as the roads connecting the intersections, where attributes of roads and intersections could be assigned into edges and nodes to be analysed for specific objectives. The fundamental concepts of graph theory are integral to the analysis and optimisation of cycling networks. These concepts include the shortest path problem, which helps to identify the most frequently passed-over routes, and centrality measurement, which helps to identify nodes and edges at critical junctions and road segments (Derrible and Kennedy, 2011). In addition, community detection techniques in graph theory help to identify potential clusters or communities within a network, thereby facilitating focused planning efforts(Clauset, Newman and Moore, 2004). Attributes of the whole network such as connectivity and directness are often used to measure the performance of a cycling network.

2.3 Designing the Cycling Network

2.3.1 Building weighted cycling networks and estimating cycling demand

Street network consists of walkable, drivable, and bikeable urban networks (Boeing, 2017). Bicycle networks could be a part of roadways as cycle lanes, separated cycle tracks away from roadways and bikeable paths in parks (Buehler and Dill, 2016). The weighted undirected network is often used in studies of cycling transport networks. As mentioned above, safety and cyclists' perception of the road are decisive factors in commuting mode choice. Klobucar and Fricker constructed a bicycle compatibility index (BCI) measure that included safety and distance in a 'safe length' for each link in the network (Klobucar and Fricker, 2007). Mahfouz embedded in road type to construct a weighted road network, concerning stress level (Mahfouz, Arcaute and Lovelace, 2021). Folco defined a weighted distance between nodes with the number of crashes and trips in that edge for demand and safety(Folco *et al.*, 2022). Akbarzadeh created a weighted graph based on taxi trip patterns to capture the commuting demand (Akbarzadeh, Mohri and Yazdian, 2018).

Cycling demand was often estimated from OD data which is transformed to edge flow in the network. For example, Akbarzadeh, Mohri and Yazdian used taxi trip data to construct a trip network, extract demand points and estimate cycling demand (2018), and Folco et al. used e-scooter positions data to estimate bicycle flow on roads (2022). Alternatively, other common approaches include using transport stations as demand nodes (Shui and Chan, 2019), extracting population information and land use to aggregation points which are used as demand nodes subsequently (Zili, 2020) and so on. These fill a need for research where OD data is not available. Beyond this, however, there is relatively limited research on target nodes.

2.3.2 Methods and algorithms of designing cycling network and their limitations

The subject of cycling network design has received rapidly increasing attention in the last decade or so, and the main problems identified in descriptive and evaluative studies include the fragmentation of cycling infrastructures in cities, which means the networks consist of many sub-networks disconnecting from each other(Schoner and Levinson, 2014). Evidence was found in cities around the world, no matter in leading bicycle cities such as Copenhagen (Natera Orozco *et al.*, 2020), in developed car-focused cities such as Montreal (Boisjoly, Lachapelle and El-Geneidy, 2020), or in many US cities such as Seatle (Lowry and Loh, 2017).

In the last few years, there has been a surge of research into specific cities to design growth strategies for cycling infrastructures. Utilitarian approaches pursue in to maximise the global benefit. For example, Largest-to-Closest Algorithms add links between existing large components and links in the whole graph (Natera Orozco *et al.*, 2020). This method can leverage the existing infrastructure and reduce investment costs for greater benefits (Natera Orozco *et al.*, 2020). The utilitarian approach, however, was also reflected as enhancing the spatial inequity of existing cycling infrastructure resources (Pereira, Schwanen and Banister, 2017).

Some specific concepts in graph theory were applied in cycling network analysis. It is assumed that there are communities in complex networks and some of the communities are linked more tightly to each other than nodes in other communities (Blondel *et al.*, 2008). Community detection is used to identify potential communities in a network and trip patterns if used in a trip flow network. For example, Olmos *et al.* (2020) used percolation theory to filter out flow links in mobility networks. Mahfouz, Arcaute and Lovelace(2021) used community detection to help propose an Egalitarian Expansion Algorithm, spatially dividing the whole area into coherent clusters and addressing the ethical problem by ensuring the resource was allocated for every community.

Moreover, in the context of transport network planning, Shui and Chan (2019) compared Minimal Spanning Trees (MST) and Shortest Paths problems, the different design objectives of the existing application of the algorithm. The methods on shortest path and MST problems are commonly used in transport planning models (Magnanti and Wong, 1984), yet it is less used in the design of cycling networks. The insufficient literature on planning networks from nodes does include a few empirical research, for instance, Shui and Chan (2019) used the MST algorithm to generate a new network based on the target nodes, which were selected by iteration of the Genetic Algorithm; Akbarzadeh, Mohri and Yazdian (2018) used a more complex bi-objective optimization model to construct bicycle networks which connected key points in each community. Zili (2020) applied MCLP, a facility location problem, to translate the population area map to the population aggregation points map and used MST to design the new cycling network to reach the target node.

Yet transport networks generated by MST are argued as not user-friendly and time-consuming for travellers (Szell *et al.*, 2022). The nature reference graph should be the Minimum Weighted Triangulation (MWT) theoretically, which consists of triangulars and also has a minimal total weighted length (Cardillo *et al.*, 2006). As a result, the MWT network has good connectivity and a short overall travel time for travellers but also has a relatively low total length, i.e. low construction costs. Yet it is hard to compute by known polynomial time algorithm, so Greedy Triangulation (GT) is considered as an approximation of MWT but easily computable (Cardillo *et al.*, 2006). Szell *et al.* (2022) proposed a structured GT methodology in a growing topological cycling network.

Following this method, Folco *et al.*(2022) explored the data-driven approach to design the cycling network in Turin and added more practical factors like crash data, trip data and existing cycling lanes.

2.4 Research Gap

Up to the present day, significant advancements have been made in the development and enrichment of algorithms for generating networks. A multitude of research efforts have been dedicated to designing methodologies for bridging the gaps and establishing missing links within existing networks. In the idea of creating a network connecting nodes, however, the consideration of node selection has been neglected. Firstly, Due to the nature of the algorithm itself, the choice of nodes largely determines the shape of the network. Secondly, the design approach from nodes is newer and less systematic. Researchers have applied several methods in selecting target nodes in different aspects, such as extracting nodes based on the trip flow, randomly chosen on a grid when no OD trip data is available or chosen by rail stations and POIs, or more complexly, using Genetic Algorithm to optimise. But it is not sure, yet which methods might be more suitable for what situations. Also, these methods are applied in different case cities so it's hard to compare.

In summary, the corresponding research gap is few papers focus on exploring different node selection approaches, which can thoroughly affect the resulting networks. Therefore, this study is going to focus on the node selection methods in node-based cycling network generation algorithms**.** On the other hand, although multi-factor evaluation guidelines are proposed in papers and design criteria documents, they are difficult to consider substantively in the design of the network. The design of the network is often divorced from the real situation of the road and fails to take into account, such as safety factors such as road width, speed limits, segregation, slope; attractiveness factors such as green spaces and public spaces. For example, the GT algorithm proposed by Szell *et al.* (2022) is a topological method. We could add more practice and real-life factors to this promising approach.

2.5 Theoretical Framework

After reviewing the relevant literature and understanding graph theory, it is possible to summarise two rather clear ideas of network generation. One approach is to connect existing sub-graphs or paths and grow incrementally. Another approach is to grow network by connecting nodes**.** Szell *et al.* (2022) structure three steps in the methodology: seed points, greedy triangulation and order by growth strategy, which is a node-based network generation. The design methodological frameworks in many studies can be fitted to this broad idea. For instance, Akbarzadeh, Mohri and Yazdian (2018) defined node pairs in each community as key points to pass bike lanes between points. Shui and Chan (2019) used GA to optimize demand points and routing networks.

In contrast to the previous approach, which primarily concentrated on augmenting existing links, this method proves to be better suited for regions characterized by a deficiency in cycling facilities. It offers the advantage of devising a comprehensive network that uniformly covers the entire area, making it an ideal foundation for a long-term plan from the ground up. This objective-oriented approach empowers planners to selectively identify target nodes based on specific demand criteria. For instance, in addition to addressing commuting demands, planners can easily incorporate priorities related to green spaces, open areas, or other spatial factors into our design procedures.

The objective of this study is to develop a framework for designing a cycling network with a wide vision for safety and equity. The node-based generation approach allows us to represent design goals more directly on nodes. Therefore this study is going to execute the design process along with Szell’s research. At the same time, sub-steps are inspired by other research mentioned above, aiming at optimising the process of seeding points and capturing the cycling demand using different methods, including facility location problem, community detection and spatial clustering. Target nodes refers to the selected nodes for connecting to get a whole network in the following steps.

3. Methodology

The main purpose of this study is to propose a methodology to design a cycling network with a wide vision for safety and equity according to demands flexibly. As mentioned above, in the idea of Greedy Triangulation network generation approach, referenced and inspired by Cycling Level of Service Tools, this study embedded indicators of road situation in edges to present safety issues and considered demographic data to present cycling demand.

In the first step for seeding point, this study would apply several methods to get target nodes. They included Community Detection, Spatial Clustering and Location Set Cover Problem (LSCP). Then in the second step, GT network algorithm is presented to generate a whole triangular-connected network connecting to the target nodes, which follows the method of Szell *et al.* (2022). In the ordering stage growth strategy, prioritise the link with higher weighted betweenness in safety.

Lastly, compare the results generated for the same existing street network within the investment scenario. The validation of the network proposal will be given to evaluate the performance of each approach result. The metrics contain both global measurements and specific objective-based measurements which focus on how conveniently cycling would be in the new networks. Then based on the performance of the network in different indicators, we make a critique and reflection on the three methods of node selection.

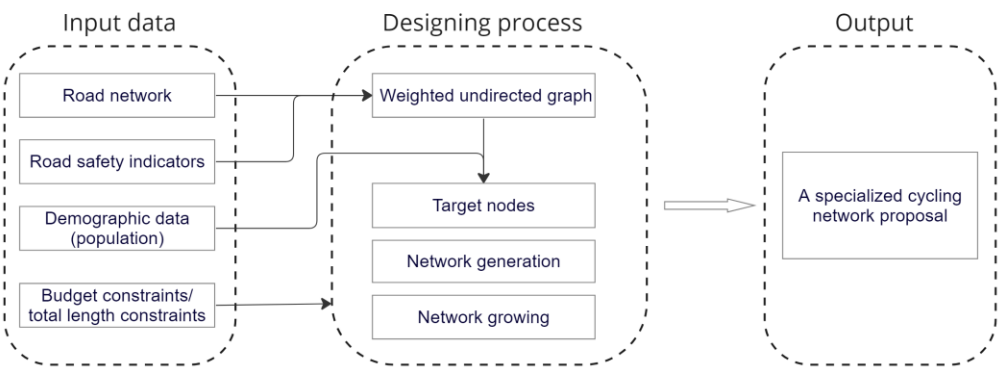


Figure 1. Methodology framework

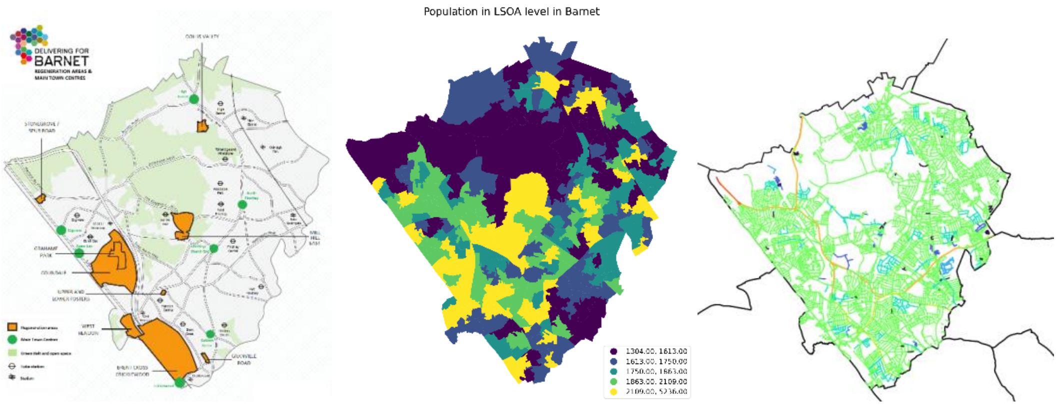
3.1 Data Source and Study Area

Street network data including walkable, drivable, or bikeable urban networks is obtained from OpenStreetMap using OSMOX python library (Boeing, 2017). This data includes road attributes like road type, speed limit and so on. Existing cycling routes and lanes in London are downloaded from TFL. These could be used in the validation part. Demographic data, the population of each LSOA in 2020 and the LSOA boundaries in 2020 are obtained from ONS.

London Borough of Barnet is chosen as the case study area to execute the methodology. According to recent research on evaluating cycling infrastructure in London, the London Borough of Barnet performs worst in London in terms of "density of cycle lanes with some form of separation", with only 0.05 km per square km (Tait *et al.*, 2022). This borough is an example to set out to build a full model and it is feasible to expand the analysis scope to other or larger areas.

3.2 Data Exploration and Preparation

3.2.1 Structure of street network in London Borough of Barnet



(a) (b) (c)

Figure 2. (a) Regeneration and town centres in Barnet, green dots are town centres; (b) population distribution in LSOA level; (c) road max speed map, green (lower speed) to orange-red(higher speed)

In Barnet, the population is mainly concentrated in the southern and eastern areas. The maximum speed limit on most of the roads in the area is 30 mph, while the speed limit on the orange colour trunk roads is higher in Figure 2(c), ranging from 40 mph to 60 mph. The street network is treated as an undirect graph in the analysis. In Table 1, N is the number of nodes, K is the number of edges, $\langle k \rangle$ is an average degree, $Cost$ and $\langle l \rangle$ are the total length of edges and average edge length (in metres). $P(k)$, probability of having nodes with degree respectively equal to 1, 2, 3, 4, 5 and 6 are shown in Figure 3 (Cardillo *et al.*, 2006). Around twenty small and disconnected components were not included in the candidate graph, and all of which were internal roads in car parks. Most node degrees are 2 and 3, which revealed that the road network is relatively evenly distributed in Barnet, the average degree is 2.28. There is a comparatively empty area in the northern area of the borough centre, which is the green belt and open space. And seven town centres are spreading around this green space in the vicinity of main roads.

Table 1. Summary data of street network in Barnet

N	K	$\langle k \rangle$	$Cost$	$\langle l \rangle$
11541	13171	2.28	800697.71	60.79

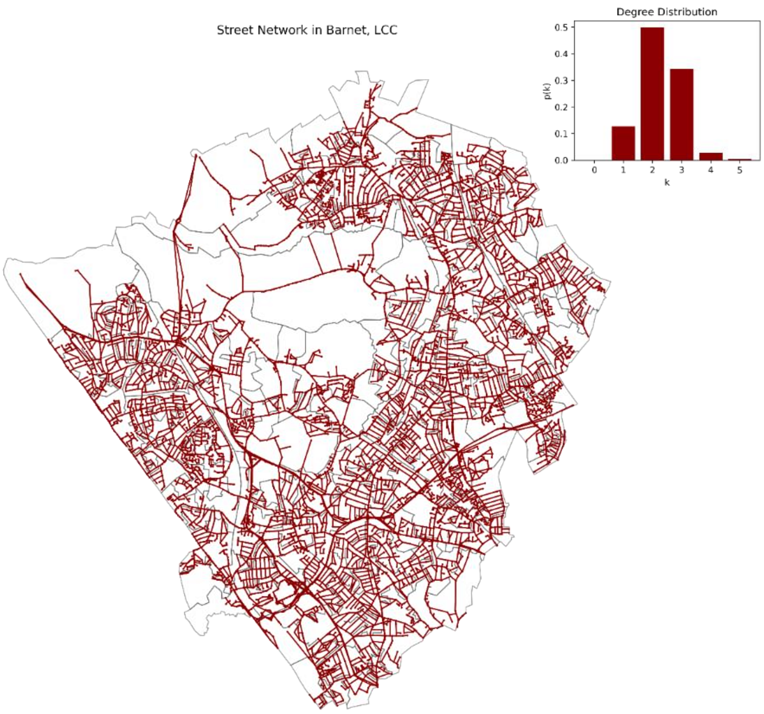


Figure 3. Street network (exclude park paths and residential internal roads), London Borough of Barnet

3.2.2 Selecting Candidate Edges and Nodes

Firstly, to filter out the proper roads for using and construction and simplify the edges, there are some criteria: the road is safe for cyclists, no self-loop road, no single disconnected nodes. There are five design principles in the Cycling Level of Service Tool (GREAT BRITAIN: DEPARTMENT FOR TRANSPORT, 2020). Due to the data availability, we choose one factor in Safety principles, the motor traffic speed on sections of shared carriageway. The 'maxspeed' attribute of the road was assigned in the graph as the weight of edges to represent the motor traffic speed limit. Roads without the 'maxspeed' attribute were found to consist of two main parts, endpoint paths to individual residency buildings and paths in parks. These roads were removed as the former is not suitable for building as cycle lanes, the latter can be considered for being re-added to the candidate edges. The roads whose 'maxspeed' is over 37mph are considered critical (GREAT BRITAIN: DEPARTMENT FOR TRANSPORT, 2020) so they were removed. Some of the roads contain two 'maxspeed' then we took the lower value one. Notes that there are several trunk roads whose 'maxspeed' is over 37mph, but separate cycle lanes or tracks exist alongside the motorway on the roads according to TFL data. They are very much the main roads for Barnet, thus, these edges were not removed from candidate edges, including Watford Way, Watford By-pass, Great North Way and Barnet Way.

All self-loop edges and single nodes were removed. We scored roads in a four-level classification based on the Bicycle Level of Service Tool (Table 2). Finally, the largest connected component was taken as a candidate graph grating for computation, after ensuring almost all the nodes and edges are included.

Table 2. Safety score classification on roads

Key requirement: Safety				
Classification	Critical	0	1	2
Motor traffic speed on sections of shared carriageway	85th percentile >37mph(60kph)	85th percentile >30mph	85th percentile 20mph-30pmh	85th percentile <20mph
Score for the road in this study	1	2	3	4

3.3 Selecting cycling target nodes

Regarding the comparatively large number of intersections in cities when selecting nodes, it is not feasible to iterate over all nodes. There are three approaches developed in this study to get fairly distributed nodes that satisfy population demand: Community Detection, Spatial Clustering and Location Set Covering Problem. The outcome in this stage will be a subset of nodes from road intersections which scatter over study area.

3.3.1 Approach 1 – Location Set Covering Problem, Facility Location Problem

This study used the fundamental Location Set Covering Problem model to do population-based facility location optimise problem, with population as the demand. The scenario is set as: cycling lanes intersections should supply more service for local people considering the distance between the two sets of sites: demand points which are the population weighted LSOA centroids and candidate supply sites which are the cycling lanes intersections. The 'spopt' python library is used.

As the model cannot set polygon data as input, there is a transformation from area to points. The steps: 1) Generate a representative set of points for LSOA polygon. Calculate the geometry centroids of LSOA area as demand points with population number as the demand. 2) Define the possible facility points location. This was taken from all the road intersections. 3) Set the parameters and run the model. The service radius depends on how far people are willing to walk with their bikes to the cycle tracks, which was set as 800 metres in this study accordingly. Finally, the outcome was an optimal least number of nodes that cover all population aggregation points.

3.3.2 Approach 2 - Community Detection

Community detection method could help identify critical cycling infrastructure places to connect different communities. In the context of planar street network, community detection involves identifying neighbourhoods and groups of intersections which are closer to each other spatially than in other groups, which helps ensuring the equitable distribution of cycling facility. The Clauset-Newman-Moore method was used to separate intersections into communities. Population data in

LSOA level was assigned to nodes in street network by location, then we obtained a geospatial division considering population distribution.

After applying community detection, randomly choose one node in each community and ensure that the minimum distance *d* between each pair of nodes is kept so as to keep an ideal nodes spatial distribution. Then set d as 300 metres in this study. Additionally, there is a research finding that the distance variations within the set interval from 250 to 1000 metres are not enough to effectively impact networks' performance (Folco *et al.*, 2022). Then the target nodes will be delivered in the following step.

3.3.3 Approach 3 - K-means spatial clustering analysis

K-means clustering analysis is used to divide nodes into clusters spatially. 'Sklearn' python library was used to apply K-means clustering analysis to intersections' coordinates. Also, the cluster result and choice of K should be checked by elbow method and be plotted to see if it makes sense. Given the clustering result, randomly choose a node in each cluster and ensure that the minimum distance d between each pair of nodes is kept. The distance (metres) *d* and reason for it is the same as stated in last section. Then the target nodes will be delivered in the following step.

3.4 Create and grow network

Greedy Triangulation algorithm is computed to create networks connecting to all the points concerning route distance. The growing principles are adding straight-line links in distances order between node pairs and ensure that no intersections of edges exist, then routing the links to the weighted shortest paths for each pair of nodes in street networks (Szell *et al.*, 2022). For investment strategy, the shortest path was a weighted choice, which considered road safety and spatial travel distance. The weighted distance of edge is inversely correlated with the betweenness; therefore, the weighted distance is defined as in Equation 1. All the available road links were given a safety score according to the criteria in table 2 and it was normalized and reversed to range [0.1, 1], then to get the weighted distance for each edge.

$$d_w(\mathrm{i},\mathrm{j}) = s * d(\mathrm{i},\mathrm{j}) \qquad (1)$$

Where: $d(\mathrm{i},\mathrm{j})$ is the path length between node i and j, and $d_w(\mathrm{i},\mathrm{j})$ is the weighted distance between node i and j. Safety score $s_o \in \{1,2,3,4\}$.

$$s = 1 + 0.1 - ((\mathrm{s_o} - \min(1,2,3,4) / (\max(1,2,3,4) - \min(1,2,3,4))) * (1 - 0.1) + 0.1$$

In the scenario of an investment budget, all the edges in the whole greedy triangulation network are re-added in the descending order of weighted betweenness. The betweenness calculation considered the weighted distance obtained in the last result. The higher the betweenness of the edge, the more priority it will have when growing.

3.5 Validating the performance of networks results

Since the number of nodes elected and the size of the generated network differ from each of the three methods, and the total size of the complete network is too large, it is impractical to construct the complete triangulation networks. Therefore, we used the investment length of the network as a benchmark for cross-sectional comparison. As mentioned above, the growth strategy was based on road safety-weighted betweenness and we would add a cut when the total length reached each set investment limit as the network grows.

To assess the quality of network proposals, we used some metrics to measure the growing network. The indicators contain common global properties and specially constructed indicators: coverage, connectivity (number of components), global efficiency, directness, population coverage rate and accessibility. Coverage is defined as the 400-metre buffered area of network edges, following the parameter distance that London and TFL often use (Norman, 2023). Global efficiency, E, is the average efficiency of pairs of nodes given in Equation 2 (Latora and Marchiori, 2001). $\mathrm{d}_{\mathrm{w(i,j)}}$ is the weighted shortest path length between node i and j, and it is normalised by the geographical distance between node i and j, d_{ij}^{Eucl}, allowing us comparing different sizes of graphs. Global efficiency presents how efficiently a cyclist can travel between any two positions, taking road safety and travel distance into consideration.

$$E = \frac{\sum_{i \neq j \in G} \frac{1}{\mathrm{d}_{\mathrm{w(i,j)}}}}{\sum_{i \neq j \in G} \frac{1}{d_{ij}^{Eucl}}} \quad (2)$$

Directness is measured by metric D as in Equation 3, d_{ij}^{Eucl} is the distance between node i and j along a straight line, d_{ij} is the shortest path length between node i and j, K for the number of nodes (Cardillo *et al.*, 2006).

$$D = \frac{1}{K} \sum_{i \neq j \in G} \frac{d_{ij}^{Eucl}}{\mathrm{d}_{ij}} \quad (3)$$

This study constructed a population coverage rate to estimate the population covered by the network coverage spatially. By calculating the proportion of covered area in each LSOA area and multiplying it by the population in each LSOA, to get the estimate after summing up.

Accessibility, refers to the average distance from each population aggregation points to the nearest cycling lanes. It measures how convenient riding bikes could be in the new networks for residents. Then we see how the metrics change along with the network's total length growth. From the plots of the three approaches, we make an evaluation and reflection on the three node selection approaches and the greedy triangulation network method.

4. Results and discussion

In this section, we delve into the process of node selection, which arises from three distinct approaches. These approaches provided how local population demand and division in Barnet are effectively addressed. We then proceed to examine the cycling network proposals that have emerged from these approaches, evaluating the performance of these networks in terms of safety, equity, and overall functionality. At the same time, we also examined the consistency of the results with our initial research topic and their contribution to the wider field of cycling network design. This analysis provides us with detailed insights into the methodology, the results and their relevance to enhancing cycling infrastructure and urban transport planning.

4.1 Target nodes Selections

In the LSCP model, 35 nodes were able to cover all the population aggregation points when the service radius of each intersection node was set as 800 metres. The nodes were distributed mainly in the Southern area of Barnet, as in Figure 4. The set of nodes covered a part of the population for each LSOA area, and this result would be used to create a network connecting all the nodes.

82 communities were found in the community detection on the street network in Barnet, with population weighting nodes. The spatial structure of the non-homogeneous street network is revealed by the division into dozens of small aggregations, with distinct spatial boundaries and no well-marked overlap was observed. It reflects the different aggregation communities and social and geographical divisions. Node selection result is in Figure 4(d). In the third approach K-means Clustering analysis, the number of clusters is pre-defined. To align with the result of Community Detection, K was set as 82 and the cluster result was as in Figure 4. Modularity values of two community groups were: 0.968 (Fast-Greedy Community Detection) and 0.931 (Clustering analysis). Overall, the results of all three node selection methods reflected the structure of the road network in Barnet, which is sparse in the north central area, whilst the rest of the area is spread across the main trunk ways and the seven town centres.

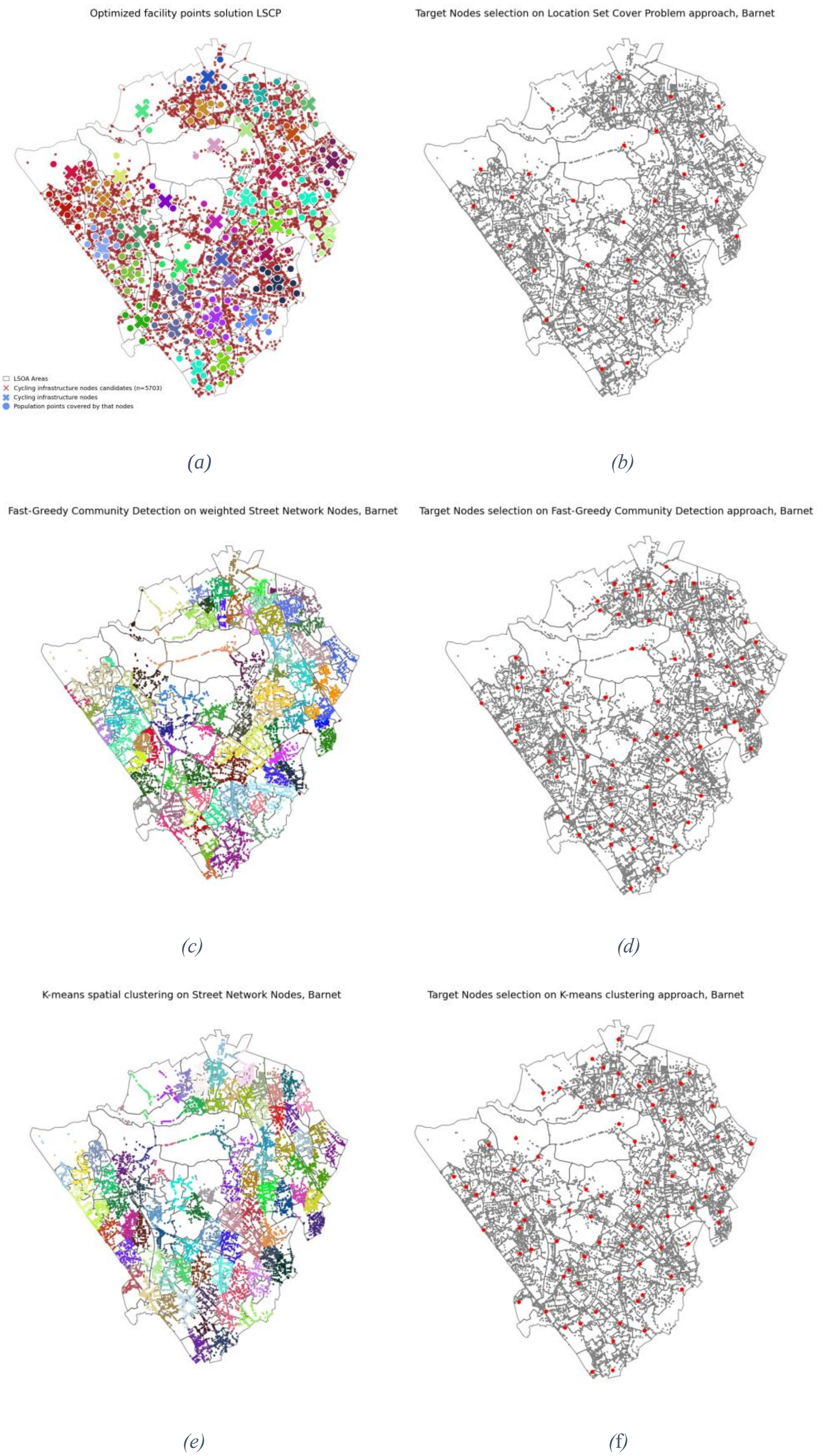


Figure 4. Node selection results for three approaches: LSCP, Community Detection and spatial clustering

Even though the modularity of community detection result was high, however, we need to reflect on the classification result on road intersections, that the road network was too finely segmented. Thus, we were getting too many target nodes if aiming at the highest modularity. This study tried to pre-define the maximal number of communities equal to 20 to force the communities to stop separating, and the corresponding modularity was still high as over 0.9. This might be because street network is a planar complex network, and there is no hub node like in social network or flow network. Most junctions have 2-3 connections, with only a few reaching 5-6 nodes, which work as local transport hubs but are not going to have great further effect. In addition, the weight factor embedded in the road edges was demographic data based on location in this study, which did not present the mobility flow on roads, therefore the interaction relationship among intersections cannot be well presented in community detection result. An improvement could be assigning motor volume or trip data of roads into network to do community detection, then the regional traffic character and junctions detection will be captured (Mackaness and Mackechnie, 1999).

4.2 Proposed networks

Three complete GT networks are as in Figure 5, where all the selected nodes were planned to be connected by cycling circles. Abstract networks provide us with a clear idea of the road network (Figure 5a-5c), and then after assigning routes based on actual roads, we get routing networks (Figure 5d-5f). In this part we are going to explain the features of network planning results in brief and compare the three results horizontally. There are 92 edges connecting 35 nodes in the network result of LSCP approach, 230 edges connecting 82 nodes for community detection approach and 233 edges connecting 82 nodes in the clustering analysis approach. The average degrees of all three networks are over 5, showing that the GT network holds a high connection efficiency.

Table 3. Summary information on abstract network results

Network information on abstract network	N	K	Average degree
LSCP approach	35	92	5.26
Community detection approach	82	230	5.61
Clustering analysis approach	82	233	5.68

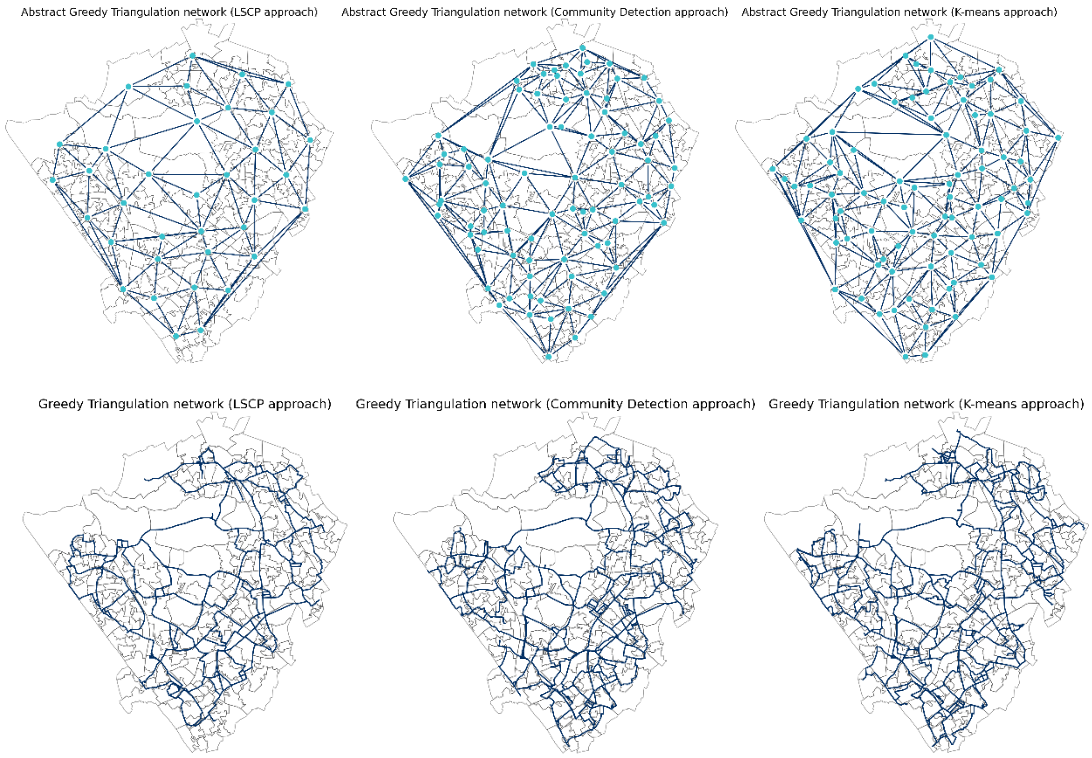


Figure 5. Network generation results for three approaches, LSCP, Community Detection and spatial clustering

Reflecting on our research question concerning the strategic placement of cycling infrastructures to establish a safe and evenly connected network in the Barnet area, our network proposals have yielded promising results. The planned network links are spread across Barnet, exhibiting a distribution that, while not entirely uniform, is proportionate to the local population and the characteristics of the road network, as depicted in Figure 2.

Comparing the abstract and routing networks as illustrated in Figure 5, reveals some disparities in the paths before and after routing. These disparities can be attributed to the structural characteristic of the street network in Barnet. The non-homogeneity of the road network is evident, especially in the north-central part of the natural park area where the road network is sparse, which results in rarely shortest path through the node pairs of the area in a nearly straight line, leading to a lower level of directness. Also, there are some disconnected edges in our candidate graph due to restrictions imposed by road speed limits during the selection of candidate edges. Thus, the shortest paths between some node pairs are tortuous, contributing to differences in the structural makeup of routing networks compared to their abstract counterparts.

Another concern when looking at the different network generation results is that the three methods differ greatly in nodes selection and abstract networks results, but the network presented after routing have a high structural similarity, particularly for the community detection and spatial clustering methods**.** First there is a speculation to explain it, that there are some roads as common shortest paths between node pairs and many other shortest pathes go through them, therefore they would be connected first and occur on the different routing results of triangulation network based on different target nodes as hub edges. Based on this, we further speculate whether this suggests that the number of nodes selection in the GT algorithm has a limited impact on the results of the generated network? The general structure is settled at early stage, then the following edges simply cut the space more finely in the local area when the number of node pairs exceeds a limit.

A comparison between the existing road network and cycling lanes and the envisioned cycling network in Barnet reveals a noteworthy distinction. The cycling networks are designed to encompass only a part of major trunk roads like Barnet Way and Watford Way, denoted by the orange lines in Figure 2, and to incorporate more sections of quieter and safer residential roads situated in proximity to these main arterial roads. This approach underscores an important insight: trunk roads may not always represent the fastest or most convenient option for cyclists. Instead, the proposed cycling networks emphasize the utilization of capillary neighborhood roads, offering a wider array of routes that can potentially be more efficient for cyclists. However, it is essential to acknowledge that the choice between trunk roads and residential streets as the preferred cycling route depends on specific origin and destination points, as well as individual preferences and safety considerations.

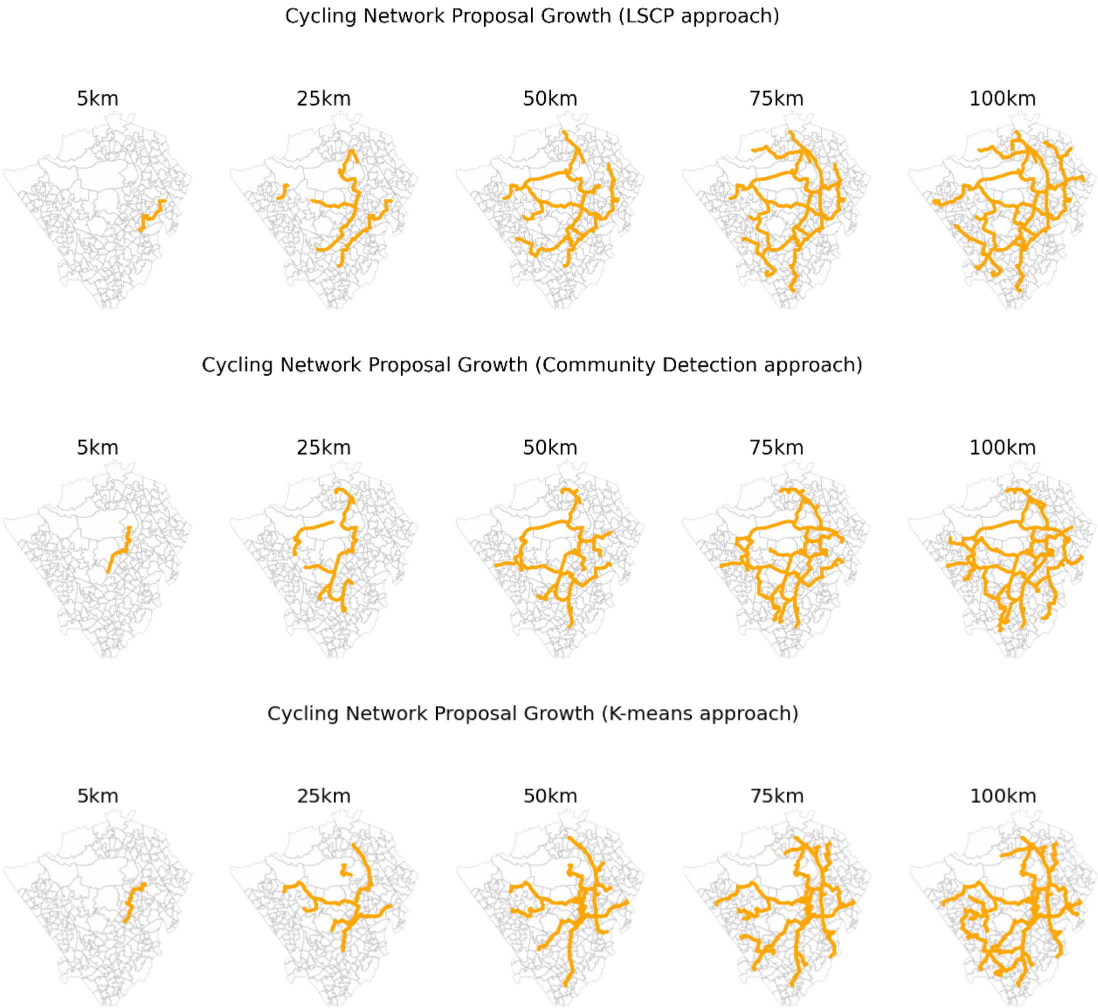


Figure 6. Cycling network growing process, total length 5km – 100km, in London Borough of Barnet.

In the scenario of investment constraints, the construction of cycling infrastructure networks should be prioritised in accordance with local economic capacity and historical levels of investment, taking into account the various needs of each borough (Norman, 2023). Based on the full road network planning from the previous step, we go for incremental growth of the network by weighted betweenness, referring to London's planned 1400km cycle network (Mayor of London, 2021), capping the total length of Barnet's cycle construction investment at 100km. Figure 6 shows the network growth from 5 to 100 km in segments of 25 km.

There are some common characteristics for the growing network on three approaches. Firstly, in the stage of 5km-25km, The proposed links developed first as a straight road running north-east to south-west in the area to the east of Barnet, and then several linear roads emerged in close proximity but disconnected from each other. In the 25km-50km stage, several adjacent straight roads were connected to form a tree-like structure, which gradually generated a net structure and extended from the eastern side of Barnet to the western side. Then in the 50km-100km stage, gaps in the network were filled and the branches of tree structures were expanding further.

For the K-means approach, its cycling network growth is inclined more to a tree structure, edges form closed loops mainly at 75-100km stage, whereas the results for the other two approaches have formed a large loop by 50km stage, and then smaller cyclic loops are formed around the periphery. Looking at the structure of the network results at the 100km stage, the network on the LSCP approach has a more spatially spread out network with wider loop coverage, while the one on community detection approach has a comparatively more compact network. This could be due to the fact that the LSCP method maximises the use of facility resources and the facility points are relatively far apart. The community detection approach has a large number of nodes selected and the nodes are more concentrated in the south-eastern part of Barnet, where they are also closer to each other. In contrast, it is more difficult to find the network pattern on the K-means approach. Possible reasons might include that the pre-defined number of clusters (k=82) was not optimised by validation, but was taken for comparison with community detection. Thus the classification of road intersections was not able to present the spatial structure of road networks properly. We also reflect on the way that how three approaches are differentiated. The nodes are selected based on population demand and the road network's structure and the difference is clear between the three approaches before generating a complete network, however, in the process of creating the network, the priority of cycleway edges depends on the safety-weighted betweenness, which means it depends on the structural features of the complete network proposal. As a result, it becomes difficult to capture the characteristics of each of these three approaches as the network grows.

4.3 Performance of networks

Following the network growth segmentation the same as above, we calculated six measurements in each stage and compared the performance of networks with growth. Growth strategy by road speed-weighted betweenness was used to order links' growth. Segmentation was performed at 5km intervals from 5km to 100km. The metrics include coverage, connectivity, directness, efficiency, population coverage rate and accessibility. The details of defining metrics are in the Methodology part. In Figure 6, the three lines show the performance of greedy triangulation networks under three node selection strategies respectively. The three lines end up in networks with a total length of 100km, but all the resulting networks are different, some of which every 25 km length can be seen in Figure 6.

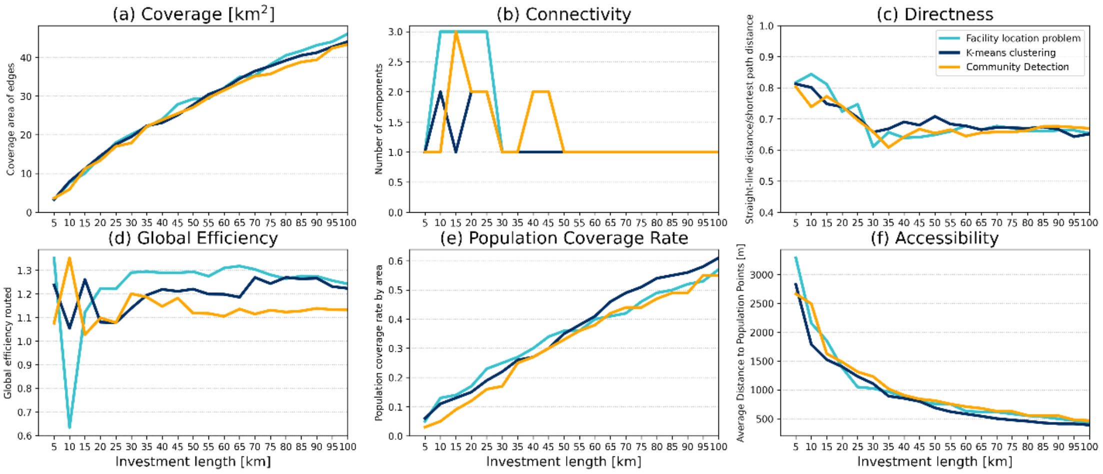


Figure 7. Comparison of network performance in the growing process

From Figure 7(a), the coverage maintains a constant growth and the curves are convex, indicating that the coverage area grows rapidly at the early stage. The changes in the coverage area of the three strategies are pretty close to each other, suggesting that the cycling network in the 5km-100km stage is reasonably distributed without overlapping. The total area of Barnet is 86.74 km^2, and ultimately in the 100km stage, the coverage area reached about half of the area. The growing trend of population coverage rate is slightly similar to the increase in coverage as can be seen in Figure 7(e). It maintains stable growth and there are no major differences between the three strategies. The LSCP (facility location problem) strategy line is slightly above the other two in the first half, and in the latter half, the K-means strategy has a better capture of population coverage. The population coverage rate reaches about 60%, which compares with a coverage area of about 45% in the same period, indicating that all three strategies in this methodology capture the spatial distribution of population demand well.

From Figure 7(b), connectivity lines fluctuated among one to three components and then decrease to combine into one large component. Overall, they have a low level of fragmentation. Networks in the LSCP approach and K-means clustering approach reached the combination of one component after 30km length, but the one in the community detection approach was more unstable, which might indicate that links in further areas might be prioritised when growing.

From Figure 7(c), directness was as high as 0.8 at first, then decreased in fluctuation and slightly increased after the lowest point, and finally levelled off to around 0.65. The high directness at the start point is due to the nearly straight link, then directness had a sharp decrease for the subsequently developed tree-like structure where the shortest path between node pairs can be quite convoluted. When more edges were added to form loops gradually, directness rose to a stable value, at around 0.66. Directness for the LSCP approach showed a fast drop until around 30km length, from 0.85 (l =10km) to 0.6 (l = 30km). The change points for the other two strategies were also around 30-35km, suggesting that this mileage might be a referable turning point for Barnet's cycling network construction. When the investment reaches this point, the directness performance will change qualitatively.

There are significant changes and fluctuations in global efficiency for the three lines in Figure 7(d). Note that as edges have smaller weighted lengths than their routing distance, especially for those with better safety classification, the normalised global efficiency can reach a higher value than 1. Overall, it indicates that the network is highly efficient at travelling between nodes. Global efficiency for the LSCP approach is overall higher than the other two. The global efficiency for LSCP has a large dip at the early stage while the other two increase and decrease at similar stages. After l $\approx$ 30km, all of the three tend to be comparatively stable and the values are around 1.1- 1.3.

From Figure 7(f), with the networks growing, accessibility, and the average distance from population points to the nearest cycling infrastructures, rapidly decline before the curve slope flattens out more and more. The one for the LSCP approach has the greatest degree of change, with l = 25km as the turning point, where 20km of cycleway investment before head rapidly reduces the average distance from 3,000m to 1,000m and then 80km (20km-100km) length of investment reduced the average distance to only 500m. This may be instructive for planners to set a threshold in order to balance the costs and benefits. Prior to this point, the return rate on investment is large, and accessibility to nearby cycleways increases rapidly as investment is made in cycling construction, however beyond this point, the return comes slower, and the service area of the cycleways may have more duplication of coverage, requiring more investment to achieve the same level of accessibility improvement.

In summary, the methodology in this study provides a way to design a cycling network design in a wide vision. The triangulation and tree-like network connecting scattered nodes are able to cover the main study area, betweenness growth strategy achieves fast coverage. Moreover, it is also found that the number of target nodes does not affect the trend and critical value of some indicators. For instance, in the 100 km length of cycling way construction, the directness values for all three approaches are leaning to around 0.65 and accessibility distance approaching 500 metres. The performance of the three methods in terms of global efficiency is different, and the network of the LSCP method shows overall better efficiency.

In a practical context, planned expansions of the cycling network have not reached out to Barnet yet and there is no published scheme or planned length for the cycling network that could be referenced. If we refer to the estimated cover rate for residents in Cycling Action Plan II, the proposal in this study can reach the goal of 40% of Londoners living within 400 metres of a qualified cycle network by 2030 (Norman, 2023).

5. Reflection and future work

5.1 Reflection and limitation

The research objective of this study is to construct a design framework. According to the results, there are some reflections both in the conceptual framework and practical work. Firstly, in terms of the framework idea, it needs a deeper consideration, of how could the design results present the necessity of selected intersection nodes. Should we focus more on selecting nodes among the street network and Points of Interest, or the growing strategy and algorithm which are more meaningful in real-life context? Also, the validation of network performance should also be justified more carefully. Two measurement metrics were built to connect back to the population demand assumption we had in the early stage of this study, however, they are not very effective in helping to identify different approaches to node selection. This problem might be attributed by the fact that what information could population data convey. As mentioned before in the Result and discussion chapter, traffic volume and trip data are more common data on this topic. It also matters a lot how we process the population information into the road network.

Then regarding the implementation of this methodology, there are several limitations on the choice of road factors, node selection rules and growth strategy. Max speed of the road was used as a factor to construct weighted length in this research method to influence the links of the network, however, road speed can be changed flexibly. For example, London is committing to reduce speed to 20mph, and 19 London boroughs have committed to this limit on all their road, with regard to road safety for pedestrians and cyclists and motorcyclists (Norman, 2023). Thus, we should note that the road speed factor is a dynamic factor. Cycling networks constructed on this basis are likely to lose relevance as circumstances change. More stable and deeper indicators should be focused on, such as road width, slope, and so on. Secondly, looking at the results of greedy triangulation networks, there are cases where three nodes are close to collinear, resulting in some paths nearly duplicating existing paths. It is more likely to happen when there are more nodes. This should be able to be solved by adding grid on the street network to help seed target nodes. Also, due to the data limitation, existing cycling lanes were not added to the proposal. Because the geographic information from TFL data is inconsistent with OSM road data, and if we look for the existing cycling ways in OSM data, the filtering results return rarely any cycling tracks in Barnet.

Last but not least, one of our research objectives is the equity of the cycling network. Whilst the growth strategy by betweenness results in the network starting to grow from a centre, so that the cycling facilities are still not able to reach out to further areas during the initial build-out period, and it is only after a series of build-outs that cycling services would gradually percolate through to these areas and become practical for them. There should be more innovation on how to develop a growth strategy for the network, based on the growth strategy by betweenness or closeness in previous research.

5.2 Future works

Firstly, more detailed future work could be done by combining more factors in different aspects of cycling infrastructure designing criteria, such as comfort and cohesion. One common way is to construct an equation that takes several factors as independent variables and to calibrate the parameters for the one combined weight.

Also, sensitivity analysis could be expanded further to test the robustness of this methodology and the findings in this study, such as by checking the performance changes of network results on varying street networks input. Different street patterns could be involved in data input, such as grid street patterns in New York and Barcelona and radial patterns in Paris. It would be valuable to compare the cycling network design results for cities with different street patterns and densities. Alternatively, it will also be helpful to assess how the results compare with the existing infrastructure and the cycling infrastructure plan declared by the London government.

Accessibility is an essential aspect of human capability, emphasizing the importance of respecting the rights of individuals to ride bikes, particularly prioritizing disadvantaged groups, as highlighted by Pereira, Schwanen and Banister (2017). Consequently, future research endeavours could provide insight into the cycling needs of different population groups. This includes examining minimum accessibility standards tailored to the specific requirements of each demographic and dissecting demand further by factors such as age group and gender proportion.

6. Conclusion

In conclusion, this study embarked on a comprehensive exploration of the methodology for designing a cycling network from scratch, with a specific focus on safety and equity considerations in the case of Barnet. It addresses various methods of extracting cycling demand from demographic data to target nodes of street network when trip or commuting flow data is not readily available. Thus, this approach ensures an equitable coverage of the diverse needs of different regions, employing methods like facility location problem to benefit a large number of residents.

Node selection, coupled with weighted road distance, are pivotal determinants of the network's layout. Road max speed was utilised as a weighting of links to address safety concerns in this study. Importantly, the flexibility of weighting factors allows for the incorporation of various road attribute considerations, further enhancing the adaptability of this network designing approach. Safety factors, as a result, influence routing choices between target nodes, with preference given to safer roads.

The application of Greedy Triangulation network algorithm for connecting target nodes balances between investment and travel time, while also incorporating global planning considerations for the entire study area. Then the growing process of network provides a practical solution for places with varying investment levels. The evaluation of network performance through several indicators revealed an overall improvement in connectivity, with directness and global efficiency fluctuating and before stabilizing at a more optimal level. Accessibility for residents also saw enhancement, albeit with varying return rates in the early and late stages of construction.

In summary, the results of this study, encompassing both the abstract network and routing network, provide valuable insights and inspiration for the establishment of a well-connected cycling infrastructure system in urban areas. Furthermore, the reliance on open-source resources and data makes the automation of this cycling network design process feasible, offering opportunities for reproducibility and application in other regions. This research contributes to the broader effort to enhance cycling infrastructure for safety, equity, and sustainability in urban environments.

Bibliography

Akbarzadeh, M., Mohri, S.S. and Yazdian, E. (2018) 'Designing bike networks using the concept of network clusters', *Applied Network Science*, 3(1), p. 12. doi:10.1007/s41109-018-0069-0.

Blondel, V.D. *et al.* (2008) 'Fast unfolding of communities in large networks', *Journal of Statistical Mechanics: Theory and Experiment*, 2008(10), p. P10008. doi:10.1088/1742-5468/2008/10/P10008.

Boeing, G. (2017) 'OSMnx: New methods for acquiring, constructing, analyzing, and visualizing complex street networks', *Computers, Environment and Urban Systems*, 65, pp. 126–139. doi:10.1016/j.compenvurbsys.2017.05.004.

Boisjoly, G., Lachapelle, U. and El-Geneidy, A. (2020) 'Bicycle network performance: Assessing the directness of bicycle facilities through connectivity measures, a Montreal, Canada case study', *International Journal of Sustainable Transportation*, 14(8), pp. 620–634. doi:10.1080/15568318.2019.1595791.

Buehler, R. and Dill, J. (2016) 'Bikeway Networks: A Review of Effects on Cycling', *Transport Reviews*, 36(1), pp. 9–27. doi:10.1080/01441647.2015.1069908.

Cardillo, A. *et al.* (2006) 'Structural properties of planar graphs of urban street patterns', *Physical Review E*, 73(6), p. 066107. doi:10.1103/PhysRevE.73.066107.

Clauset, A., Newman, M.E.J. and Moore, C. (2004) 'Finding community structure in very large networks'. doi:10.48550/ARXIV.COND-MAT/0408187.

Cycling and Walking Investment Strategy: Report to Parliament (2020). London: Dandy Booksellers Ltd.

Derrible, S. and Kennedy, C. (2011) 'Applications of Graph Theory and Network Science to Transit Network Design', *Transport Reviews*, 31(4), pp. 495–519. doi:10.1080/01441647.2010.543709.

Folco, P. *et al.* (2022) 'Data-driven micromobility network planning for demand and safety', *Environment and Planning B: Urban Analytics and City Science*, p. 239980832211356.

doi:10.1177/23998083221135611.

GREAT BRITAIN: DEPARTMENT FOR TRANSPORT (2020) *CYCLE INFRASTRUCTURE DESIGN*. S.l.: TSO.

Hull, A. and O'Holleran, C. (2014) 'Bicycle infrastructure: can good design encourage cycling?', *Urban, Planning and Transport Research*, 2(1), pp. 369–406. doi:10.1080/21650020.2014.955210.

Klobucar, M.S. and Fricker, J.D. (2007) 'Network Evaluation Tool to Improve Real and Perceived Bicycle Safety', *Transportation Research Record: Journal of the Transportation Research Board*, 2031(1), pp. 25–33. doi:10.3141/2031-04.

Latora, V. and Marchiori, M. (2001) 'Efficient Behavior of Small-World Networks', *Physical Review Letters*, 87(19), p. 198701. doi:10.1103/PhysRevLett.87.198701.

Lowry, M. and Loh, T.H. (2017) 'Quantifying bicycle network connectivity', *Preventive Medicine*, 95, pp. S134–S140. doi:10.1016/j.ypmed.2016.12.007.

Mackaness, W.A. and Mackechnie, G.A. (1999) 'Automating the Detection and Simplication of Junctions in Road Networks'.

Magnanti, T.L. and Wong, R.T. (1984) 'Network Design and Transportation Planning: Models and Algorithms', *Transportation Science*, 18(1), pp. 1–55. doi:10.1287/trsc.18.1.1.

Mahfouz, H., Arcaute, E. and Lovelace, R. (2021) 'A Road Segment Prioritization Approach for Cycling Infrastructure'. arXiv. Available at: http://arxiv.org/abs/2105.03712 (Accessed: 13 March 2023).

Mayor of London (2021) *Record-breaking growth in London's cycle network continues*. Available at: https://www.london.gov.uk/press-releases/mayoral/mayor-and-tfl-announce-work-on-four-new-routes (Accessed: 10 August 2023).

Natera Orozco, L.G. *et al.* (2020) 'Data-driven strategies for optimal bicycle network growth', *Royal Society Open Science*, 7(12), p. 201130. doi:10.1098/rsos.201130.

Nelson, A.C. and Allen, D. (1997) 'If You Build Them, Commuters Will Use Them: Association

Between Bicycle Facilities and Bicycle Commuting', *Transportation Research Record: Journal of the Transportation Research Board*, 1578(1), pp. 79–83. doi:10.3141/1578-10.

Norman, W. (2023) 'Cycling action plan 2'.

Olmos, L.E. *et al.* (2020) 'A data science framework for planning the growth of bicycle infrastructures', *Transportation Research Part C: Emerging Technologies*, 115, p. 102640. doi:10.1016/j.trc.2020.102640.

Pereira, R.H.M., Schwanen, T. and Banister, D. (2017) 'Distributive justice and equity in transportation', *Transport Reviews*, 37(2), pp. 170–191. doi:10.1080/01441647.2016.1257660.

Quinn, B. (2020) 'London cycling could increase tenfold after lockdown, says TfL', *The Guardian*, 6 May. Available at: https://www.theguardian.com/world/2020/may/06/cycling-could-increase-five-fold-in-london-after-lockdown.

Schoner, J.E. and Levinson, D.M. (2014) 'The missing link: bicycle infrastructure networks and ridership in 74 US cities', *Transportation*, 41(6), pp. 1187–1204. doi:10.1007/s11116-014-9538-1.

Shui and Chan (2019) 'Optimization of a Bikeway Network with Selective Nodes', *Sustainability*, 11(23), p. 6531. doi:10.3390/su11236531.

Szell, M. *et al.* (2022) 'Growing urban bicycle networks', *Scientific Reports*, 12(1), p. 6765. doi:10.1038/s41598-022-10783-y.

Tait, C. *et al.* (2022) 'Is cycling infrastructure in London safe and equitable? Evidence from the cycling infrastructure database', *Journal of Transport & Health*, 26, p. 101369. doi:10.1016/j.jth.2022.101369.

Zili, W. (2020) 'São Paulo Cycling Network Development Design: A Minimum Spanning Tree Approach', in *CICTP 2020. 20th COTA International Conference of Transportation Professionals*, Xi'an, China (Conference Cancelled): American Society of Civil Engineers, pp. 2916–2929. doi:10.1061/9780784483053.246.

Research Log

Date	Supervisors	Record
24/04/2023	Ioannis Toumpalidis	Next Steps: • Add Details to sub-questions about Identifying Demand, Community detection, and Algorithmic Design (which steps are planned to be discussed, reviewed or be developed) • For the processes that needs to be developed
25/04/2023	Hannah Fry, Ioannis Toumpalidis,	Next steps: A focused research question and objective. Literature review
01/06/2023	Ioannis	Progress: • Would like to focus on the spatial equity issue of cycling infrastructure in London • Proposed some methods: GA, Greedy triangulation. Next steps: • Justify research areas in London: which boroughs • Methodology: 1. Genetic Algorithm: define fitness function, greedy triangulation network generation; 2. Comparing different node selection solutions: GA started from randomly; MCLP; by betweenness, closeness. Coding work
15/06/2023	Hannah, Ioannis	Progress: assured the adjusted research topic and went through the literature review sketch. discussed the methodology: the GT algorithm I plan to use and the referenced

		codes. discussed the feasibility of Genetic Algorithm (NSGA-II) Ioannis created a Teams Channel and GitLab for updating work materials and codes notebooks. Next Steps: Workshop on SPACED and tech instructions will be given by Ioannis. I will continue codes work at the study area London Borough of Barnet and update it online.
30/06/2023	Ioannis	Ioannis, Jin Progress: SPACED workshop Did a simple greedy triangulation network generation (code) Next Steps: Keep going. Mini conference presentation
05/07/2023		Mini-conference
19/07/2023	Ioannis	Progress: Helped arrange the method and progress. To use community detection help select nodes
26/07/2023	Ioannis	Progress: Did three approaches to select points. Next: Validation part.