Bus Network Sketch Planning with Origin-Destination Travel Data

by

Cecilia A. Viggiano

M.S. in Transportation, Massachusetts Institute of Technology (2013)
B.A. in Environmental Analysis, Pomona College (2008)

Submitted to the Department of Civil and Environmental Engineering in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Transportation

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2017

Author .... Signature redacted
Department of Civil and Environmental Engineering
May 19, 2017

Certified by.............................. Signature redacted
Nigel H.M. Wilson
Professor of Civil and Environmental Engineering
Thesis Supervisor

Accepted by .... Signature redacted
Jesse Kroll
Professor of Civil and Environmental Engineering
Chair, Graduate Program Committee

Bus Network Sketch Planning
with Origin-Destination Travel Data

by
Cecilia A. Viggiano

Submitted to the Department of Civil and Environmental Engineering
on May 19, 2017, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Transportation

Abstract

Public transport plays an important role in urban mobility. Public transport planners seek to improve existing public transport networks to better serve existing passengers and recruit new passengers, particularly as demand patterns change with evolving demographics and land use. One strategy for network improvement is to add new routes, which can improve service by reducing circuity in the network. In large, complex, and multi-modal public transport networks, it is challenging to determine where new routes should be added. A systematic approach for incremental network improvements, such as adding new bus services, is needed. This research proposes a new approach to network-level public transport planning by combining origin-destination (OD) level analysis with new spatial aggregation methodologies, and develops a comprehensive framework for the identification of corridors for new bus services.

In the context of this framework, this dissertation contributes several new methodologies. First, it proposes a methodology for defining zones that reflect the spatial characteristics of a public transport network. This produces zonal pairs that are appropriate for OD level analysis of travel in the network. Second, the dissertation develops metrics and rules for the identification of OD pairs that can benefit from new bus services, and proposes methods for estimating the expected benefits of such services at the OD level. Finally, a new methodology for spatially clustering OD pairs into corridors is developed, based on trajectory clustering methods. This final methodology represents a new way of aggregating OD level information to accomplish the first step in bus network design: the definition of corridors for new services.

The framework is demonstrated for the identification of corridors for new bus services in the London public transport network. Bus stops and rail stations are clustered into 1,000 zones. A subset of zonal OD pairs with circuitous service are identified as candidates for improvement through new bus routes. An algorithm that clusters OD pairs into corridors for bus service is developed and applied. Several promising corridors are identified, and their potential is confirmed in post-analysis.

Thesis Supervisor: Nigel H.M. Wilson
Title: Professor of Civil and Environmental Engineering

Acknowledgments

I would like to thank my advisors: Professor Nigel Wilson, Professor Haris Koutsopoulos, and John Attanucci for following my progress and sharing their wisdom at every stage of this research. Thank you also to Professor Jinhua Zhao and Chancellor Cindy Barnhart for providing feedback at my committee meetings and in the periods between. And to Fred Salvucci, Mikel Murga, and Gabriel Sanchez Martinez: thank you for always being willing to listen and advise.

This research was funded by Transport for London and benefited immensely from the expertise of TfL's Network Development team, including John Barry and Alex Phillips. And thank you as well to Lauren Sager-Weinstein, Dale Campbell, and Andrew Hyman of TfL for their support.

Thank you to the Transit Lab students, past and present.

To all of the wonderful CEE staff including Ruth Yiu, Shannon McCord, Ginny Siggia, Kiley Clapper, and Kris Kipp: thank you for supporting me throughout my time at MIT.

Thank you to my friends, near and far, for endless encouragement, and to my water polo team, for keeping my spirits high.

Thank you to my parents, Sharon and Frank, and my sisters, Vanessa and Emily, for always believing in me.

Last of all, thank you to James for listening to me talk about my research for hours upon hours, for cooking me countless dinners, and for many, many walks around MIT.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

As the world's population increasingly concentrates in urban areas, the importance of urban transportation solutions grows. Public transport systems have been shown to improve accessibility in cities, increasing the locations people can move between in a given amount of time (Newman and Kenworthy, 1999). More broadly, well-planned public transport systems foster economic, environmental, and social benefits.

In a review of studies on the impacts of investment in transportation on economies, Bhatta and Drennan (2003) conclude that such investment generally results in economic benefits, including increased productivity and output, reduced costs of production, and increased incomes. In cities, agglomeration benefits produced by spatial concentration of economic activity, are of particular note (Graham, 2007). Evidence suggests that public transport plays a critical role in enabling the density of activity that produces these benefits (Venables, 2007). From an environmental perspective, shifting travel from single-occupancy vehicles to public transport, helps achieve energy conservation and emissions reduction goals (Shapiro et al., 2002). Therefore, many cities seek to encourage public transport use to promote sustainability. Effective public transport systems also support equity, as they enable individuals who cannot drive or cannot afford to own a car to be mobile.

Good public transport systems are fast, convenient, and comfortable. Planning an effective public transport system includes decisions ranging from strategic to operational. At the strategic level, planners determine the public transport network design, consisting of routes and lines and the connections between them. Good network design can be measured across many dimensions. Generally, metrics take either the passenger perspective, the operator perspective, or the system perspective.

Designing public transport networks that serve passengers well, adhere to operator constraints, and improve the overall transportation system is challenging. In cities that have existing public transport networks, the network design needs are unique. Most existing methodologies for public transport network design aim to identify an optimal set of routes or lines without serious consideration of the existing network. While these methods can be effective in cities without existing public transportation, or in cases where planners wish to completely re-design the network, they do not serve the need for incremental improvement of established networks. In these networks, planners could benefit from guidance on strategies that can improve the network.

One important aspect of network design is circuity. Reducing deviation from the shortest path and/or eliminating interchanges can improve the quality of service. This dissertation presents a framework for bus network sketch planning. Specifically, the framework identifies opportunities for new bus services in existing networks to reduce circuity and eliminate interchanges.

Section 1.1 defines sketch planning. Section 1.2 discusses the factors that motivate this research, including reducing circuity, and the need for network planning methodologies that can be applied to existing, multi-modal complex networks. Section 1.2 also identifies a special opportunity for new approaches to network planning using newly available information on complete journey origin-destination (OD) level travel. A case study, the London bus network, is introduced in Section 1.3. This if followed by descriptions of the objectives, approach, and major contributions of this research.

1.1 Sketch planning

Network design often begins with a sketch planning phase, which is the focus of this research. Sketch planning methods identify and select opportunities for network design changes, such as the introduction of new routes. Sketch planning often includes efficient methods to estimate costs and benefits of alternatives and select between them. In sketch planning, new routes or lines may be defined in an abstract way, leaving the exact design to later phases in the planning process. While many sketch planning approaches assume that alternatives are already defined, and therefore focus only on the evaluation of these alternatives, this dissertation proposes a framework for both the generation and evaluation of alternatives.

1.2 Motivation

A methodology that can be applied to an entire public transport network to identify opportunities for improvement fills a gap in the public transport planning literature. Traditional approaches to public transport network design use partial optimization methods to identify the set of routes or lines which best serve a specified demand matrix, be it theoretical or observed. These approaches may produce a network that bears little resemblance to the existing network. As a result, in established networks, planners generally have little use for such results.

A complete redesign of an existing network is typically impractical for several reasons. First, it is costly. For rail networks, which require significant fixed costs to build, such a redesign is rarely considered. Even for bus networks, there are costs associated with new bus stop infrastructure, maps, and information provision. Of even greater importance, network changes are disruptive to passengers, who are accustomed to the existing network and will have to adapt to alterations. Passengers may also have chosen their home locations based on the existing public transport system. Similarly, businesses often take transportation infrastructure into account when selecting locations. For these reasons, a total overhaul of the public transport infrastructure is generally impractical and is likely to be met with significant opposition. This is particularly true, because the expected results of public transport network design recommendations are subject to assumptions inherent in the methodologies, and to the inputs used, and therefore are always subject to uncertainty. Planners are (rightfully) reluctant to seek uncertain benefits in the face of such high costs.

Instead, in established networks, planners typically make incremental changes over time. While the issues related to cost, information, opposition, and uncertainty may persist with more minor changes, the magnitude of each issue is generally reduced. Planners are willing to endure these costs to make needed changes to the network in light of changing demand. For good planning decisions, the benefits from such changes outweigh the associated costs.

Improvements to the network help serve current passengers better and can draw new passengers to the network from other modes. In the longer term, these improvements may also induce new journeys. Network improvements can take many forms. Ceder (2007) groups the design strategies planners can use to improve service into five categories:

1. New forms of service such as express, zonal, and on-demand

2. Increased network coverage or span-of-service coverage

3. New and adjusted routing to connect routes, lines, or locations

4. New and adjusted scheduling, including interlining, coordination, and frequency

5. Improved amenities such as passenger facilities or vehicles

Determining how to apply these strategies in large public transportation networks is a complex task. Planners may use their own knowledge of the system, public input, or the results of performance analysis to decide which actions to take. They may also react to political pressure. Public transport agencies are unlikely to have the resources to analyze more than a small set of alternatives when deciding which changes to make.

For bus planning in particular, trial-and-error approaches are common (White, 1995). Because changes to bus services are significantly less costly than in the rail network, planners may pilot routes in order to assess the impacts. While this can be effective, it is clear that in large, complex networks, it is an inefficient method to identify those corridors that will provide the most benefits. This leaves an unaddressed need to systematically identify promising opportunities for improvement.

This dissertation provides a general framework for sketch planning based on systematic, network-level analysis to be applied in existing networks, and a detailed methodology for implementing one particular strategy identified by Ceder: new and adjusted routing to connect routes, lines, or locations. Ceder distinguishes between this strategy, and actions that increase coverage. While increasing coverage may also require new routes, the objectives are distinct. Increasing coverage expands the network to areas that were previously not served. The framework proposed in this dissertation is intended to be applied to an area that already has public transportation coverage. In areas that already have coverage, new services can either parallel existing ones, or follow new paths. While there may be some cases in which adding service that parallels existing routes or lines is desirable (for example, adding a bus route that parallels a rail line provides benefits to passengers who prefer bus to rail), generally redundant service is avoided. This framework focuses on the introduction of new routes to reduce circuity in the public transport network. The implementation of this strategy requires a network perspective, as is taken in this research.

1.2.1 Reducing circuity

Circuity in public transport can be measured in different ways. Often, it is measured as the level of deviation from the shortest path. Many measures incorporate the number of interchanges (or transfers) required to make a journey, and some include the number of stops. Public transport systems cannot provide direct point-to-point service for all origins and destinations. Thus circuity will never be eliminated. However, reducing circuity is beneficial.

Huang and Levinson (2015) found evidence that commuters locate themselves in places where their commute trip will have low circuity, with circuity measured as the ratio of public transport distance to straight line distance. In addition, they found that public transport accessibility measures for 36 metropolitan areas were negatively correlated with circuity. This suggests that not only is circuity undesirable for specific trips, but it degrades the overall quality of the public transport system. In support of this notion, many graph theoretic metrics designed to assess public transportation networks characterize them in terms of circuity or directness (Garrison and Marble, 1964; Gordon, 1974; Derrible and Kennedy, 2010).

Interchanging, in particular, is widely regarded as a deterrent from public transport attractiveness and quality of service. Understanding the interchange penalty incurred when journeys require an interchange is the subject of a large body of research. While the penalty can be reduced by planning and operating actions such as coordination of service or high frequency of service, eliminating interchanges altogether is beneficial.

Introducing new services to eliminate circuity requires investment of money and resources. A methodology that systematically assesses and quantifies the benefits of new bus routes in terms of reduced circuity will help planners make decisions about adding routes to existing networks.

1.2.2 Bus planning in a multi-modal context

Developing such a methodology that can be applied to a large, complex, multi-modal public transport network is particularly challenging. Public transport networks often include heavy and light rail services that operate in conjunction with bus services. Bus routes may feed the rail system, or provide stand alone services. In some places, bus is an alternative to rail, particularly serving origins and destinations that require multi-stage rail journeys or serving journeys that originate (end) between rail stations. Buses typically stop more frequently than rail services, and therefore provide denser coverage that is especially important for passengers who are unable (or unwilling) to walk significant distances to access a rail station. A methodology for bus planning that accounts for multi-modal public transport journeys and origins and destinations served by multiple public transport modes is needed.

1.2.3 An opportunity for a data-driven approach

Currently, there is an opportunity for significant improvement in bus planning methods presented by the increasing availability of travel information at the origin-destination (OD) level. Automatic fare card (AFC) and automatic vehicle location (AVL) data are routinely collected in many public transport networks. Recently, methodologies have been developed to infer complete journey information from these data sources (Chu and Chapleau, 2008; Gordon et al., 2013). These methodologies can infer destination and interchange information for a large share of boardings that use fare cards, often upward of 70% (Gordon et al., 2013). Compared to OD data collected from surveys, which is generally collected for a small number of passengers on a few selected days, this vastly increases the sampling rate of public transport OD information, improving estimates of OD-level demand and travel time, and enabling these estimates to be made at increasing levels of detail.

OD data is important for evaluating circuity. It provides information on the paths individuals choose, including the prevalence of multi-stage journeys. When there are multiple paths serving an origin and destination, journey data provides insight on the choices individuals make. For example, it can identify cases when passengers choose circuitous bus paths even when more direct alternatives exist.

Detailed OD matrices for modes outside of public transport rely on emerging data sources and methods. GPS-equipped mobile devices can provide traces of individuals' movements and researchers are developing methodologies to infer the mode of travel from these traces (Stenneth et al., 2011; Zheng et al., 2008; Patterson et al., 2003). While these data sources and methodologies are promising, some hurdles remain before they can be widely applied. There are data quality, missing data (mobile devices are sometimes off or not reporting location), and data privacy and ownership issues that are the focus of ongoing study.

Making use of these existing and emerging information sources for planning purposes requires new methodologies, including metrics and methods for evaluation and aggregation. This dissertation develops a framework for the use of OD-level information in bus planning. In this framework, detailed OD information for public transport journeys is critical, while detailed information about travel on other modes can enhance the application of the framework.

1.3 London as a case study

The approach developed in this dissertation is applied to London's bus network, as an example. The London example contextualizes the need for a bus planning framework that can be applied to an entire network to identify opportunities for new services. London is densifying and growing, and has a vast and complex public transport network. Leaders in London recognize the need to support population and economic growth with transportation services. A framework that guides the growth of the bus network helps address this need.

Like many urban areas, London has experienced extensive population growth in recent years. From 2001 to 2011, the population grew at an average rate of 87,000 people per year. The population in 2011 was 8.2 million and is projected to continue growing to 9.3 million by 2021 and 10.1 million by 2036 (Greater London Authority, 2016). Some population segments are growing faster than others. The population over age 64 is expected to increase by 64% in the 2011 to 2036 time frame (Greater London Authority, 2016). The economy is also growing. The number of jobs in London is projected to increase from 4.9 million in 2011 to 5.8 million in 2036 (Greater London Authority, 2016). Overall incomes are growing and expected to continue to grow, but the distance between low-income and unemployed individuals and high-income earners is also increasing (Greater London Authority, 2016).

"The London Plan," (2016) authored by the Greater London Authority, states the goals of the Mayor of London, including ensuring that London is internationally competitive and establishing London as a world environmental leader. The document emphasizes maintaining a good quality of life for all Londoners and specifically notes the importance of a "providing a transport network enabling easy access to jobs, opportunities and facilities while mitigating adverse environmental and other impacts."

London has an extensive public transport system that plays a critical role in urban mobility. The system includes the Underground, which opened in 1863 and serves 4.8 million passenger journeys per day (Transport for London), as well as the Overground, trams, Docklands Light Railway, and many National Rail lines. However, the greatest number of passenger journeys are made on the bus network, which carries approximately 6 million passenger journeys on the average weekday on over 700 bus routes (Transport Committee of the London Assemply, 2013).

Public transport, in general, and the bus network in particular, play vital roles in mobility in London. Nearly 50% of Londoners use the bus at least two days per week (Transport Committee of the London Assemply, 2013). Demand for bus services grew from 1999 to 2013. In that period there was a 64% increase in bus journeys. However, from 2014 to 2017 bus journeys decreased by 5.4%, correlating with decreases in bus speeds in the same period (Transport for London, 2017). The greatest declines in bus speeds and ridership were observed in central London, and the 2017/2018 Transport for London budget states as an objective: "Matching bus capacity with demand by reducing the underused services in central London and reallocating them to where they are needed" (Greater London Authority, 2017).

Strategic planning is needed to ensure that the bus network continues to support mobility and attract new demand. "The London Plan" (2016) recommends "regular review of [the] bus network to cater for population, housing and employment growth, maintain ease of use, attractive frequencies and adequate capacity, reliable services, good coverage, effective priority and good interchange with other modes." It also emphasizes the need to improve links between town centers in Outer London including through orbital and radial bus services. The Transport Committee of the London Assemply (2013) reported that London boroughs are concerned that bus planning is primarily done on a route-by-route basis. They suggest there is a need for full-network area-based planning and evaluation of routes that cross borough boundaries.

Given the vast size of the London public transport network, and the bus network in particular, ad hoc and trial-and-error approaches to planning are likely to be costly and have only limited efficacy. Planners recognize the importance of network-wide sketch planning to identify needs for bus services that cross borough boundaries and serve orbital journeys. Moreover, London has the public transport data needed to implement a data-driven planning approach. London collects extensive AFC and AVL data. The Oyster card, a smart payment card that was first issued in 2003, was used by approximately 80% of rail passengers and 90% of bus passengers by 2013 (Gordon et al., 2013). Data used for the London application of the framework is discussed in more detail in Chapter 7.

1.4 Objectives

The objective of this research is develop a framework for the use of OD-level data in sketch planning. More specifically, a methodology is required that identifies opportunities for new bus routes that can be added to an existing network to provide incremental service improvements by reducing circuity and interchanges. The methodology should systematically analyze the full network to identify opportunities. It should not require planners to specify alternatives, but rather should identify them based on the analysis. It should quantify expected benefits of the corridors identified to help planners choose between potential corridors. Finally, the method must be flexible in order to account for differences between cities and networks and variation in planner priorities.

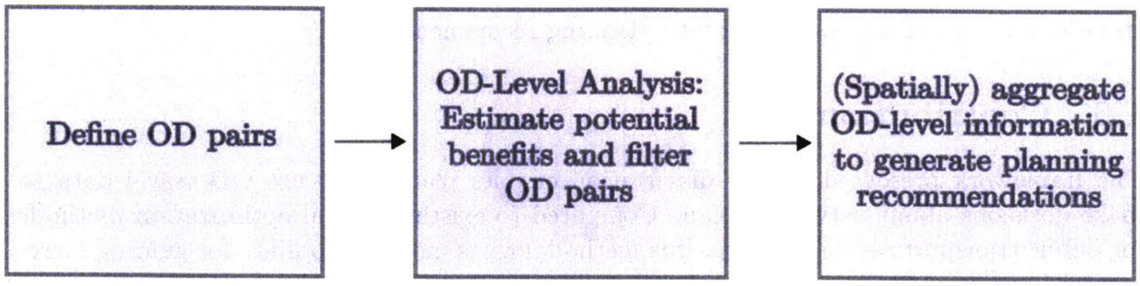

Figure 1-1: Approach

1.5 Approach

This dissertation proposes a framework for the use of OD travel information in planning consisting of three steps, summarized in Figure 1-1. The general approach can be applied to a variety of planning objectives, such as improving coverage, improving reliability, or addressing crowding. The OD-level analysis and the type of aggregation will be particular to each objective. Here the approach is applied to bus network planning and specifically to the addition of new bus services to an existing network.

The first step is to define the network as a set of OD pairs, which requires the definition of zones. The second step is to perform analysis at the OD level. The analysis characterizes OD pairs using metrics that quantify their performance and opportunities for improvement in terms of a specified planning objective. At this stage, OD pairs can be filtered based on these characteristics. In the specific methodology proposed in this dissertation, OD pairs are assessed in terms of the directness of existing services, and the expected benefits of a new direct bus service are estimated for each pair.

While OD-level analysis captures current passenger experiences adeptly, it is usually too disaggregate to be used directly to inform planning decisions. The third step in the approach aggregates the OD-level information to a unit of analysis that is appropriate for bus network planning. For the identification of opportunities for new bus service to reduce circuity, OD pairs are aggregated into corridors, each of which can be served by a single bus route. Expected benefits are estimated at the corridor level and used to prioritize the corridors identified.

Distinctive features of the approach, which makes it useful for many network planning decisions, are the analysis at the OD-level, and the aggregation of OD-level information to generate corridors. This is a departure from more traditional analysis of either existing routes or zones. Traditional corridor analysis of existing routes or groups of routes, is unable to analyze multi-stage journeys or the impacts of path alternatives outside the corridor. Zone level analysis usually considers services within a zone (such as density of stops or stations) or analyzes journeys entering (or leaving) a zone. This can miss important OD-level effects. For example, a zone may have a high density of bus stops but is served only by east-west bus lines. If individuals wish to travel north-south, they may have to endure longer access distances and/or in-vehicle travel times. Whenever performance depends on both origin and destination, and particularly in any system with multi-stage journeys and multiple paths, an OD approach can capture passengers' experiences much more fully.

While there are existing examples of the use of OD-level data to understand performance and passenger experience (See Chapter 2), the methodologies generally do not extend to service planning. By aggregating OD pairs into potential corridors, this method directly translates OD-level understanding into planning recommendations.

1.6 Contributions

The framework presented in this dissertation enables planners to use OD travel data to make decisions about network design. Compared to existing partial optimization methods for public transport network design, this methodology is more appropriate for guiding incremental improvements in existing networks. The framework uses granular OD travel data, which enables accurate and detailed analysis of performance and develops a procedure for aggregating this information to a meaningful level in order to use it for planning decisions, enabling these decisions to be made based on a systematic understanding of the full network.

The framework includes a method for clustering individual journey data into zonal OD pairs. These zonal OD pairs form the building blocks for the definition and selection of corridors. The corridor identification algorithm developed in this dissertation generates alternatives for bus services. This is an improvement over ad hoc methods for generating potential corridors for new services.

As discussed in Section 1.2.1, reducing circuity is beneficial in public transport networks. This research contributes methodologies for evaluating OD pairs in terms of directness of service. In addition, it proposes methods for quantifying the expected benefits of improved directness for OD pairs and corridors. The dissertation also discusses how information about a particular network, such as access distance distributions and desired bus route length can be used to adjust parameters used in the OD analysis and corridor identification processes. This results in a framework that is sufficiently flexible to be applied to many different networks.

1.7 Dissertation outline

Chapter 2 summarizes prior research on public transport network design, sketch planning, analysis of OD travel data, and clustering and aggregation methods. An overview of the framework is provided in Chapter 3. Chapters 4 through 7 discuss the zonal definition, OD analysis, and corridor identification methodologies in more detail. The application to London's public transport network is presented in Chapter 7, and concluding remarks are offered in Chapter 8.

Chapter 2

Literature review

The design of public transport systems is the subject of a large body of research. In the literature, public transport network planning is typically divided into five steps: network design, frequency setting, timetable development, vehicle scheduling, and crew scheduling (Ceder and Wilson, 1986). Most methodologies address a single step in this planning process, although Schöbel (2017) proposed an integrated method for line planning, timetabling, and vehicle scheduling.

The first step, the design of routes, is the subject of the framework presented here. The design of routes can be at the route, group of routes, or full network level (Ceder, 2007). The objective of this research is to design a framework that can be applied to a full existing network to identify corridors for new bus services. There are many approaches to full network design, summarized in Section 2.1. However, these methods generally do not fully and realistically capture the intricacies of real, complex networks. Moreover, they are intended for complete network re-design, which is rarely the objective for planners in existing, well-established public transport networks.

In existing networks, planners are more likely to employ sketch planning methods to efficiently compare alternatives for incremental network improvements. Examples of such methods are summarized in Section 2.2. These methods generally assume that a set of alternatives exist, and focus on selecting the best alternative(s) from the set. The framework presented here improves on these methods by systematically generating and evaluating alternatives for new bus services.

The framework relies on new origin-destination (OD) level travel data, including data from automated fare collection (AFC) systems. Smart card data has been used to gain an improved understanding of passenger behavior and travel patterns, as discussed in Section 2.3. Methods that infer complete multi-stage journeys from automated vehicle location (AVL) and AFC data enable analysis of OD level flows. The framework extends research that uses AFC and AVL to improve understanding of travel behavior, by using this data to inform planning decisions.

Achieving this objective requires aggregation of discrete OD-level data. The methods proposed in the framework build on methodologies including graph theoretical models and spatial and non-spatial clustering techniques described in Section 2.4.

2.1 The Transit Network Design Problem

The Transit Network Design Problem (TNDP) is usually formulated as an optimization problem consisting of an objective function and constraints. Some methodologies are tailored for a specific mode, either bus (Ciaffi et al., 2012; Bielli et al., 2002) or rail (Guan et al., 2006), while others have been developed for general application to public transport networks (Bagloee and Ceder, 2011). The methodologies aim to identify a set of optimal or approaching-optimal routes. In general, these methods assume the network will be designed from scratch. Advancements in both computing and algorithm development have allowed partial optimization and heuristic methods to be demonstrated for actual networks. However, Ceder (2007) notes that in practice, partial optimization methods that researchers have developed are rarely applied. He states that this implies a need for methods that are "more practical and less complex".

The inputs to the TNDP are OD demand matrices and network topology (either idealized or existing roads and physical barriers). Early versions used ideal networks rather than real inputs. Hasselstrom (1981) developed a method designed to use detailed OD matrices. Most methods also incorporate demand models to assign demand to proposed routes or lines. Hasselstrom (1981) originally used a direct demand model. Later, a route choice demand model was added to his method (Jansson and Ridderstolpe, 1992). Yan et al. (2013) incorporated stochastic travel time estimates to make the demand assignment and evaluation process more realistic, and Ng and Lo (2016) proposed a robust model, using lower, mean, and upper bound demand estimates in network design.

Network topology may be represented as a set of potential nodes and links, or may be a full representation of actual road topology (Guihaire and Hao, 2008). Hasselstrom's method requires users to specify terminals of routes, as does the method proposed by Baaj and Mahmassani (1995). Ceder (2001) uses heuristics to develop a skeleton network. Similarly, Carrese and Gori (2002) assume a skeleton network and develops a method to identify feeder routes.

Good public transport network design can be defined in a variety of ways, which is reflected in the diversity of objective functions in different methodologies. Ceder (2001) notes that there are multiple perspectives on network design: that of the operator, that of the passengers, and that of the community. His method defines metrics for each of these perspectives and optimizes with respect to all three. Travel time or travel time savings are commonly used as the criterion for optimization, while Ceder's method also includes waiting time, empty seats, and deviation from the shortest path. Most methodologies also include a set of constraints including fleet size, line length, number of lines, coverage, and percent of demand that is unsatisfied (Yan et al., 2013). Performance aspects, including directness, are sometimes included as either constraints or objectives (Guihaire and Hao, 2008).

A variety of approaches have been proposed to solve these problems, including neighborhood search algorithms, and evolutionary algorithms (Guihaire and Hao, 2008). All methodologies make use of heuristics to simplify either the input topology or the optimization problem. As a result they are considered partial optimization approaches (Guihaire and Hao, 2008). Bielli et al. (2002) used geometric shapes to define a set of possible routes with emphasis on routes connecting large demand modes. Cipriani et al. (2012) defined three types of routes (high-demand direct routes, connector routes, and existing routes) to guide the optimization process. Bagloee and Ceder (2011) used data to inform the heuristic step. They clustered demand to identify locations for bus stops.

While many methods are limited to theoretical applications, some have been applied to actual networks. Bagloee and Ceder (2011) applied their method to the Winnipeg bus network and to the Chicago rail network. Guan et al. (2006) applied theirs to a simplified version of Hong Kong's MTR network. Ciaffi et al. (2012) defined a skeleton and feeder network in their methodology which identifies a set of routes to serve a given station. They applied the methodology to a suburb of Rome currently served by 15 routes. They report a running time of approximately 15 hours for this application.

Most methodologies can be applied to existing or new networks, but are intended for full network redesign. As an example, Cipriani et al.'s method considers existing routes in the choice set of all possible routes, but the method is not intended to augment or improve the existing services, but rather find the best set of routes to serve as a full network. In the application of their framework to Rome, only 20% of existing routes were kept. The design from scratch or full network redesign approach is one reason that the methods in the TNDP literature are rarely applied.

2.2 Sketch planning and decision support tools

Much more commonly applied are sketch planning methods and decision support tools for network planning. Sketch planning consists primarily of efficient and effective methods to evaluate early stage alternatives, such as new routes or lines or extensions. These methods are designed to help planners systematically choose between alternatives. Cascetta and Carteni (2014) developed a way to use EU standards to evaluate public transport proposals. They developed measures to evaluate performance including accessibility, travel time, comfort and security and propose methods to estimate each metric for planned projects.

Deakin et al. (2004) present a methodology developed by Bay Area Rapid Transit (BART) planners for initial sketch planning consisting of demand estimation, cost approximation, identification of possible station locations and access needs, evaluation of impacts on the existing system, and consideration of community support. This evaluation framework serves as an initial filter to determine if the project merits further study. This provides a systematic evaluation of proposed services, but does not include a method for identifying opportunities for new services. In a rail network like BART, the options for new service are likely limited, making this aspect less important. However, in large, complex bus networks, the opportunities for network expansion can be vast.

One example of the use of performance indicators to identify improvement opportunities comes from highway infrastructure planning in Minnesota. Planners prioritized existing corridors based on demand, growth trends, and connections to regional trade centers. Prioritized corridors were evaluated on how closely actual speeds matched target speeds along the corridor. This process identified corridors for improvement. Planners then identified improvement strategies for these corridors ranging from construction of new roads or public transport services, to travel demand management, and pricing (Zemotel and Montebello, 2002).

Another framework for incorporating performance measures into transportation planning is presented by Miller et al. (2013). They define livability indicators and propose methods for weighting and aggregating the indicators at the community level (Miller et al., 2013). They suggest that these indicators can be used in transportation policy and planning to track performance and select between alternatives. Shah et al. (2013) present a diagnostic tool for transportation networks at the metropolitan level. They make comparisons across cities to benchmark performance and identify key problems and discuss the planning implication of these results.

Other decision support tools do not make recommendations about specific improvements, but rather make recommendations about network design more generally. Daganzo (2010) proposed a methodology to determine the appropriate network shape (grid or hub-and-spoke) and modes (bus, BRT, or metro) to produce desired performance, given network size and demographics. Badia et al. (2016) compared the performance and cost of four different network structures (radial, direct trip-based, grid, and a hybrid transfer-based system) given various idealized spatial travel patterns. Compared to the framework presented in this dissertation, the recommendations from these methodologies require more interpretation and analysis to be translated into planning decisions, particularly in networks that already have public transport service.

2.3 New sources of travel data

New sources of data present an opportunity for new sketch planning and decision support tools. While automated vehicle location (AVL) systems have been incorporated in public transport systems for several decades, the early 2000s saw a rapid adoption of contactless smart cards for public transport fare payment (Deakin and Kim, 2001). Automated data collection systems for AVL and smart card data have been lauded as a source for better estimates of existing measures and the development of new measures (Wilson et al., 2009). Trépanier et al. (2009) demonstrated the estimation of a variety of supply-oriented measures (vehicle-kilometers, vehicle-hours), demand-oriented measures (passenger-kilometers, passenger-hours), and measures that combine supply and demand (bus occupancy) from smart card AFC systems. Uniman et al. (2010) developed new metrics for service reliability based on AFC data. Pelletier et al. (2011) affirmed that smart card data can be informative for strategic, tactical, and operational planning of public transport systems.

Utsunomiya et al. (2006) combined smart card data with personal information provided by passengers to analyze typical patterns for different user groups and develop profiles for users of specific stations. Morency et al. (2007) analyzed the regularity of daily patterns and classified passengers based on their public transportation usage. Bagchi and White (2005) analyzed smart card churn to inform targeted campaigns to retain passengers. These studies can inform decisions about advertising, promotions, and fare structure.

Other studies provide information that can inform operations and service planning. Csikos and Currie (2008) analyzed arrival patterns at heavy rail patterns to understand how unreliability affects waiting time. Zhu (2014) used AFC data, including passengers' station entry and exit times, and train tracking data to assign passengers to specific trains. Guo (2008) used AFC data to estimate transfer penalties and explain path choice.

Smart cards record passengers' entries and (in some cases) exits from the public transportation system. Several researchers have developed methodologies to reconstruct individuals' itineraries, inferring origins, destinations, and transfers, based on AFC and AVL data (Chu and Chapleau 2008; Gordon et al. 2013). In networks where smart card usage is prevalent, these itineraries can provide a comprehensive picture of public transport travel. This creates opportunities for analysis of full, multi-stage journeys.

Vanderwaart (2016) developed a framework for using complete OD-level journey data for service planning. The framework consists of five steps. First, target locations are selected based on economic, demand, accessibility, and congestion factors. Once target locations are identified, they are analyzed in terms of speed, mode, transfers, journeys per capita, and travel time variability. Then, service changes are proposed either in terms of system structure or system performance. Finally, the service change proposals are evaluated in terms of costs and benefits. She also emphasized that the service planning framework should include a review of service changes after they are implemented.

Compared to Vanderwaart's research, the framework presented here focuses on generating service alternatives directly from the data, rather than relying on planner judgment to interpret statistics and maps. This makes the framework presented here more appropriate for very large and complex networks. To do so, this framework focuses on a specific service change: the addition of new bus routes, rather than the diverse service changes considered in Vanderwaart's framework.

Mobility data is also increasingly available from GPS-equipped mobile devices, primarily mobile phones. Hung et al. (2015) developed a methodology to infer complete trajectories from GPS-equipped mobile devices. Calabrese et al. (2010) used mobile phone data to analyze travel patterns of people attending special events. This type of data is not widely available for all cities due to ownership and privacy issues, and is not used in the implementation of the framework presented here, as it is not strictly necessary. However, mobile phone data could augment the framework by providing additional information on demand outside the public transport system and on the access and egress portion of public transport journeys.

With the exception of Vanderwaart (2016), the existing research uses OD travel data primarily to understand travel behavior and patterns rather than to make planning decisions. The framework presented here focuses directly on the planning implications, using OD travel data to generate corridors that are candidates for new bus service.

2.4 Aggregation, clustering, and statistical methods

In order to use OD data for sketch planning, methods are needed to convert the data into planning recommendations. Ackoff (1989) summarizes this process in four steps, the progression from data to information to knowledge to wisdom. Ackoff (1989) defines data as the most direct representation of objects or events. Information is also a representation of these objects or events, but the representation has become more compact and useful through data processing. Knowledge is the use of information to generate instructions, answering questions about how to do something. Finally, understanding is the explanation for the instructions, an answer to questions about why to do something.

Andrienko et al. (2008) delve into the process of converting data to information in a summary of methods for visualizing large spatio-temporal data sets. They note that some data can be visualized through direct depiction, with each record presented visually, but large data sets become illegible when all data is depicted. In these cases, researchers use two types of methods to visualize data: summary methods and pattern extraction. Summary methods consist of aggregating, generalizing, and in some cases sampling to extract abstract summary statistics. These statistics may be summarized in plots or figures, but the spatial elements of the data set are not directly displayed. Pattern extraction methods consist of analysis that is dependent on the spatio-temporal representation of the data (Andrienko et al., 2008). Pattern extraction methods serve as the basis of the corridor identification methodology in this framework. In another review of visualization methods Rae (2011) notes that these methods must strike a balance between simplicity and complexity.

2.4.1 Modeling networks as graphs

Much of the early work on network structure addresses this challenge by modeling transportation networks as graphs composed of nodes and edges. Garrison and Marble (1964) defined summary metrics to assess connectivity of transportation networks that can be estimated from a graph. Other studies built on these measures to characterize network shape and levels of circuity (Gordon, 1974), and critically evaluate network structure (Vuchic and Musso, 1991). More recently, Derrible and Kennedy (2010) used concepts from graph theory to develop indicators of the maturity of a network, the relationship of the network to the built environment, and the network directness. They applied the metrics to 33 metro networks and classified them by accessibility and coverage.

Highly complex graphs cannot be visualized with every node and edge depicted. For use in planning, information from the graphs can be summarized through statistics, or the graphs can be simplified. Holten and Van Wijk (2009) developed a sophisticated method to simplify directed graphs intended primarily for visualization. The method models graph edges as springs and assesses the compatibility of edges based on angle and length. Similar edges are grouped to reduce visual clutter, allowing flows to be more legible than if each edge was depicted individually. While this methodology can aggregate a large number of edges, it is effective only for reasonably small number of nodes.

Girardin et al. (2008) used mobile phone and geo-referenced photo data to identify the main points of interest for tourists in Rome, Italy. They then visualized the desire lines connecting the points of interest. Limiting the points of interest makes the visualization possible. Nielsen and Hovgesen (2005) use a grid-based approach to visualize a large OD data set of community trips in Denmark. Each trip is represented as a desire line connecting origin to destination. Then, the map is divided into a grid and desire lines passing through, originating, and ending in each grid square are aggregated to a single statistic representing the overall commuting flow in that square. They applied a similar method to commuter trips in England and Wales (Nielsen and Hovgesen, 2008).

2.4.2 Clustering

Many of the methods that summarize or identify patterns from data rely on clustering techniques, which group data based on feature similarities. These include density-based methods, hierarchical methods, and partitioning methods.

Density-based methods assume that the data consists of a mixture of points drawn from many different probability distributions, with each cluster representing a specific distribution. The algorithms grow initial clusters as long as the density in the neighborhood of the clustered points meets a density threshold (Maimon and Rokach, 2005).

Hierarchical clustering methods either merge or divide data points into clusters using rules about cluster linkage, which define the distance between clusters. Hierarchical methods are flexible and can identify non-isotropic clusters including chain-like or concentric clusters. However, they have non-linear complexity, meaning they may not be appropriate for clustering large data sets (Maimon and Rokach, 2005).

Partitioning methods begin with an initial partitioning and then reassign data points based on feature similarity. They typically require the user to specify the number of clusters. The k-means algorithm is the most used partition method and works well for data that forms isotropic, convex clusters. It is unable to accurately identify non-convex clusters such as concentric circles or clusters that are highly variable in size (Maimon and Rokach, 2005).

Non-spatial clustering of OD data

Yang et al. (2005) developed a method to identify poorly served areas by dividing a network using existing traffic zones, and defining a metric to assess each zone based on existing public transport supply and demand. The metric combines public transport network density and population density (taken as a proxy for demand). and then clustering zones based on this metric. In this case, the clustering is not based on any spatial factors. Instead, the clusters are mapped and performance patterns identified by eye. Similarly, Rao et al. (2012) clustered OD pairs based on demand and identified the highest demand OD pairs, then identified corridors formed by these OD pairs visually.

Spatial clustering of OD data

One of the most common spatial clustering methods is density-based spatially clustering of algorithms with noise (DBSCAN). DBSCAN is based on intuitive definitions of clusters and noise, where clusters are contiguous spatial areas with a consistent density of data and points outside of clusters are noise. DBSCAN defines a neighborhood surrounding each input point, and introduces the concepts of density-reachable and density-connected to separate out clusters from noise. DBSCAN is efficient on large databases and can identify clusters of any shape and size (Ester et al., 1996).

For spatial clustering of OD data, representing data as a single point is insufficient. Instead, methods have been developed based on line segment or trajectory clustering. Nanni and Pedreschi (2006) adapted DBSCAN for spatio-temporal trajectory data by defining a euclidean-based metric for the distance between two objects, consisting of the average distance between the two objects. Zhu and Guo (2014) defined a neighborhood around each origin and destination point and used these neighborhoods to detect the similarity of flows. The flows were clustered using a hierarchical clustering process.

For the methods proposed by Nanni and Pedreschi (2006), Zhu and Guo (2014), and a similar method for detecting "hot flows" developed by Tao and Thill (2016), trajectories must both start and end in approximately the same place to be considered similar. Only trajectories of similar lengths will be clustered. In contrast, Lee et al. (2007) developed a method that is able to cluster similar parts of trajectories, even if other parts of the trajectories diverge. The method decomposes non-linear trajectories into line segments and clusters them based on perpendicular, parallel, and angle distances between the segments, following a method based on DBSCAN.

This methodology was adapted for desire line clustering by Bahbouh and Morency (2014), and then further by Bahbouh et al. (2015). Their methodology preserves the decomposition of trajectories (desire lines) into line segments, but imposes a maximum width and maximum angle requirement to ensure that the clusters identified form a reasonable travel shed. They applied their methodologies to identify corridors with a high concentration of potential walking trips, with the aim of targeting these corridors for walkability-focused improvements.

The research presented here builds on the methodology developed by Bahbouh et al. (2015) to cluster OD pairs into corridors. However, here the objectives are different, as the resulting corridors must be appropriate for bus service, and are prioritized by expected benefits. As discussed in Chapter 6, this requires several adaptations to the methodology in Bahbouh et al. (2015), including minimum and maximum length constraints, the requirement that full desire lines are clustered, and a new prioritization method.

2.5 Summary

Previous work on public transport network design includes partial optimization methods that are focused on full network redesign, and sketch planning methods focused primarily on evaluation. Like many sketch planning methods, the framework presented here is designed for incremental improvements to an existing network. However, it provides additional value by systematically generating bus corridors based on OD-level data.

Many studies make use of new OD-level data to understand travel behavior and evaluate public transport networks, but there are few examples that attempt to use this data for network planning. This research solidifies the connection of data to planning by drawing on aggregation and clustering techniques developed to extract information from data.

Chapter 3

Framework

Bus network sketch planning includes decisions about where to add new bus routes to improve existing public transport networks. As discussed in chapters 1 and 2, while various network planning methods exist, a framework for incremental improvements to a complex, established network is lacking. This thesis presents a data-driven framework to identify corridors where new bus services can provide benefits to existing and potential users of a network. The framework is built on the premise that corridors with concentrations of demand that are not well-served by the current network represent opportunities for new public transport services. This premise is intuitive and not controversial. However, identifying corridors that fit this description can be challenging in large and complex networks.

The framework relies on analysis of zonal origin-destination (OD) pairs, an approach that is useful for understanding network effects in multi-modal networks with multiple paths. The framework includes a methodology to define zonal OD pairs, and an algorithm to cluster OD pairs into corridors that form catchment areas for new bus services. These corridors are the outputs of the framework, and are prioritized by a planner-specified combination of induced demand and benefits to existing passengers. Planners can select from this set of corridors for further study as potential new bus routes.

This chapter begins with a discussion of the scope of the proposed framework, an overview of three main steps, and an introduction to important methodological choices. This is followed by a description of the inputs required and a summary of how the framework addresses the challenges and objectives specified in Chapter 1.

3.1 Scope of framework

The framework consists of a data-driven approach to bus network sketch planning that can be applied to an existing multi-modal public transport network for a variety of planning time horizons.

3.1.1 Existing, multi-modal networks

Chapter 1 discusses the challenges of identifying and systematically evaluating opportunities for new services in a large, complex network where planner observations and intuition fall short. This framework can be applied to an entire metropolitan region, helping planners make strategic decisions about which corridors to prioritize for the addition of new routes. The corridors may cross different areas of the city, and all are evaluated using the same metrics to identify which have the potential to provide the greatest benefits. This yields a systematic generation and evaluation of bus route alternatives, taking into account data from the full network.

In addition to being appropriate for full network level planning, this framework is tailored for application to an existing network. Rather than a complete redesign or network synthesis approach, the objective of the framework is to develop a method for marginal or incremental improvement to an existing system. The framework addresses this objective by identifying potential corridors for new bus services. Corridors are prioritized by expected benefits so that planners can choose how many corridors to develop based on their budget and network improvement objectives.

Multi-modal public transport networks (e.g. including rail, bus, streetcars, etc.) present a unique planning challenge, as discussed in Chapter 1. It is important to account for interactions between modes, including multi-modal journeys. The framework includes all public transport modes and multi-modal public transport journeys in the assessment of existing performance and takes into account the existing modes used and journey stages to select which OD pairs have the potential to be improved. As a result it is appropriate for application in a multi-modal public transport network.

3.1.2 Planning level and strategies

As discussed in Chapter 1, the framework focuses on a specific strategy to improve service in an existing network: the introduction of new routes to reduce circuity and number of stages. Thus, it only considers potential improvements related to network design, ignoring aspects of service quality related to scheduling or operations. For instance, reliability and crowding are not evaluated. Furthermore, the framework is intended for identifying opportunities for new bus routes within an area with public transport coverage. Bus routes that expand coverage to new areas are not considered.

The framework is designed for high level sketch planning and does not include the development of specific route alignments or bus stop locations. In addition, while frequency setting is sometimes considered to be part of route planning, this framework does not address frequency setting. Route alignment, stop location, and frequency-setting can be performed in more detailed downstream analyses. Therefore, the framework exclusively evaluates potential benefits resulting from reduced circuity and number of journey stages. These are a direct consequence of network design at the sketch planning level, and do not depend on downstream decisions about route alignment and stop locations.

The framework identifies a set of promising corridors, so that planners need only conduct more detailed analysis for corridors likely to produce benefits. Each corridor identified by the framework consists of a set of OD pairs expected to be in the catchment area of a potential bus route in the corridor. All corridors identified have expected benefits quantified at the corridor level, which can also be disaggregated to the OD pair level to understand within-corridor variation. The expected demand on each corridor identified as potentially promising is sufficient to support a new bus route.

The framework does not present a methodology to make decisions about the removal of routes. At a local level, the potential to remove routes that provide circuitous service along an identified corridor may be part of the detailed planning that would follow the implementation of this framework. The systematic evaluation of the full network to identify routes for removal would require additional methods and analyses not included in this dissertation, but discussed in Chapter 8 as potential future work.

3.1.3 Mode

The framework is tailored for planning new bus services. Public transport modes have different characteristics and planning requirements. Compared to rail, bus services are flexible, and can operate on most roads. The framework developed here places few restrictions on where corridors can be located, taking advantage of this flexibility. For rail, there are many more construction constraints. Appropriate space for tracks must be procured and different corridors may have very different construction costs. These issues are not considered in this framework and therefore it would require substantial adaptation for rail planning applications. The corridors identified by the framework are required to have size, shape, and demand attributes appropriate for bus service, and the estimated benefits assume new services are provided by bus. The possibilities for adapting the framework for sketch planning of other modes are discussed in Chapter 8.

3.1.4 Time horizon

Strategic planning decisions are made with respect to a specific planning time horizon. Because bus services can be changed at relatively low cost, bus service planning can take place for a variety of time horizons, from short to long. The framework developed here does not assume a specific planning time horizon. Instead it can be used for any time frame for which OD-level expected performance and demand estimates are available.

In this framework, expected travel time savings and expected demand at the OD level are used to estimate benefits and prioritize corridors. Over longer time horizons, traffic conditions may change and the road network may also change, affecting the expected travel times, and making them more difficult to estimate. Estimates of expected travel time influence expected demand estimates. In addition, expected demand depends on other factors, such as land use and population changes, which can vary significantly over time. Over very long time horizons, people may change their home and work locations based on the location of new services. Whatever the desired planning time horizon, demand estimates for the appropriate planning horizon should be used. In short, the framework can be applied with a variety of planning time horizons, as long as reasonable estimates of performance and demand for the time horizon can be made.

Given planner-specified parameters, the framework has the potential to be automated, and relies largely on information that can be collected through automated data sources (See Section 3.4). With its emphasis on incremental network improvements, the framework is appropriate for regular monitoring of an established public transport network.

3.2 Overview

The first step in the framework is the definition of geographic zones. The combinations of these zones form a set of zonal OD pairs. This allows for the analysis of existing and predicted performance and demand at the zonal OD pair level in Step 2. The first component of the OD-level analysis uses mode, stages, and distance, and travel time information to identify OD pairs that can be improved by new bus service. For those OD pairs, the second component of the OD analysis quantifies the potential travel time savings for existing public transport passengers, and the third component quantifies the expected induced demand on the OD pair, given the new service.

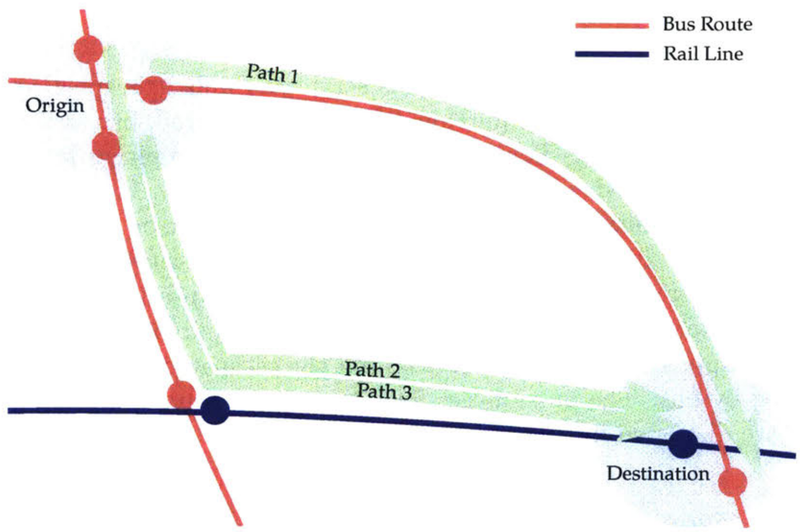

Figure 3-1: Example of multiple paths serving a similar origin-destination pair

Step 3 clusters the expected OD-level demand to identify bus corridor-shaped clusters and prioritizes them by their expected benefits. The results are a set of corridors each consisting of a set of OD pairs, with an expected corridor-level demand that satisfies the demand requirements of a bus route and is within an area that could form a bus route catchment area. The corridors identified are mutually-exclusive and bus service on each corridor identified is expected to provide benefits. Planners should consider the corridors as recommendations for adding service at the full network level. More detailed analysis of these corridors is required for final selection and service planning.

3.2.1 Step 1: Defining OD pairs

In many cases, individuals may consider multiple paths within the public transport network to make their journey. For example, they may have a choice between parallel bus and rail lines, or between neighboring stops on the same route, or they may choose between different paths within the public transport network that start and end in similar locations. Figure 3-1 illustrates an example of three paths that serve the same origin and destination. Path 1 uses a bus route that provides direct service from the origin to the destination. Paths 2 and 3 use the same combination of a bus route and a rail line, but start at different origin stops.

To account for multiple paths, such as the three paths in this example, transportation analysis often groups origin and destination points into origin and destination zones. Aggregating journey data into discrete and mutually exclusive zonal pairs allows for journey-based analysis that assesses the existence of multiple paths in the network and can summarize performance for the entire network without risk of double-counting journeys.

The available paths for each OD pair are used in Step 2 of the framework to differentiate between pairs that are well-served by the existing network structure and those that are not. For the OD-level quantification of benefits, all existing paths are considered. For both assessments, the zonal structure allows multiple paths to be grouped. Zones should be defined such that individuals traveling from an origin to a destination zone would consider all the paths connecting the zones, and would not consider paths originating outside the origin zone or ending outside the destination zone. This aim will never be perfectly achieved due to variation in the distances individuals are willing to travel to access stops and stations and boundary issues for stops and stations that are close to the edge of a zone. Chapter 4 explains how the k-means clustering algorithm, applied to public transport stop and station location data can produce zones of appropriate shape and size for this framework.

3.2.2 Step 2: Analyzing performance and opportunities of OD pairs

In an existing network, not all OD pairs can be improved by new service. Some OD pairs are already well-served by the existing network and/or have no anticipated benefits in terms of travel time as a result of new service. The first part of the OD-level analysis filters out these pairs. Planners may also manually eliminate any OD pairs they know are not good candidates for bus service due to other constraints, such as road network limitations.

Following the filtering step, the potential benefits of new service are quantified for each remaining OD pair. This is a critical step, as the framework identifies corridors prioritized by expected benefits. Corridor level benefits will be aggregated from OD-level benefits and OD-level demand will be used to identify viable corridors for new service.

The benefits of new service are twofold. There are travel time savings accumulated by existing public transport passengers who shift to the new service and there are new passengers attracted to the service from outside the public transport system. In Step 2, both of these benefits are estimated. The expected benefits and current and expected demand will be inputs to the corridor clustering algorithm. The expected demand will be used to determine if potential corridors meet a minimum corridor demand threshold and the benefits are used to prioritize corridors.

Identifying improvable OD pairs

The framework uses specific definitions of well-served and improvable OD pairs. OD pairs are defined as well-served if they are served by direct, single-stage service. In this framework, direct and circuitous are used to refer to how closely a path through the road network resembles the shortest possible path. While some research includes number of stages in the definition of direct (circuitous), in this dissertation the definitions refer exclusively to the characteristics of the path in relation to the shortest path.

Number of stages is also considered in the definition of well-served. Only OD pairs with single-stage service are considered well-served. Single-stage always refers to a journey requiring only a single vehicle. It can refer to one bus stage (one boarding, one alighting) or one rail stage (one entry and one exit without any in-station or out-of-station interchange).

Improvable OD pairs are defined as a subset of the OD pairs that are identified as not well-served. Improvable OD pairs are not well-served and are expected to see improvements in travel time as a result of new service.

As Chapter 1 described, route structure dictates the directness and number of stages required for journeys in the network. OD pairs that are served by direct, single-stage service are well-served by the existing network structure and therefore should be excluded when clustering OD pairs into corridors in Step 3. If an OD pair is already served by direct, single-stage service, a new service would parallel existing service. At best, it would relieve congestion on the existing service, but this could be accomplished more efficiently by increasing service on existing routes. For these OD pairs, there are no expected benefits from the new service. It should be noted that even if an OD pair is structurally well-served, it may not have good performance. However, poor performance on OD pairs that are structurally well-served is caused by factors other than network structure, such as operational issues or insufficient capacity. These performance issues are not of interest for this framework.

A decision tree is used to identify and exclude OD pairs served by direct, single-stage service. In order to classify OD-level service as direct or not, a target distance standard is required, which approximates the distance required for a new route to serve the OD pair. As discussed in Section 3.1, determining the final alignment of bus routes on selected corridors is beyond the scope of this framework. Therefore, precise target bus distances are unknown, but can be estimated based on the shortest path distance. The target bus distance should represent an achievable distance, given constraints that typically dictate bus route alignment, but should still represent good, direct service. Chapter 5 discusses the definition of target bus distance in detail.

The decision tree can be applied to a multi-modal public transport network. In complex multi-modal networks, there are often multiple paths available for many OD pairs. Available paths can be identified based on the physical network. Alternatively, when journey data is available, it can be used to identify the paths used for each OD pair. As long as one of the paths available is direct and single-stage, the OD pair is defined as well-served. Even if poor performance on the direct, single-stage path results in it having little demand, its existence means there is no opportunity for new direct service.

To classify OD pairs as improvable or not, the travel time given new service must be estimated. The expected travel time should reflect the expected path length for the new service, expected congestion, and the time required to make expected stops. Travel time in this framework consists of the sum of in-vehicle and inter-stage time for each journey. This special definition of travel time isolates the travel time components that can be improved through reduced circuity and journey stages. In-vehicle time is influenced by the circuity of the path the vehicle takes and inter-stage time, consisting of the transfer and waiting time between stages, which can be eliminated by replacing multi-stage journeys with single-stage journeys.

New bus services may also impact other components of journey time. If a new route has shorter headways than existing alternatives, it will reduce waiting time for passengers. If the new route has stops closer to passengers' origins or destinations, it can reduce access and egress time. However, waiting time improvements are dependent on the frequency of service selected, and access and egress time improvements depend on the specific route alignment and stop locations. As discussed in Section 3.1, this framework does not inform these downstream decisions. Therefore, current and expected travel times exclude access time, egress time, and the waiting time preceding the first journey stage.

OD pairs that were not identified as well-served but have no expected time savings can result when the current network provides a circuitous but high speed path or if the current network requires multi-stage journeys but coordinates the transfer so that it adds little travel time. If the new service does not provide any expected benefit over existing service, the OD pair is excluded

Estimating potential travel time savings

To estimate the potential travel time savings for existing public transport passengers, the current and expected travel time for the OD pair is required. Some OD pairs will have greater expected travel time savings than others. It follows that existing public transport passengers are more likely to shift to a potential new service if it provides greater time savings. Given this, Chapter 5 describes three methods of varying complexity to estimate total potential travel time savings to existing public transport passengers on an OD pair. The first method assumes that if there are any expected travel time savings, all current passengers will shift to the new service. The second method proportionally assigns current passengers depending on the magnitude of the travel time savings. The third method develops an assignment model for the public transport network and assigns current public transport passengers to the new service based on its anticipated characteristics.

Given the uncertainty in the actual path of the new service, it can be useful to estimate a range of possible travel time savings. The corridor identification algorithm in Step 3 requires benefits for each OD pair to be summarized with a single expected value. However, after clustering is complete, the range of benefits for the corridor can be estimated based on OD-level ranges. This range can show planners which corridors have a narrower range of expected savings and which have a wider range.

Estimating new demand

The expected demand on a potential new service includes journeys that shift from the existing public transport services, as well as new journeys. Current journeys (those shifting from existing public transport services) are estimated using any of the methods described in Section 3.2.2. New journeys include passengers who shift from car or other modes, and latent demand: new journeys that previously were not made. Individuals may choose to make more trips due to improved convenience, and over longer time periods, individuals may change their home and work locations in response to new bus services, generating additional demand.

Expected new demand depends on many factors including the change in bus travel time, the demand and performance of alternative modes, and socio-demographic and land use factors. Predicting demand changes is challenging, both because reliable data on all of these factors is not always available and because of complexity and variation in how these factors influence individual behavior.

Demand estimation is not the main focus of this research. Therefore the framework does not recommend a specific demand estimation method. If attracting new demand is not important, the framework can be applied using only current demand estimates and benefits to current passengers. However, new services are likely to attract new demand, and understanding these potential benefits will be of interest to many planners. The demand estimates should, at a minimum, estimate the impact of the expected travel time changes, and ideally also reflect additional information. Models that use external data such as land use and population should be able to predict new demand even on OD pairs without any current demand. The best estimates available for the planning time horizon desired will produce the most insightful results.

Similar to the expected travel time savings, the corridor identification algorithm requires a single value for the estimated new demand for each OD pair. This will be used to determine if demand on the corridor meets the demand threshold and to prioritize corridors in terms of expected benefits. Once again, even with a single value selected for the clustering algorithm, a range of expected new journeys can be maintained and presented at the corridor level, so that planners can consider the uncertainty of the expected benefits.

3.2.3 Step 3: Clustering OD pairs into prioritized corridors

The aggregation of OD level demand into corridors is a critical component of the framework. New services will not influence OD pairs in isolation. Rather, they will provide benefits along a linear (or pseudo-linear) corridor. Therefore, planners making decisions about adding bus services want to know the expected benefits of potential routes. Step 3 clusters OD pairs into corridors that can be served by a single bus route and have demand characteristics that support new service. For each corridor, expected benefits are defined as the weighted sum of travel time savings for existing public transport passengers and the total journey-minutes of new journeys attracted from outside the public transport network.

Step 3 consists of the development of an algorithm to systematically identify corridors that meet shape and demand definitions, prioritizing those with the greatest expected benefits. Potential algorithms draw on work on trajectory clustering. However, compared to existing trajectory clustering algorithms, the identification of corridors for bus service requires additional constraints on cluster shape and size.

The result is a set of mutually exclusive clusters of OD pairs, called corridors, each of which has sufficient expected demand within a reasonable bus route catchment area to support a new bus route. Each corridor identified is expected to bring some benefits. Out of overlapping candidate corridors, those that generate the greatest expected benefits are prioritized as final corridors.

Some OD pairs may not be assigned to any corridor. These OD pairs are not good candidates for improvement through new bus services because they are not part of a corridor with sufficient demand for new bus service. Depending on the distribution of the OD pairs identified for improvement across space and the demand threshold applied, differing numbers of OD pairs will be assigned to corridors. In networks where a small percentage of OD pairs identified for improvement are assigned to corridors, planners may wish to consider services other than traditional bus routes to improve these OD pairs.

Each corridor identified consists of a set of OD pairs with current and expected demand and expected travel time savings. As described in Section 3.2.2, expected demand and expected travel time savings include uncertainty and may be better represented using a range of values. Thus, the ranges of expected travel time savings and demand can be estimated for each corridor. These ranges can be combined with the results of post-analysis of corridors discussed in Section 3.2.4 to manually fine-tune estimates of expected benefits for the corridors identified.

3.2.4 Post-analysis

While all the corridors identified using the framework will have some potential expected benefits, planners may not wish to add services on all of these corridors. Due to budget constraints, they may only be able to add a limited number of routes, or they may only be willing to add service on a route over a certain threshold of expected benefits. The estimates of expected benefits allow planners to compare the addition of a bus route on a given corridor to assessed benefits of other uses of the same funds, such as additional service on existing routes or other types of service improvements.

For each corridor identified, planners can also perform more detailed analysis of demand in the existing network and potential other sources of demand in the corridor. This information may be used to update the corridor-level benefits estimated within the framework and better understand how flow and benefits are distributed along each corridor.

In post-analysis, planners should also re-consider the feasibility of adding new bus routes on particular corridors, which depends on many factors such as characteristics of roads and sidewalks influencing bus stop location possibilities, geographic barriers, and constraints on termini locations. Section 3.3 discusses how these factors can be considered within the framework. However, some details may come to light with post-analysis of a specific corridor. The post-analysis should define potential route alignment and stop locations, assess demand and flow in more detail, and determine appropriate service frequencies. At this stage, existing routes serving the corridor can also be considered for removal.

3.3 Methodological choices

Planning new bus services in existing networks presents many unique challenges, which can be addressed in a variety of ways. Choices regarding spatial representation and how to account for the existing public transport network are made throughout the framework.

3.3.1 Spatial representation

The framework emphasizes spatial analysis. The spatial representation of travel in a network can range from realistic to abstract. At each step in the network, appropriate choices about spatial representation must be made.

Real networks are incredibly complex. Public transport networks, which are often multimodal are layered on top of a road network. Potential new bus routes are limited to the road network, at a minimum, and likely face several other constraints. Some roads are not appropriate for bus services, either because of space constraints (very minor roads sometimes cannot accommodate large buses and accompanying infrastructure) or speed (it may be unsafe to locate bus stops on very high speed roads). Additional factors often constrain locations that are appropriate for bus route termini. The road network may not allow linear, direct paths between some locations, particularly when geographic barriers limit road network connectivity. Finally, in most networks, where passengers access bus routes on foot, the geographic area that can be served by a new bus route is defined not by the road network that is appropriate for bus service or cars, but rather by the network of walkable paths.

All representations of these complex, layered networks simplify reality, but some mimic it more closely than others. More abstract representations model OD pairs as straight desire lines connecting origins and destinations, and can represent corridors as straight lines surrounded by a fixed buffer. More realistic representations include actual paths through the road network. Distances, which play an important role in all three steps in the framework can be defined as euclidean, road network distance, or walking network distance. Similarly, shortest paths between two points can be defined as a straight line, or as the shortest path through the road or walking network. In some cases, multiple paths are very similar in length. Often referred to as hyperpaths, these sets of paths more realistically represent available shortest paths.

While abstract representations reduce complexity, they may miss important aspects of the actual network. Appropriate spatial representation in the context of the methodologies proposed for defining zones, evaluating circuity of OD pairs, and defining corridors are discussed in chapters 4, 5, and 6.

3.3.2 Planning in an existing network

This framework is designed for incremental improvements to existing networks, which are often complex and multi-modal. In multi-modal networks, different modes have different characteristics, in terms of right-of-way, travel speed, capacity, and flexibility. These differences are accounted for in the definition of zones and the evaluation of circuity. Some OD pairs are served by multiple paths within the existing network, which is also accounted for in the evaluation of circuity and the estimation of expected OD level benefits.

The OD-level evaluation ensures that OD pairs that already have direct service are not identified for improvement within this framework, even if the existing service is poor for other reasons, such as unreliability, low frequency or over-crowding. If an OD pair already has direct single-stage service a new route can only mirror existing service.

Another important aspect of an existing network is demand. Information on demand on different paths in the existing network allows for the estimation of how many journeys will shift to a potential new service. As discussed in Section 3.2.2, this demand shift can be estimated with a variety of methods. Understanding the demand shift allows the estimation of demand for a potential new service and also provides information about the expected effect of this new service on demand in the existing network. Chapter 8 discusses the potential to use this information in an iterative application of the framework proposed here, extending the framework to both recommend corridors for new services and flag routes for removal.

3.4 Inputs

Table 3.1 shows the potential inputs required for each step in the framework. The framework takes advantage of data sources that are readily available for most public transport networks. The most critical inputs can be derived from a combination of smart card or other automated fare card (AFC) data and automated vehicle location (AVL) data. Using existing methodologies, individual journey itineraries can be inferred from AFC and AVL data (Gordon et al., 2013; Chu and Chapleau, 2008).

The framework also includes several parameters and standards, also listed in Table 3.1. Parameter-setting and methods for defining the required standards are discussed in Chapters 4 and 5. Below, several supplementary inputs that can aid in selecting parameters and making estimates are summarized.

Table 3.1: Inputs, parameters, and outputs by step

Step	Substep	Potential inputs	Parameters and standards	Outputs
Step 1: Defining Zones		stop and station locations, demand by stop/station, access (egress) distributions	zone size	zones
Step 2: OD-Level Analysis	identifying improvable OD pairs	existing paths (modes and stages), path lengths for existing bus stages, shortest path distances through the road network, current travel time, shortest path travel time, bus stop spacing, time required per bus stop	target distance standard, expected travel time	set of improvable OD pairs
	estimating travel time savings	current travel time, expected travel time, current demand		potential OD-level travel time savings
	estimating new demand	expected travel time savings, current demand, sociodemographic information, non-public transport demand, land use and population		OD-level estimates of new demand
Step 3: Corridor Clustering		OD pair endpoints, OD-level expected demand, OD-level potential travel time savings, distribution of existing bus route lengths, access (egress) distributions, existing route-level demand, fares and policies	corridor length, maximum distance, maximum angle, minimum flow, coefficients to weight benefits	corridors with expected benefits

In Step 1, the locations of public transport stops and stations are used to define zones. Current public transport travel time, demand, path lengths, and journey mode and stages data are needed in Step 2. For Step 3, the primary inputs are the results of the previous two steps.

All inputs used in Step 2 must be available for the zonal OD pairs defined in Step 1. In the past, modeled estimates of demand and performance for zonal OD pairs were often used for planning purposes. While the general framework presented here could be applied using these estimates, more accurate inputs can be collected by aggregating actual individual journey data to the zonal OD pair level.

AFC and AVL do not generally provide information on access or egress time. However, this framework is focused on high level service decisions, not choices about specific route alignments. The impacts of a new route on access and egress times depend greatly on the specific route alignment, therefore the performance analysis in this framework does not consider access and egress time. Similarly, because individuals do not tap or swipe their fare cards until they board a bus, the waiting time preceding an initial bus stage typically cannot be directly measured from fare card and AVL data (although it can be estimated based on service frequency). However, the waiting time preceding an initial journey stage is not solely impacted by route design choices, but rather depends on the selected service frequency, which is beyond the scope of this framework. In contrast, route design decisions may impact the number of journey stages for passengers in the network by replacing multi-stage journeys with single-stage journeys. Therefore the waiting and interchange time between stages is an important component of performance for this analysis. Helpfully, this time can be inferred from AFC and AVL data.

The framework introduces several parameters that should be tailored to the particular system where the framework is applied. In Step 1, the size of zones should be informed by expected access and egress distances and modes. Step 2 requires estimates of the target bus distance and expected travel time for new services and the expected demand. Shortest path length and time estimates can inform the target distance standard and target travel time. To adjust the shortest path length and time values, information about the road network and current levels of directness, stop spacing, and time required to make stops is needed. As discussed in Section 3.2.2, demand estimates can rely on a variety of other external factors for which data may be available.

The algorithm used in Step 3 to cluster OD pairs into corridors uses several parameters that define the size and shape of the area served by a bus route, including the minimum and maximum length of the corridor, the width of the corridor and the allowable variation of the angle describing the direction of OD pairs assigned to a single corridor. Corridor length may be informed by typical route lengths in the network as well as operational constraints. Corridor width depends on access and egress distances that passengers are willing to travel. The algorithm also takes as an input the minimum demand along the corridor to justify the addition of a bus route. The minimum demand may depend on fare and cost structures as well as policy and planning objectives.

3.5 Summary

Broadly, the framework helps planners develop and maintain a public transport network that meets the needs of existing passengers and attracts new passengers. It does this by identifying corridors with expected benefits in terms of travel time savings to existing passengers and new journeys generated, and quantifying these benefits.

The framework addresses the challenge of planning in an existing network where a full route design overhaul is not desired. Instead of attempting to design a completely new network, the method identifies OD pairs that are not well-served by the existing network structure and clusters these OD pairs into corridors that are promising candidates for new services. This is achieved by differentiating between OD pairs that are improvable and those that are not. This allows planners to select corridors for new services in order to incrementally expand and improve the network.

Network planning in large, complex, multi-modal networks is especially challenging. The proposed framework meets this challenge through OD-level journey-based analysis and systematic aggregation of OD pairs into corridors. First, using zonal OD pairs as the primary unit of analysis allows for the identification of OD pairs that require multi-stage journeys, and the evaluation of performance on these pairs. Analysis at the route or line level cannot produce this same type of analysis. Second, use of actual journey data helps reveal possible paths for each zonal OD pair. Possible paths do not have to be enumerated from the physical network, which may be challenging in complex networks.

Finally, the framework includes an algorithm to cluster OD pairs into corridors, with priority given to the corridors that provide the greatest expected benefits. This helps planners generate alternatives for new services in a systematic way. While the identified corridors must meet the length, width, and demand requirements of a bus route, they may traverse any part of the network. This can identify alternatives that may be missed through more traditional methods. Furthermore, the clustering algorithm allows planners to differentiate between isolated performance issues and those that could be addressed with new bus service. If an OD pair is not contained within a corridor with sufficient demand to support new service, it is not assigned to any corridor. All identified corridors have sufficient expected demand to support a bus route.

An important component of the framework presented here is the flexibility to incorporate planner knowledge and understanding of the public transport network and region and to tailor the process to address planner goals. While the methodology is designed to be applied to a full network, it can also be applied to a portion of the network that is of particular interest to planners. For example, if planners want to expand peripheral bus service, the analysis can be restricted to peripheral zones. Or, if planners know that service cannot be added in certain corridors or regions, they can be excluded from the analysis.

Throughout the framework, parameters can be tailored to specific systems. Flexible parameters include the size of zones defined in Step 1, and the definition of corridor shape and minimum corridor demand applied in Step 3. In the OD-level analysis, there is latitude in the methods for estimating expected distance, travel time, and demand for new services. These influence the number of OD pairs identified as candidates for improvement and the expected benefits for each OD pair.

The outputs of the framework are a set of corridors, each of which are expected to benefit from bus service. The benefits can be summarized as a single number which adds (weighted or unweighted) values for the travel time savings for existing passengers and the expected new demand, or these two benefits can be considered separately. In addition, a range of expected values can be presented for each type of benefit. This helps planners compare the uncertainty of benefits across corridors. Planners can decide which corridors to analyze further. They may also wish to peruse the more detailed break-down of benefits by OD pair. Because each corridor consists of a set of aggregated OD pairs, this more-detailed analysis is straightforward.

In short, the framework accurately assesses opportunities in an existing network. It makes use of data to systematically analyze large, complex networks, and is flexible enough to account for the intricacies of planning in these networks.

Chapter 4

Defining origin-destination pairs

Origin-destination (OD) level analysis is an important component of the overall approach used in this framework. It can assess the ability of the network to serve full journeys, even where multiple stages are required. In addition, because this framework informs planning decisions about adding new service to the network, the analysis should not be limited to existing routes or lines. To estimate expected benefits of new services, the framework incorporates information about shortest path lengths. The shortest path may deviate from the paths on existing services. Therefore it is beneficial to make comparisons at the OD pair level, rather than, for example, comparing car and public transport speeds in the same corridor.

In the first step of the framework, travel in the network is summarized by defining a set of OD pairs. At the most disaggregate level, the origin and destination address for each journey can be used to define OD pairs. However, in large networks this will produce a vast number of OD pairs. More importantly, it does not allow for direct comparison of multiple paths that serve similar, but not identical origins and destinations. As discussed in Chapter 3, understanding the opportunities in a complex network requires consideration of multiple paths that serve similar origins and destinations. When multiple paths are available, individuals may judge the quality of their journey based on the set of paths available to them rather than on a single one. Defining zones enables the downstream OD evaluation and estimation of benefits to account for the diversity of existing service.

If complete journey data, including access and egress stages, is available, true origin and destination points (such as the individual's home or office building) can be grouped into zones, allowing for accurate assignment of complete journeys to zonal OD pairs. In most cases, however, data on the access and egress portion of public transport journeys is much less widely available than within-network journey data, as it is typically more costly and difficult to collect.

Automated fare collection (AFC) and automated vehicle location (AVL) data is available in many networks, and usually provides good coverage of travel throughout the network, but generally does not provide information about true origins and destinations for each journey. When true origins and destination are unknown, it can be assumed that journeys begin (end) in the area surrounding their origin (destination) stop or station. Stops and stations can be grouped into zones with journeys assigned to zonal pairs based on their starting (ending) stop or station. This methodology is reasonable where walking is the main mode to access the public transport system.

This chapter includes some parts heavily drawn from Viggiano et al. (2015)

The definition of a zonal scheme introduces several challenges. First, there is variation in the access distance and egress distance that individuals are willing to accept in their journeys. Therefore, the access area around a stop or station cannot be defined deterministically. Secondly, all zonal schemes present boundary issues. True origin and destination points as well as stops and stations may be located on, or close to, the boundaries of zones. When this is the case, it becomes more likely that the origin (destination) stop or station is not in the same zone as the true origin (destination).

One way to group nearby stops and stations is to use existing zonal schemes, such as postcodes or census tracts. In these schemes, the zones are typically defined using roads as boundaries. Consequently, bus stops and rail stations tend to be at the boundaries of zones, increasing the likelihood that individuals access the stop or station from a neighboring zone.

In networks with dense public transport coverage, it may be impossible to entirely avoid having stops or stations at the boundaries of zones, but zonal schemes can be designed to minimize this issue. Clustering algorithms assign data into groups based on in-group similarity and between-group separation. By defining each stop (or station) by its geographic coordinates (latitude and longitude or equivalent), clustering algorithms identify distinct clusters of stops (and stations) located in close proximity. Defining zones using a clustering method, rather than using an existing zonal scheme, can reduce the incidence of stops or stations on, or close to, zonal boundaries, in turn reducing the probability that stops or stations are accessed from true origins in neighboring zones.

4.1 Clustering methodology

In order to assess the similarity or dissimilarity of data points, all clustering methods use a distance measure to define data closeness. Because data points in this application represent physical points in space, defined by their latitude and longitude, euclidean distance is an obvious choice of distance measure. Of course, the straight line distance between stops or stations is an approximation of the actual walking distance along streets and pedestrian ways. Using distance through the walking network would be more accurate, reflecting the actual distances individuals would need to travel between stops, including when they must travel around barriers. However, this requires more detailed walking network information and assumptions about walking paths.

There are many methods for clustering data based on feature similarities. These include density-based methods, hierarchical methods, and partitioning methods, which were summarized in Chapter 2.

One common density-based method, DBSCAN (density-based spatial clustering of applications with noise) is efficient for large spatial databases, such as the stop (and station) location data in this application (Maimon and Rokach, 2005). However, DBSCAN does not assign all data points to clusters. Clusters are only identified where the minimum density threshold is met, with points outside these clusters left unassigned. This aspect of the methodology is problematic for the application presented here, as all stops and stations must be assigned to a zone for the downstream analysis.

Hierarchical algorithms output a dendogram, which presents nested clusters representing different levels of similarity. While hierarchical clustering does not require the user to predetermine the number of clusters, the user must determine which level of the dendogram to use to define clusters (Maimon and Rokach, 2005). In contrast, partitioning methods typically require the user to specify the number of clusters.

While both hierarchical and partitioning methods may be able to produce reasonable clusters from the stop and station location data, the k-means algorithm (a partitioning method) was selected for several reasons. The k-means algorithm assigns points to clusters based on the distance to the cluster centroid, producing voronoi polygon-shaped zones surrounding each cluster centroid. This method matches the aim of the application: the spatial clustering of coordinates, in which ideal zones consist of a cluster of stops (and stations) close to the zonal centroid, with a buffer surrounding them. In addition, applying k-means is computationally efficient compared to hierarchical methods (Maimon and Rokach, 2005).

Furthermore, the two most commonly cited weaknesses of k-means, the fact that it is limited to isotropic clusters, and the requirement that users specify the number of clusters, do not pose problems for this application. Non-uniform, concentric, and irregular shaped clusters are neither expected nor desired. And while the optimal number of clusters is unknown, the approximate number of clusters can be estimated given information about access and egress distances. Section 4.2 discusses the selection of the number of zones in more detail. Because zone size is an important consideration in this application, allowing the user to specify the number of clusters provides more control over this aspect, compared to some other methods.

The k-means algorithm consists of three steps. First, a set of data points are selected as the initial centroids. The number of centroids corresponds to the user-specified number of clusters. Then, each stop or station is assigned to the closest centroid, using the euclidean distance between the coordinates. Once all points have been assigned, the centroids are recalculated as the mean coordinate values of all points assigned to a given cluster. This process iterates until the centroid locations do not change significantly from one iteration to the next (Lloyd, 1982).

The algorithm always converges, but may reach a local (instead of a global) minimum, based on the selection of the initial centroids. Therefore, the algorithm is typically run multiple times to find the result with minimum in-cluster variation (MacQueen, 1967). In addition, in order to improve performance, the k-means++ initialization process can be used to select the initial centroids, ensuring that they are distributed throughout the data set. To achieve this, initial centroids are sequentially sampled from a weighted version of the data set, with weights assigned proportional to the squared distance from the closest initial centroid already selected. The process is fast and improves the performance of the k-means clustering methodology (Arthur and Vassilvitskii, 2007).

The k-means algorithm clusters data points, assigning equal importance to each point. However, rail stations and bus stops have different properties. Typically, networks have fewer rail stations than bus stops, but each rail station draws much higher ridership. Positioning high ridership rail stations in the center of zones can reduce violations of the assumption that journeys' true origins (destinations) are in the same zone as origin (destination) stops or stations. Therefore, it can be advantageous to weight stops and stations by number of boardings prior to applying the k-means algorithm. The impact of stop and station weighting is discussed for the London application in Chapter 7. A simpler option of weighting rail stations by a factor over bus stops, is also discussed, and shown to be more effective in this application.

4.2 Number of zones

The number of clusters is a user-specified input to the k-means algorithm. For this particular application, the size of zones is important. If there are too many zones, individuals are likely to consider stops and stations in multiple zones, and the stops and stations in a single cluster will attract passengers from multiple zones. At the same time, the stops within each cluster should be in comfortable walking distance of one another, as the analysis assumes that paths originating from any stop or station in a zone constitute valid alternatives for travel on zonal pairs. Too few clusters can result in stop and station groupings that violate this assumption.

Given these objectives, information about access and egress distance distributions in a network is important in choosing the appropriate number of zones, and resulting zone size. Acceptable walking distances may vary by city based on road characteristics and walking infrastructure or cultural attitudes. Information about the distance individuals walk to access the existing system ensures that the zone size is appropriate. Additionally, data on complete journeys, including the access and egress portions (if available) can be used to validate the chosen zonal scheme, by assessing the extent to which journeys' true origins (destinations) are in the same zone as their origin (destination) stop or station. The use of access distance and complete journey data in the zone definition process is demonstrated for the London application in Chapter 7.

In addition to information on walkability, clustering evaluation metrics can be used to help determine the appropriate number of zones. Various metrics have been developed to evaluate the quality of clustering results, which allow for the comparison of results for different numbers of clusters. One metric is the silhouette score, which measures both cluster tightness (distances between points in the same cluster) and cluster separation (distances between clusters). Another metric that is sometimes applied is the sum of squared errors, which sums the squared distances between all points in a cluster and the cluster centroid. The sum of squared errors always decreases as the number of clusters increases, but a distinct "elbow", where the reduction in the sum of squared errors begins to taper off and the slope of the sum of squared errors plot changes from steep to shallow, signals the appropriate number of clusters. Chapter 7 discusses the efficacy of such metrics for selecting the number of zones in the application of the framework to London's network.

4.3 Step 1 outputs

The results of the stop and station clustering are cluster assignments for each stop and station. The centroid of each cluster is the weighted average location of all stops or stations assigned to the cluster. The centroids are used to define a voronoi diagram, in which the area served by the public transport network is partitioned into a set of cells, one for each centroid. Each cell represents the area for which its centroid is closer than any other. The boundaries of these cells define the set of zones used throughout the framework. All combinations of zones are zonal OD pairs. For the OD-level analysis in Step 2, current journeys are assigned to the zonal OD pairs based on their origin and destination stops and stations.

Chapter 5

Origin-destination level analysis

The objective of the framework is to identify opportunities to add service to improve poorly served OD pairs. Step 1 defined a set of OD pairs. Step 2 consists of analysis of these pairs. Improvable pairs are identified and the expected benefits of new service for these OD pairs are estimated. Subsequently, Step 3 of the framework (described in Chapter 6) clusters improvable OD pairs into corridors to identify those corridors where new services can provide the greatest benefits.

5.1 Analysis constraints

The full set of OD pairs consists of every combination of the zones defined in Step 1. However, there are some limitations on which of these OD pairs can be analyzed in Step 2. First, some journeys are very short, starting and ending within the same zone. These journeys are assigned to OD pairs where the origin and destination zone are the same. The analysis in this framework summarizes current services and performance and expected benefits at the zonal OD pair level. Analysis of journeys that start and end within the same zone depends on where they originate and end within a zone. In the OD analysis, expected travel time and target distance estimates assume travel begins and ends at the zonal centroid, which precludes estimating these values for intrazonal OD pairs. Therefore, OD pairs with the same origin and destination zone are excluded.

The OD analysis uses current public transport travel times estimated based on existing public transport journeys. However, some OD pairs may have few or no existing journeys in the period analyzed. To ensure travel time estimates are reliable, a minimum number of journeys constraint is imposed. The minimum sample size for an estimate is:

$$n = \frac{z^2_{\frac{\alpha}{2}}\sigma^2}{E^2} \tag{5.1}$$

where $z^2_{\frac{\alpha}{2}}$ is the critical value for confidence level α. For a 95% confidence interval, the critical value is 1.96, which is applied here. σ is the population standard deviation for travel times for each OD pair. E is the allowable error. For OD pairs that do not meet the sample size constraint, it is not possible to reliably estimate current travel time. Therefore, travel time on these OD pairs cannot be compared to the expected travel time with new service to determine if the OD pair can be improved by new service.

Eliminating these OD pairs may remove OD pairs that have little demand because public transport service is poor, including cases where service is poor due to circuity. Therefore, there may be opportunities to improve these OD pairs. One method to account for these opportunities is to estimate travel time for these OD pairs based on the network, rather than from journey times. This requires data on link travel times in the public transport network. Here, these OD pairs are excluded, but this work around should be considered for future implementations.

Demand and travel times vary by time of day. Therefore, expected benefits of new services will depend on the time period analyzed. The framework should be applied to journey data for each time period of the day (such as AM peak, inter-peak, PM peak) separately. Planners can ultimately make decisions based on the corridors identified for each period, according to their service priorities.

Planners may also prefer not to change some parts of the network or there may be constraints that prohibit the introduction of new bus services in some areas. If OD pairs are not candidates for new service, they can be excluded at this stage.

5.2 Filtering OD pairs

The first part of the OD-level analysis filters OD pairs to identify the set of OD pairs that can be improved by new bus service. To do this, OD pairs are first classified as well-served or not well-served. While good public transport service encompasses many factors including travel time, reliability, service frequency, crowding levels and others, here a special definition of well-served is used, focused only on network design factors. OD pairs are defined as well-served if they are served by direct single-stage service. Recall that in this framework direct is defined as the opposite of circuitous. Direct service uses a path that closely approximates the shortest path.

Of those OD pairs that are not well-served, a smaller subset are defined as improvable. Here, improvable means that an OD pair is not well-served and the expected travel time given new service is shorter than the current travel time. OD pairs may be well-served but not improvable if they are served by bus routes that are indirect but fast, or by coordinated (and therefore fast) multi-stage service. Determining if an OD pair is improvable requires the definition of current travel time and the estimation of expected travel time given new services (see Section 5.2.2).

5.2.1 Identifying OD pairs that are not well-served

Well-served OD pairs are filtered using the decision tree in Figure 5-1. The first node in the tree evaluates whether the pair is served by single-stage rail service. Single-stage rail service is defined here as travel on a single line. OD pairs requiring in-station transfers are not defined as single-stage. OD pairs are defined as served by single-stage rail if they have at least 1 journey by single-stage rail. For OD pairs served by single-stage rail there is no additional filter to evaluate directness. Rail service is not limited to the road network. Therefore, comparing rail paths to potential paths through the road network does not make sense. Rail service is also inflexible. This framework assumes that rail networks are well-planned, providing the most direct path possible given constraints. If an OD pair requires multi-stage rail, however, it is evaluated further.

OD pairs that are not served by single-stage rail are filtered based on whether or not they are served by single-stage bus. As with single-stage rail, OD pairs are defined as served by single-stage bus if at least one journey is made using single-stage bus. If the OD pair is not served by single-stage bus, this means that the OD pair is served only by multi-stage

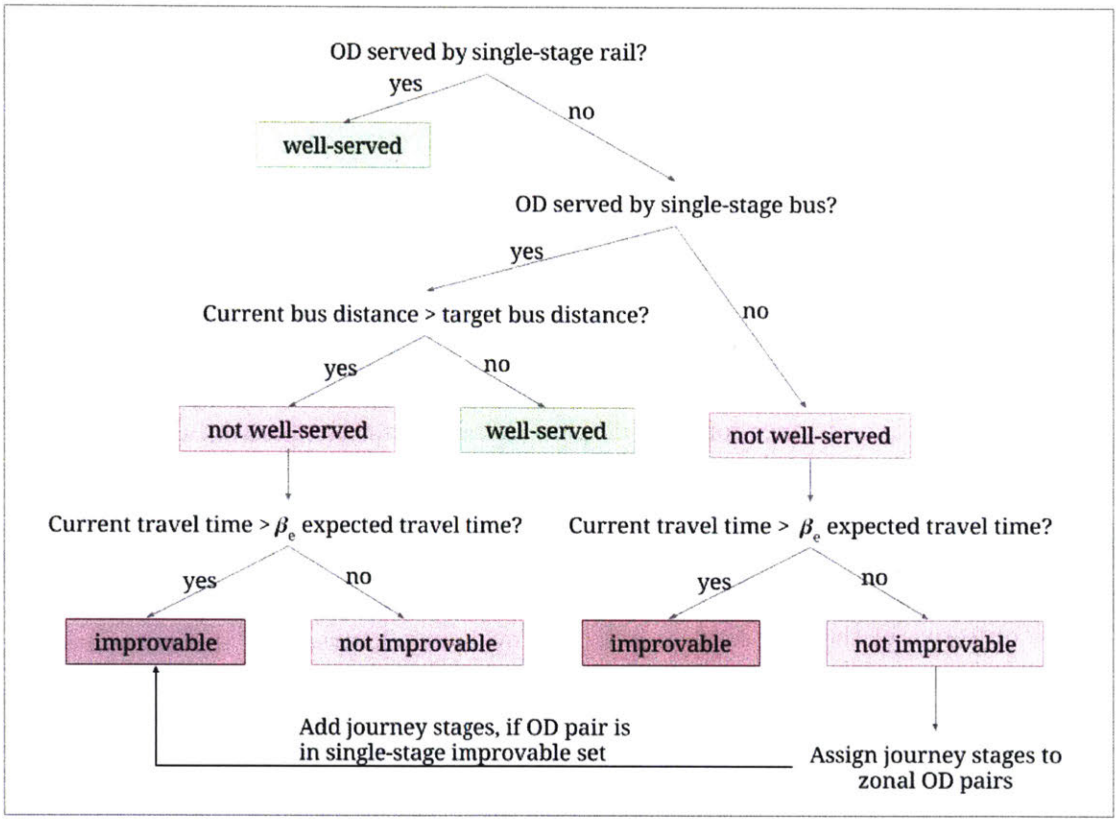

Figure 5-1: Decision tree

paths. These OD pairs have the potential to be improved by providing a single-stage bus option. Therefore these pairs are included in the set of OD pairs that are not well-served.

If an OD pair is served by single-stage bus, adding a new bus service will not reduce the number of journey stages for the OD pair. However, the existing bus routes may take circuitous paths. If this is the case, a new bus route that follows a more direct path could improve service. To evaluate this potential, the next node in the decision tree compares the current best-available bus distance to the target bus distance for a potential new bus route. If the current best-available distance exceeds the target bus distance, the OD pair is included in the set of OD pairs that can be improved. Otherwise the OD pair is excluded, as it is already served by direct single-stage bus. The set of OD pairs remaining after filtering out OD pairs that are well-served by the current network structure are pairs that are served by circuitous single-stage bus (possibly in addition to multi-stage alternatives) and pairs that are not served by any single-stage service.

Defining current bus distance

The current best-available bus distance can be determined from the physical network or may be available from journey data. If an OD pair is served by multiple bus routes, the distance on the bus route that takes the most direct path should be selected. As such, any improvement for the OD pairs will have to be better than the best (shortest) path already available. In many cases, a zone contains multiple bus stops on the same route. While the shortest distance will be found from the stop closest to the edge of the zone, the bus distance should be taken from the stop on the route that is closest to the zone centroid. This provides a more representative distance for the OD pair and allows for comparison with the target bus distance, which will also be estimated from zone centroid to zone centroid.

Defining the target bus distance

The exact path of the potential new service is not specified in early stage bus planning. Given that the objective of the new services is to improve network structure, the expectation is that new bus routes will follow a direct path through the road network to serve a given OD pair.

As discussed in Chapter 3, different methods for spatial representation can be used in the context of this framework. The shortest path serving an OD pair can be approximated as a straight line path, with the length measured as the euclidean distance from zone centroid to zone centroid. However, this ignores the road network and any geographic barriers that cannot be traversed by bus. Therefore, it is superior to use a representation of the road network, or ideally, a subset of the road network that is traversable by bus. The shortest path distance through this road network represents a lower bound on the bus distance for a new service on an OD pair.

However, the potential route serving the corridor may not follow the shortest path through the road network exactly. Planners may decide that it is important to place a bus stop in front of a location such as a school or hospital or may be required to avoid a stretch of road that is not compatible with bus operations or cannot support bus infrastructure. These factors may vary within and between networks. For example, a newer city with wide roads and ample space for bus stops may have more flexibility when selecting route alignments than an older city with narrower roads and more restrictions. A target bus distance closer to the shortest path distance can be used for the newer city compared to the older city. If the road network varies considerably in different parts of the city, definitions of target bus distance may also vary. Data on how current bus distances relate to the shortest path distances for OD pairs in the network can inform the definition of target bus distance.

In general, the definition of target bus distance (d_{target}) as a function of the shortest path distance through the road network (d_{short}) takes the form specified in Equation 5.2. m and a are constants selected based on analysis of current services or planner objectives or specifications. m reflects additional distance that is proportional to the shortest path distance for the OD pair. a reflects fixed additional distance that is independent of shortest path distance.

$$d_{target} = md_{short} + a \tag{5.2}$$

Depending on the choice of m and a, more or fewer OD pairs are identified as well-served. More conservative choices will ensure that the OD pairs that are not well-served can be improved by adding a more direct bus route. However, very conservative choices of m and a may eliminate some OD pairs from the downstream analysis that have some possibility for improvement. Ultimately, the target bus distance serves as a planning standard. Planners should set an objective for directness in the network, which can be defined in terms of m and a.

5.2.2 Identifying OD pairs that are improvable

Once OD pairs that are already well-served by the existing network have been removed, a secondary filtering step removes OD pairs that, while not well-served in terms of number of stages and directness, nonetheless are not expected to see travel time improvements from new services. As shown in Figure 5-1, based on current and expected travel time, OD pairs are identified as improvable or not improvable.

Defining current travel time

The current travel time includes in-vehicle time and interchange time between stages. As discussed in Chapter 3, while new services may impact access, egress, and waiting time, the extent of these impacts is highly dependent on route alignment details and service frequency. The framework proposed here is designed for sketch planning and leaves route alignment and frequency choices for later analyses. The current travel time therefore accounts only for the travel time components that are directly influenced by network design - this includes in-vehicle time, which is determined in part by the circuity of the path taken, and interchange time, which is eliminated when multi-stage paths are replaced with single-stage ones.

For each OD pair, the travel time for current passengers will vary by origin (destination) stop or station and by path selected. In addition, there are other sources of variation, including operational decisions, minor incidents, road congestion, and crowding and demand fluctuations within the public transport system. The current travel time can be summarized as either the mean or the median of the travel times for current journeys. Using the median is preferable as it is less sensitive to long travel times that may be the results of incidents or very inefficient path choices. Passengers who currently choose very inefficient paths relative to the paths available may have a specific reason for doing so and would not necessarily benefit from a new service.

Estimating expected travel time

Similar to the expected distance, the expected travel time will depend on the specific route alignment, prohibiting a precise determination. However, an estimate of the expected in-vehicle travel time can be made based on the shortest path travel time, the expected frequency of bus stops, the approximate time required to make a stop, and the expected in-vehicle distance. This estimate is specified in equation 5.3.

$$t_{exp} = m' \left(t_{short} + \frac{d_{short} t_{stop}}{d_{interstop}} \right) \tag{5.3}$$

Where t_{exp} is the expected travel time. t_{short} is the travel time on the shortest path and d_{short} is the distance on the shortest path. m' is a multiplier to adjust for factors that may cause the travel time to be longer than the shortest path travel time, including deviations from the shortest path described in Section 5.2.1. t_{stop} is the expected stopping time for buses, including deceleration and acceleration. $d_{interstop}$ is the distance between stops.

The time required to make a stop, t_{stop}, is highly variable, depending on boarding and fare payment technology, number of boardings and alightings, the location of the stop along a road, and the speed of traffic on the road. In many networks, buses may skip stops when there are no passengers boarding or alighting. As a result, estimating t_{stop} is challenging. Existing analysis of dwell time and the time to accelerate and decelerate to make a stop can inform the choice, as well as information about expected demand (hence boardings and alightings) on new services. The distance between stops, $d_{interstop}$, is specified as a service planning objective or requirement in many networks.

A single value for expected travel time is used to filter OD pairs that are not improvable. An optional coefficient, β_e, can be used to scale up the expected travel time for filtering purposes, such that only OD pairs with more significant expected travel time improvements are selected.

Partially improvable OD pairs

The framework evaluates complete journeys, which is necessary in order to identify cases where OD pairs requiring multi-stage service can be improved through the provision of single-stage service. However, there may also be cases in which one stage of a multi-stage journey can be improved. The potential benefits for these passengers should also be included in the OD-level estimates. The following procedure accounts for journey stages on these partially-improvable OD pairs.

All journeys on OD pairs that require multi-stage paths but are not identified as improvable (based on travel times) are split into their journey stages. Each bus stage is assigned to the zonal OD pairs defined in Step 1 based on the location of its starting and ending bus stops. If a bus stage is assigned to an OD pair that was previously identified as a candidate for improvement (based on definitions of well-served and the expected travel time), the stage level demand is added to the current demand for that OD pair.

An important distinction is that journeys on OD pairs requiring multi-stage paths that are identified as improvable are not split into journey stages. This would result in double-counting, as these stages are already included as part of the journeys on an improvable OD pair. Improving complete journeys is prioritized over improving journey stages. Chapter 6 discusses how stage-level demand on OD pairs that are not assigned to any corridor can be accounted for in post-analysis of potential corridors.

5.3 Estimating potential travel time savings

The potential travel time savings for current public transport passengers for each OD pair consist of the difference between the current travel time on the OD pair and the expected travel time for the OD pair given new bus service. As discussed in Section 5.2.2, current travel time is estimated as the median time for current journeys on the OD pair. The total potential travel time savings can be estimated as the travel time savings multiplied by the number of current passengers expected to shift to the new service. In reality, current passengers may take a variety of different paths in the existing network and experience different travel times. Thus, travel time savings could be estimated at the individual level and summed. However, using the median travel time for all journeys provides a reasonable estimate.

5.3.1 Estimating current public transport passengers who benefit from new service

The simplest approach to estimating total potential travel time savings at the OD pair level is to assume all passengers shift to the new service and that the new service has the capacity to accommodate them. Barring capacity constraints, all existing passengers have the option to use the new service. In this sense, it is reasonable to apply the benefit to all current public transport passengers.

In reality, however, not all current passengers are likely to use a new service. Some current passengers may elect to continue using the existing services instead. If a new service is introduced, the number of current public transport journeys that will shift to the new service is likely to depend on the magnitude of the improvement that the new service produces relative to existing service. At the same time, some passengers on an OD pair that was not classified as improvable, may shift to a new service, if it serves the OD pair.

To improve on assignment based only on binary classification, one possibility is to estimate current journeys that will use a new service using percentages defined heuristically in proportion to potential travel time savings. For OD pairs with large travel time savings, a large percentage of current journeys would be expected to use the new service. This percentage would decline as the expected improvement decreases. The percentage scale would need to be defined based on data from past improvements or from models. This assumes that travel time savings are the only factor influencing passengers to shift to new services. Path choices likely depend on many factors besides travel time, including the cost of each alternative, the purpose of the trip, and passenger income and age. In addition, other travel time components not included when estimating travel time savings, including access and egress time and waiting time likely influence path choice.

A more sophisticated approach could use a within-public-transport demand model to assign journeys to the potential new service based on the expected travel time on the proposed new service and travel time on the existing public transport options in addition to some of these additional factors. A discrete choice model could account for stochasticity in passengers' route choices. This would produce better estimates of current journeys expected to use a new service at the OD pair level.

However, estimating such a model that can be applied at the zonal OD pair level is challenging, and data on the many factors that influence decisions may not be available. In addition, models may be highly location specific, meaning that a model from one location may not be applicable in another location.

Given this uncertainty, a range of estimates of current passengers who will shift to (and hence benefit from) the new service can be produced. This could be combined with a range of potential travel time savings, produced by varying the parameters that define the expected travel time. While a single value for potential travel time savings is used to prioritize corridors within the corridor identification algorithm in Step 3, the range of potential benefits for each corridor can be estimated to provide an idea of the uncertainty of the expected benefits.

In addition to expected travel time savings, the estimate of current passengers who will shift to the potential new service will be used as an input to the corridor identification algorithm in Step 3. This information is used for estimates of expected flow on potential corridors.

5.4 Estimating expected new demand

For OD pairs that are improvable, the potential reduction in travel time is expected to attract new demand, in addition to existing public transport journeys on the OD pair. A variety of methods exist for demand estimation. The following sections discuss three demand estimation methods and the arguments for and against each.

5.4.1 Four-step demand models

Demand at the regional scale is often estimated using a four-step model, consisting of: trip generation, trip distribution, mode choice, and trip assignment (Willumsen and Ortuzar, 2011). The trip generation step estimates the number of journeys leaving from, and attracted to, each zone based on population and land use characteristics. The trip distribution step links the trips produced by, and attracted to, each zone, creating a zone-to-zone OD matrix. This is often achieved with a gravity model. In the mode choice step, the OD level demand is assigned to the available modes. This can be done using a variety of aggregate or disaggregate mode choice models. In this step, all public transport options are often treated as a single mode. The final step completes within-mode path assignment. This assignment is usually based primarily on travel time although other factors such as cost and comfort can also be included (Willumsen and Ortuzar, 2011).

Models of this type have been criticized as non-behavioral and not very reactive to system changes, as the first two steps do not account for many performance-related variables that may influence demand (Ceder, 2007). Predictions at the zonal OD pair level may not be reliable. On the other hand, these models can be effective for long range predictions, as expected population and land use changes (including those that result from new services) can be included in this first step of the model.

5.4.2 Direct demand models

Sketch planning analysis often uses direct demand models. These models predict total trips (often for each mode and purpose) from an origin zone to a destination zone based on a set of explanatory variables, including population and employment. Direct demand models estimate demand in a single step, rather than using multi-step distribution and assignment models. The most well-known examples are for intercity travel, including the Kraft-SARC model for the Northeast Corridor (Domencich and McFadden, 1975). There are few applications to urban areas and their successful implementation has not been demonstrated for small zone sizes (Willumsen and Ortuzar, 2011). Furthermore, direct demand models cannot easily account for demand and performance of alternative modes (Willumsen and Ortuzar, 2011). If these weaknesses can be overcome, direct demand models have the benefit of being straightforward to apply in the context of this framework.

5.4.3 Elasticities

The simplest demand estimation methods apply elasticities to predict expected demand changes based on performance or cost changes. There is a significant body of research dedicated to estimating demand elasticities. In this framework, it is important to estimate the demand response to changes in travel time. Some research estimates elasticity with respect to generalized cost which includes weighted travel time components as well as travel cost (Halcrow Fox and Associates et al., 1993). Other examples estimate elasticities for each component of public transport travel time (Small and Winston, 1999). Trip length, trip purpose, time of day, and income are all sources of variation in estimated elasticities (Balcombe et al., 2004).

For this framework, applying an in-vehicle time elasticity based on the change in in-vehicle time expected to the current demand levels represents the simplest way to predict new demand. This approach requires only the inputs already required by the methodology. However, these estimates do not utilize any information about competing modes and expected population and land use changes that may be available, and therefore cannot account for the effects of these factors. While elasticity estimates may account for some induced demand, they are unable to estimate new demand on OD pairs that currently do not have demand, and are likely to underestimate demand changes on OD pairs with very few existing journeys. For these OD pairs, the lack of existing journeys may reflect very poor current service. There may, however, be significant latent demand, which elasticity models alone will be unable to predict. In contrast, direct demand models and four-step style models can predict induced demand on these types of OD pairs.

5.4.4 Uncertainty

These estimates are uncertain. One way to recognize this is to estimate a range of expected new journeys based on known uncertainty in demand model parameters. Estimates of new demand are used in Step 3 to estimate expected flow on potential corridors and to prioritize corridors based on expected benefits, which include new journey-minutes. While the corridor identification methodology uses a single value for each OD pair for each of these inputs, a range of expected new journey-minutes for each potential corridor can be estimated and used to compare the uncertainty of benefits on different corridors.

5.5 Step 2 outputs

The results of Step 2 of the framework are a set of OD pairs that are deemed improvable. This means they are not currently served by single-stage rail or direct single-stage bus service and the expected travel time given new bus services is better than the current travel time. For this set of OD pairs, the potential travel time savings have also been estimated, as well as the current demand and the expected demand given new service.

Chapter 6

Corridor identification methodology

The origin-destination (OD) level analysis in Step 2 of the framework may identify thousands of OD pairs that could be improved by new, direct service. Adding service for all of them is unreasonable given most budgets. If improvable OD pairs are clustered along a corridor, many can be served by a single new bus route. These cases represent the best opportunities for new services.

Step 3 of the framework clusters improvable OD pairs into corridors, which are clusters of OD pairs. The expected benefits, estimated at the OD level in Step 2, are aggregated up to the corridor level. The corridors must have appropriate characteristics to be served by a new bus route, including shape, size and demand. There may be restrictions on route termini, geographic barriers, and intricacies of the road network that impact the suitability of corridors for new bus routes. As discussed in Chapter 3, these spatial factors can be accounted for in different ways. Section 6.3 describes how these factors influence methodological choices. Post-analysis of the corridors identified can further assess spatial factors and also evaluate the potential impacts of new bus routes on existing routes.

6.1 Review of trajectory clustering

Chapter 2 discussed spatial clustering, including trajectory clustering. There are existing methods for clustering both curved and straight line trajectories (Hung et al., 2015; Lee et al., 2007; Bahbouh et al., 2015). Each method incorporates a definition of similarity of trajectories that is used to produce the desired cluster type. Some methods cluster the full trajectory and assess similarity in terms of cluster starting and ending points. This is generally intended for visualization of a large number of trajectories. Other methods, such as TraClus, the trajectory clustering framework developed by Lee et al. (2007) split trajectories into segments and identify corridors by clustering segments. Therefore, the corridors identified may include part or all of a trajectory. This method is applicable where high density corridors are of interest, and origin and destinations are secondary.

Bahbouh et al. (2015) adapted the framework for clustering desire lines, developing TraClus-DL. The methodology is similar, but they defined more intuitive parameters, with physical meaning. They applied the method to identify corridors with a high potential for walking trips. The following section summarizes their methodology, and discusses the aspects that must be further adapted to identify corridors for bus services.

6.1.1 TraClus-DL

Bahbouh et al.'s method preserves the segmentation of desire lines into segments and clusters segments into corridors defined by three parameters: maximum distance, maximum angle, and minimum weight (journeys). In addition, the user specifies the length of segments that desire lines are divided into. In their algorithm, each segment is considered as a seed. The distance between the seed segment and another segment is defined as the minimum distance from the midpoint of the seed segment to any point on the other segment. The maximum distance thus defines a radius around the midpoint of the seed segment. The maximum angle limits the difference between the angle of the seed OD pair and any other OD pair in its cluster. Segments of desire lines that are within the maximum angle of the seed segment and within the maximum distance of the seed segment are added to the seed segment's cluster. Each desire line contributes at most one (the closest) segment to each cluster.

If the demand (number of journeys) on the desire lines contributing segments to the cluster meets the minimum weight parameter, it is expanded. The expansion consists of defining the same circle around each non-seed segment in the cluster and adding up to one segment from each new desire line in these circles, if the demand on the desire lines meets the minimum weight specified. (The angle similarity requirement remains and continues to be assessed with respect to the original seed.)

All expanded clusters are placed in a priority queue, ordered by weight (number of journeys). The top cluster in the queue is popped and defined as a final corridor segment. At this time, any segments that are in this final corridor are removed from any other clusters of which they are part. Clusters that no longer meet the minimum weight requirement are removed from the queue.

The method produces corridor segments, as each desire line contributes at most one segment to each cluster. They summarize each corridor segment with a representative trajectory, averaging the segments that are members of the cluster and suggest these representative trajectories be smoothed and connected in post-analysis.

Several aspects of TraClus-DL present problems if it is used for the identification of potential bus corridors. First, the decomposition of desire lines into segments means that for some OD pairs, part of the desire line will be assigned to a corridor, and some portion will not. The objective of this framework is to identify opportunities where new bus routes can reduce circuity. If new service is added for only part of the desire line, it could bring some benefits, but multi-stage journeys would still be required. Moreover, if the new service does not connect with other services to serve the origin and destination, the new service would not provide any benefits at all. A method that clusters full trajectories is needed.

Second, there are distinct spatial features that define a good bus corridor, including length and width. In TraClus-DL, the cluster expansion process produces corridors of varying widths. This expansion process is a feature of the Density-Based Scan algorithm on which TraClus and TraClus-DL are based. The objective is to continue expanding clusters as long as the density persists. In TraClus-DL, clusters are only expanded once, which places some limits on corridor width. However, some corridor segments may have a width of up to four times the maximum distance, while others may be much narrower.

TraClus-DL also does not incorporate any constraints on corridor length. Corridor segments are short, and the smoothing and connecting applied after-the-fact may produce corridors of any length. Bus routes that are too short or too long raise operational and planning concerns.

Compared to TraClus-DL, the methodology for identifying bus corridors must incorporate more constraints on corridor shape. However, some elements from TraClus-DL remain applicable. The similarity measures, including angle and distance are appropriate. The incorporation of a density requirement is also critical. In this case, it ensures that the corridor has sufficient flow to support a bus route. The use of a priority queue to eliminate overlap and ensure corridors with the greatest benefits are selected, is also applicable here.

6.2 Methodology overview

The proposed methodology for identifying corridors defines an exhaustive set of potential corridors that have appropriate spatial characteristics and meet constraints and demand requirements for bus services. Each potential corridor consists of a cluster of OD pairs; OD pairs may be assigned to multiple potential corridors. Like in TraClus-DL, the second part of the methodology eliminates this overlap by establishing final corridors in order of priority. Here, priority is defined by the expected corridor-level benefits.

Figure 6-1 summarizes the first part of the methodology. The objective of the first part of the methodology is to define an exhaustive set of candidate corridors. One way to address this objective is to begin by defining a set of potential routes. These routes do not have to be defined precisely. They may consist of an abstract representation of a potential route, such as a straight line. Figure 6-2 shows examples of possible spatial representations of potential routes. The potential routes are generated from a set of anchors and are defined according to rules about length and shape. The length, shape, and anchor points from which each potential route can be generated are discussed in Section 6.3.1.

Next, OD pairs are assigned to each potential route, based on similarity criteria. Similarity criteria can be defined and assessed in different ways, as will be discussed in Section 6.3.2. While length rules are already applied to the potential routes, OD pairs may not exist along the full length of the potential route. Therefore, the length of the corridor is evaluated based on the OD pairs assigned to it and compared to a minimum length parameter that reflects the minimum reasonable length for a bus route. Then, the expected flow on potential corridors (which consist of the set of OD pairs assigned to each potential route) is compared to a minimum flow parameter. Only corridors that meet the minimum flow requirement are placed in a queue, prioritized by the expected corridor-level benefits.

The second part of the methodology defines final corridors using a priority queue, similar to the final stage of the TraClus-DL algorithm. This process is summarized in Figure 6-3. The highest priority corridor is established as a final corridor and any OD pairs that were assigned to this final corridor are removed from any other corridors that they were assigned to. After OD pairs are removed, any potential corridors that no longer meet the minimum flow requirement are removed from the priority queue, and the priorities of all potential corridors are re-assessed, re-ordering the corridors in the queue as needed. This process is repeated until there are no corridors left in the queue.

This approach reflects some assumptions about the desired characteristics of the corridors identified. First, it does not allow OD pairs to be included in multiple corridors. This is desirable because it prevents the double-counting of expected benefits. Also, without this step, it is possible that many of the corridors identified will be very similar to one another. However, planners may not elect to add bus routes on all of the corridors identified, either because more detailed post-analysis reveals that it is not feasible to add a route in that corridor, or because of factors. In this case, it is possible that a similar potential corridor containing many of the same OD pairs is actually a better candidate. Such a situation may

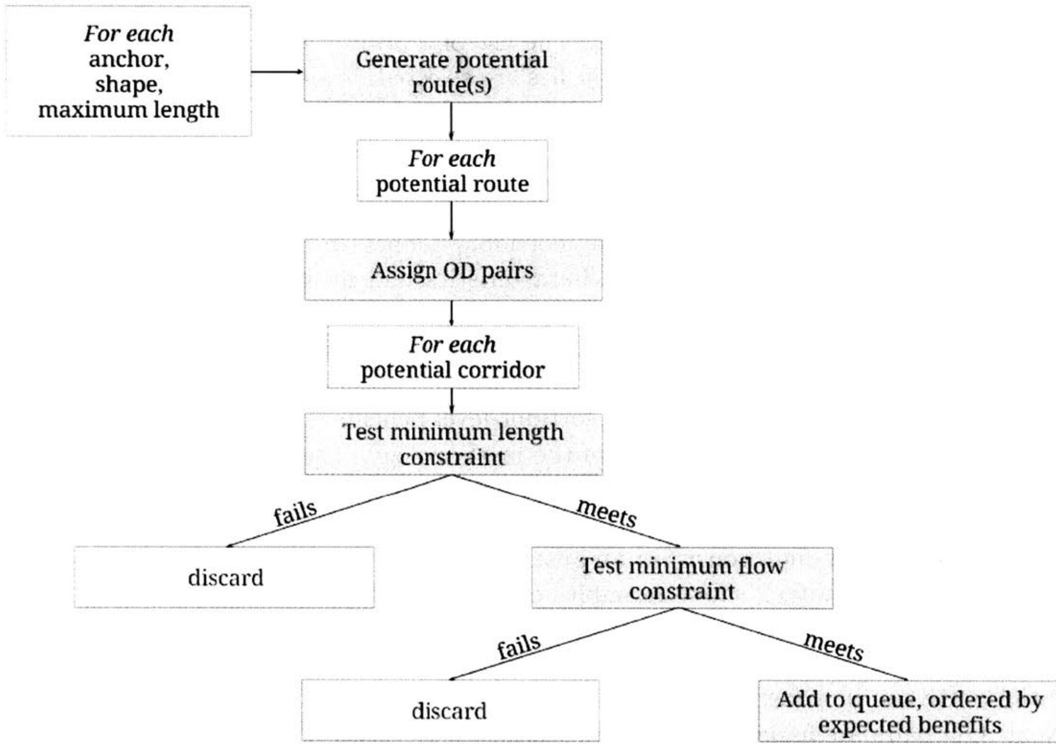

Figure 6-1: Overview of Part 1 of the corridor identification methodology

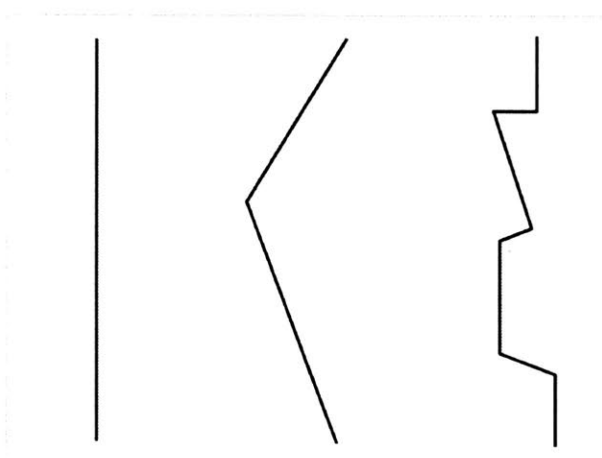

Figure 6-2: Some spatial representations of potential routes

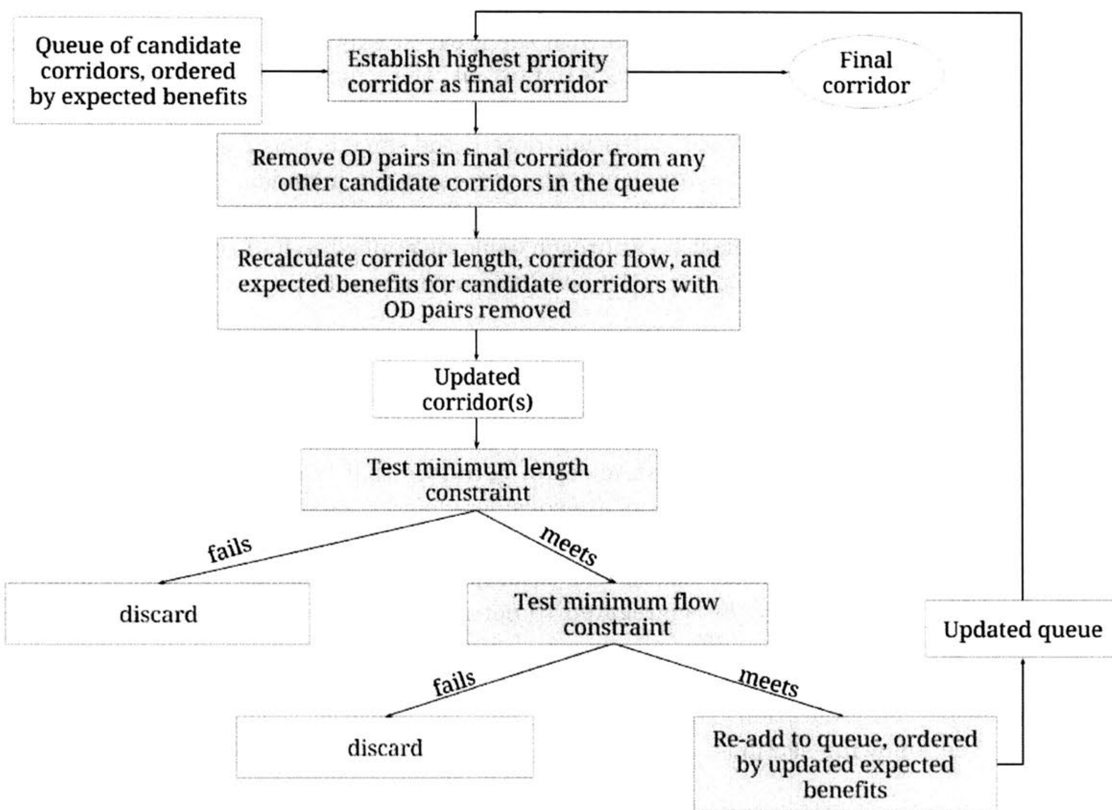

Figure 6-3: Overview of Part 2 of the corridor identification methodology

occur due to the challenges of defining potential corridors that are appropriate for new bus service described in Section 6.3.

This situation can be addressed by including a manual evaluation component. That is, after a set of mutually exclusive corridors are identified, planners can perform post-analysis of the corridors, evaluating the feasibility of adding a bus route to each corridor in more detail. If they decide not to add service in some corridors, they can repeat the second part of the methodology, manually removing these corridors first. In some cases, this will produce a different set of corridors. However, repeating this process many times may be tedious. Planners may prefer to manually adjust corridors in post-analysis in cases where they observe that the corridor identified is not a good candidate for new service, but a similar corridor is.

The approach applied in the second part of the methodology also assumes that the objective is to identify individual corridors with the greatest expected benefits regardless of other factors, such as cost. If costs are very important, the expected benefit-cost ratio for each corridor can be used as the prioritization metric instead of expected benefits. However, the assumption that the cost of each potential route is the same is reasonable for high-level sketch planning, as details such as frequency of service are not included. Finally, a greedy approach is taken to sequentially find the corridors with the greatest expected benefits per corridor. Alternatively, an optimization approach could be applied to find the set of corridors with the greatest total benefits, given a defined budget constraint.

6.3 Defining potential corridors

The first part of the algorithm defines a set of potential corridors. The objective of this part of the algorithm is to define an exhaustive set of corridors that could be served by a new bus route, such that all potential corridors are considered and compared. Potential corridors should be defined such that a potential new bus route in the corridor could serve all the OD pairs assigned to the corridor. Under this assumption, the expected benefits on all OD pairs assigned to a corridor can be aggregated to determine corridor-level benefits. Finally, potential corridors should have sufficient expected demand to justify a new bus route.

Within the context of the methodology presented in Section 6.2, there are many choices regarding spatial representation, corridor definition, and demand. These alternatives are discussed in the following sections.

6.3.1 The potential route

The first set of choices related to the set of potential routes. Potential routes are defined by a set of anchors, length requirements, and shape definition.

Anchors

The anchor defines the location in the network from which a potential route is generated. Anchors may define the route endpoint, midpoint, or another point on the corridor. From this anchor, the length and shape definitions may define one or more potential routes. Multiple potential routes can be generated from a single anchor, as the anchor may be located at the center, the end, or another point in the potential route, and potential routes anchored by the OD pair may also be of varying lengths and shapes. Section 6.3.1 explains how potential routes are generated from OD pairs, given length and shape definitions.

Technically, any point can be an anchor. One possibility is to define anchor points in a grid covering the full network, and allow any of these points to be endpoints of potential routes. However, even with length and shape constraints, this will generate a vast number of potential routes, many of which may not be appropriate for bus service or may not serve any of the OD pairs being clustered. Thus, this approach is inefficient. Two logical alternatives are to use a known set of potential bus termini as anchors or to use OD pairs as anchors.

If the potential locations for bus termini are strictly defined and known, they are the obvious choice for the endpoints of the potential routes. However, limiting potential corridors to a fixed set of endpoints will constrain the potential corridors defined. Therefore, this approach should only be taken if an exhaustive set of possible termini is known.

Alternatively, OD pairs can be used as anchors. Similar to in TraClus-DL, in which each OD pair is considered as a seed, each OD pair can be used as an anchor to define a set of potential routes. Compared to allowing any point to be an anchor, this restricts potential routes to those that serve at least one OD pair. An OD pair can be represented as a straight line or as an actual path through the road network. A potential route generated from an OD pair anchor must contain the OD pair. The OD pair may be at the end or at some point along the potential route. Section 6.3.1 discusses how the location of an OD pair along a potential route can be defined, and how constraints on termini locations and due to geographic barriers can be incorporated with OD pairs used as anchors.

Shape

Given corridor endpoints, the shape of potential routes can be defined in several ways ranging from abstract to realistic. The most abstract representation is as a straight line. This definition does not closely mimic reality, and the actual road network may not accommodate a similar path, particular if there are geographic barriers. Determining if the corridor can accommodate a bus route can be left to post-analysis. Using this shape definition, it is likely that not all of the final corridors identified are appropriate for new service. However, defining linear corridors requires less computation and does not require a representation of the road network. Using simple linear corridor shapes may enable more complex extensions of the algorithm, including the incorporation of optimization and iterative approaches to selecting corridors. The advantages and requirements of such extensions are discussed in Chapter 8.

However, one issue with linear corridors is that some OD pairs cannot be served by a linear (or even pseudo-linear) path due to geographic barriers that limit the road network. One way to accommodate this is to define a set of geographic barriers and prohibit corridors from crossing them. Instead, the corridor shape is defined as the shortest linear path that avoids the barrier(s). This produces corridors that are piecewise linear, as shown in Figure 6-4. Incorporating barriers in this way requires the definition and spatial representation of barriers, but does not require a full road network representation.

In order to incorporate more aspects of the real network, a simplified road network can be used. This may be limited to major roads that are appropriate for bus travel. In this case, corridor endpoints can be connected by the shortest path through the simplified road network. This path defines the shape of the potential route. Assuming the simplified road network does not allow roads to traverse barriers, this definition of potential routes will account for barriers.

The most accurate definition of potential routes uses a representation of the road network consisting of all roads that can accommodate bus services. Potential routes are defined as the

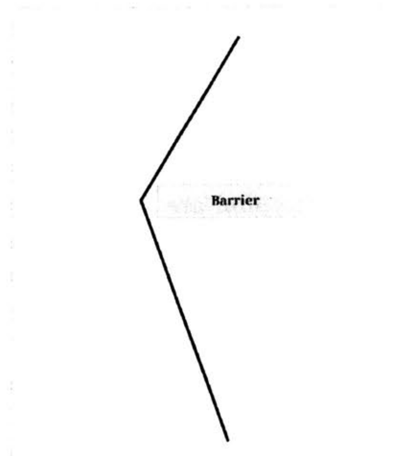

Figure 6-4: Example of defining corridor shape to account for a barrier

shortest path connecting two endpoints through this network. In some cases, two points are connected by multiple paths of very similar lengths. While all of these paths would provide direct service, some may serve more OD pairs than others. Ideally, each of these paths is defined as a potential route, so that the definition of potential corridors is as exhaustive as possible.

Defining more complex shapes requires the storage of more information about each potential route in order to appropriately assign OD pairs, and also requires more computation to determine which OD pairs to assign to each potential route. However, it will reduce the likelihood that the final corridors identified cannot be served due to the shape and constraints of the road network.

Length

Potential new bus routes must be of a reasonable length. Corridors that are too short or too long are not appropriate for bus service. Therefore, both a minimum and maximum length should be defined. These parameters can be set by planners and be adjusted for a particular network, given the desired route lengths.

Corridor length can be measured using the euclidean distance from endpoint to endpoint or using the distance through the road network. As discussed in Chapter 3, road network distance more closely reflects reality. If the road network is used to define the shape of the potential route, it is logical to use the road network to estimate corridor length as well.

However, because corridor length constraints are likely somewhat approximate (planners do not have an exact definition of the minimum and maximum acceptable route lengths) estimating corridor length as the euclidean distance is sufficient. If euclidean distance is used, it is important to account for the fact that the actual route length will be longer than the straight line length measured, when specifying minimum and maximum length parameters.

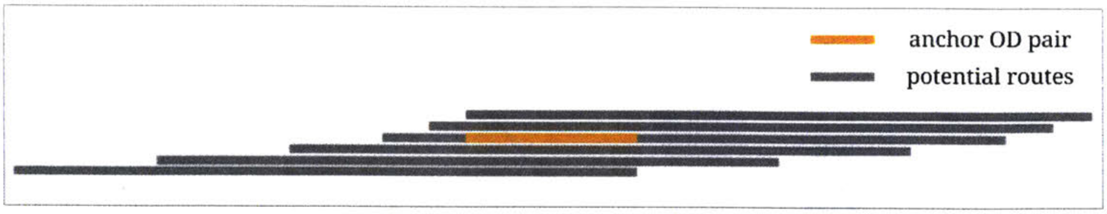

Figure 6-5: Potential routes that can be generated from a single OD pair by shifting position of OD pair in the route

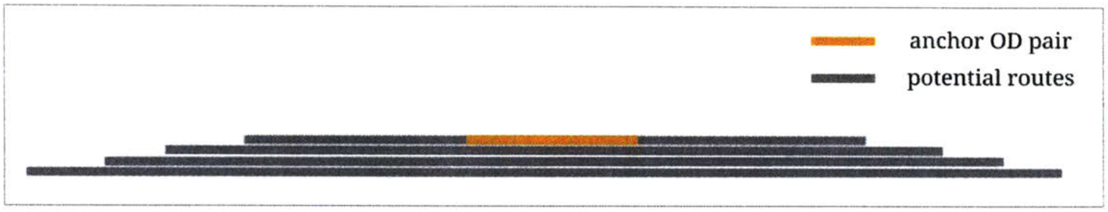

Figure 6-6: Potential routes that can be generated from a single OD pair within a specifed range of lengths

Generating potential routes from anchors, shape, and length definitions

Depending on the choice of anchor, shape, and length, the process for generating potential routes varies. If termini locations are used as anchors, defining an exhaustive set of potential routes that meet minimum and maximum length parameters is fairly straightforward. Pairs of termini points are connected according to the choice of shape (for example, as a straight line or as the shortest path through the road network). Those that meet the length requirements are defined as potential routes.

In some cases, termini locations may not be clearly defined. Using OD pairs as anchors allows more flexibility in the potential routes that can be defined. If OD pairs can be located at any point along the potential route, the set of potential routes that can be generated from an OD pair approaches infinity, as the route can be shifted infinitesimally (See Figure 6-5). Similarly, within a range of allowable potential lengths, the set of potential routes that can be generated from a single OD pair also approaches infinity (See Figure 6-6). Thus, it is necessary to define a discrete set of allowable route lengths and locations of the OD pair within the corridor. Defining smaller intervals between allowable lengths or allowing the anchor to be in more positions along the corridor results in a more exhaustive definition of potential corridors, but it also requires more computation. The number of potential corridors generated for each anchor OD pair is the product of the number of allowable positions and the number of discrete lengths.

If there are many OD pairs being clustered, and a large number of potential routes are generated for each, it is likely that many will be similar and have significant overlap. The overlap is eliminated in the second part of the algorithm. However, there is a trade off between exhaustively defining potential routes and the required computation which impacts the running time for the algorithm. Generating more potential corridors ensures that the highest priority corridors are defined, but if the potential routes are very similar, they will impact the final results only slightly, while requiring significant additional computation.

The simplest approach is to require potential routes to be centered on the midpoint of the

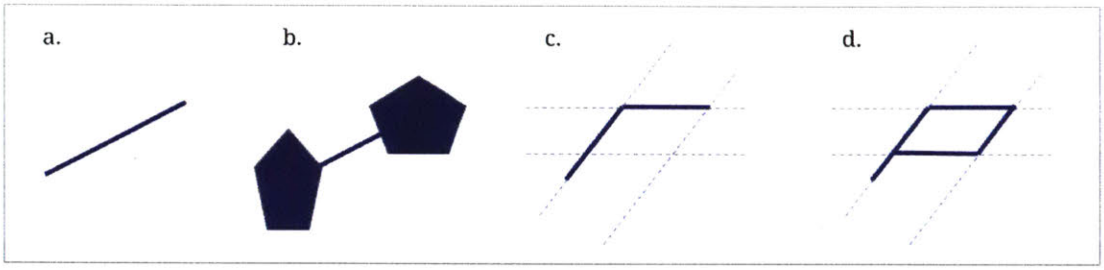

Figure 6-7: Possible spatial representations of OD pairs

anchor OD pair. This is the approach taken in the implementation discussed in Section 6.4. With this approach, OD pairs at the edge of the network or that abut a geographic barrier will generate potential routes that are illogical. These potential routes can be eliminated, or adjustments can be made in these circumstances. Given these issues, an improved approach requires the anchor OD pair to be located at one end of the potential route. Compared with the midpoint approach, this doubles the potential routes generated for each OD pair, given the same number of allowable routes. This therefore increases computation, but it produces a more exhaustive and logical set of potential routes. For OD pairs next to barriers or the edge of the network, potential routes can be limited to those extending away from the barrier or network edge.

Using OD pairs as anchors is appropriate when termini locations are not explicitly predefined. In these cases, planners may know that some locations are not appropriate for termini locations. This can be accounted for in post-analysis, adjusting or eliminating corridors as needed, or it can be addressed at this stage. To address it here, potential corridors generated with endpoints at locations that cannot be termini should either be adjusted (shifted) or eliminated.

While figures 6-5 and 6-6 show potential routes defined as straight lines, other shapes can be used, as previously discussed. In order to define a shape generated from the road network, the anchor OD pair is represented as a straight line and the endpoints of potential routes for that anchor are defined based on the discrete definitions of length and anchor position. Given the route endpoints and a representation of the road network, the shortest path (or multiple similar shortest paths) can be defined.

6.3.2 Assigning OD pairs to potential routes

OD pairs are assigned to potential routes based on similarity measures. The similarity measures should determine if the OD pair can be served by the potential route. The first methodological choice is in how to spatially represent OD pairs. They can be represented as straight lines, as a pair of connected zones, or as an actual path (or multiple paths) through the road network, connecting origin and destination (See Figure 6-7). In the context of this methodology, representing OD pairs as actual paths through the network does not provide any benefit. Whether or not a potential route can serve an OD pair depends on the location of the origin and the destination, not the path between them.

Representing origins and destinations as zones is a more realistic depiction of zonal OD pairs. In this case, similarity can be measured using the boundaries of zones, rather than the centroids. However, if zones are fairly consistent in shape and size, it is not necessary to model each one. Instead, the zonal radius should be accounted for when setting distance

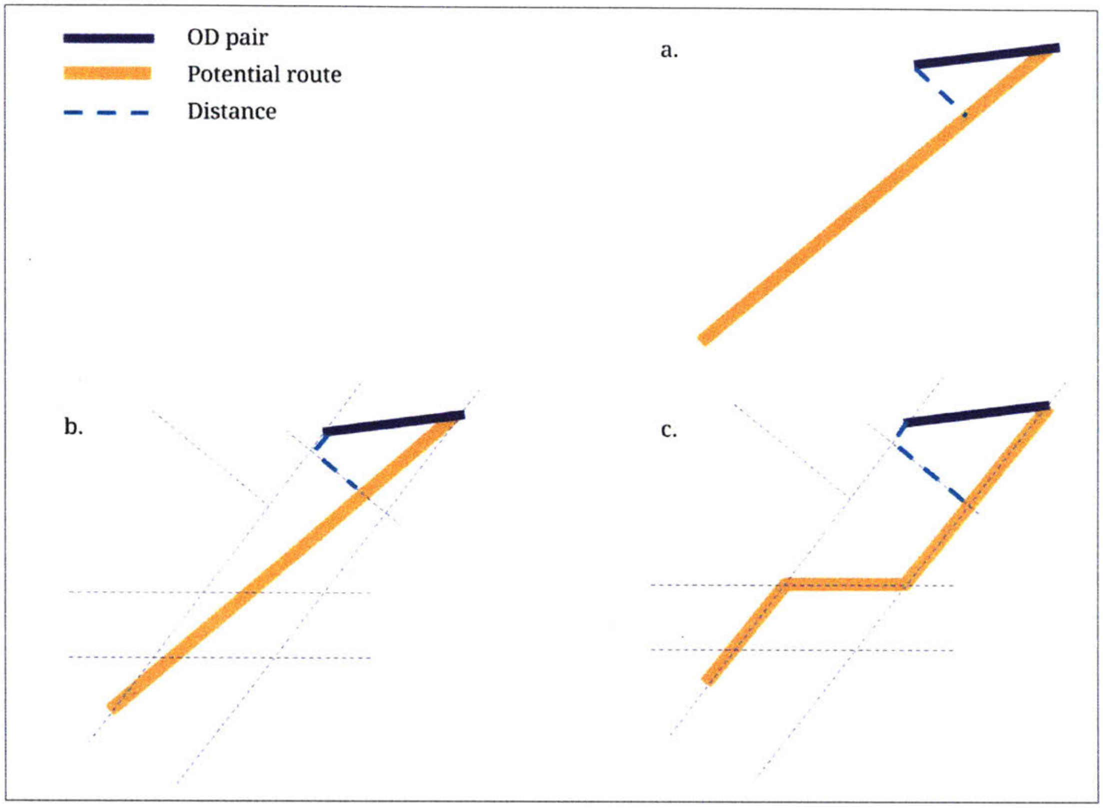

Figure 6-8: Evaluating distance from OD pair to route

parameters.

The distance from an origin (destination) point to a potential route can be assessed as the euclidean distance from the point to the route (See Figure 6-8a), or as the distance through the road network (See Figure 6-8b and c), or (ideally) the walking network, assuming routes are accessed on foot. In each case, the minimum distance from the point to the route should be taken. That is, the distance is to the closest location on the route. This is trivial for the euclidean distance case. For road network, this requires computing multiple shortest paths and finding the shortest.

The specification of a maximum distance parameter effectively defines a buffer or radius extending from the potential route. Only OD pairs with both origin and destination points within the radius are assigned to the route. The maximum distance parameter should be specified taking into account the average zone radius defined in Step 1 of the framework and information about access and egress distance in the network.

Using road (or walking) network distance defines a more realistic buffer and ensures that if there is a barrier between an OD pair and a potential route, the OD pair will not be assigned to it (as the barrier will result in a large distance estimate). Using euclidean distance ignores the intricacies of the road network and the locations of barriers, but is straightforward to compute, which is useful when assigning large numbers of OD pairs to large numbers of potential routes. One alternative is to define a network of barriers and measure distance as the shortest straight line path around a barrier, as in Figure 6-8b.

Some short OD pairs may have origins and destinations within the maximum distance

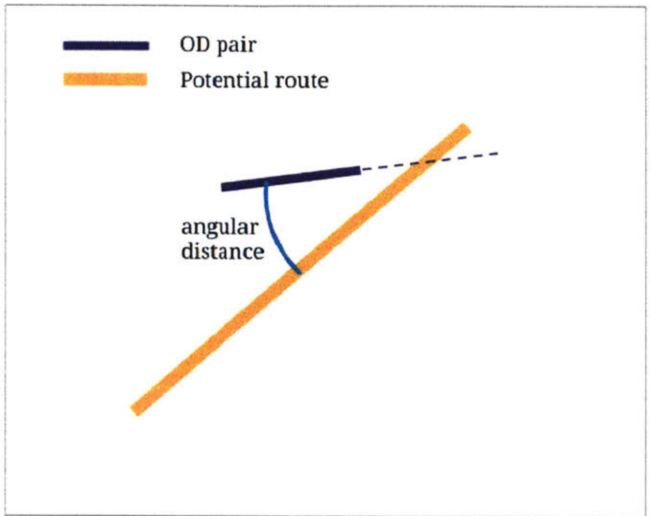

Figure 6-9: Corridor angle restriction

of the potential route, but if the direction of the OD pair and the route is not similar, the potential route does not serve the OD pair. To eliminate these OD pairs, a maximum angle parameter is defined. Angular similarity is defined as the difference in angle between an OD pair and the potential route. Only if this difference is within the maximum angle specified are OD pairs assigned to a potential route. Corridors are defined as bidirectional, as bus service is expected to be provided in both directions. The angular distance is always defined as the smaller of the two angles formed at the intersection of the route and an OD pair (or the linear extension of the OD pair), as in Figure 6-9. That is, angular distance is defined as in Equation 6.1, where a_r is the angle of the route and a_{od} is the angle of the OD pair, both in degrees.

$$\Delta a = min\left(|a_r - a_{od}|, |a_r - a_{od}| - 180\right) \tag{6.1}$$

In the case that potential routes are not represented as straight lines, the angle of the route can be defined based on the straight line connecting its endpoints, which represents a spatial simplification. Because the maximum angle rule is intended to eliminate short OD pairs with angles that diverge from the potential route (longer OD pairs with angles that diverge from the route are eliminated by the maximum distance constraint), assessing the angle locally is is more relevant, as shown in Figure 6-10. This applies a superior angle filter. As angles may differ in either direction from the seed OD pair, the difference in angles between any two OD pairs that are members of the same provisional corridor will not exceed twice the maximum angle parameter specified.

Finally, OD pairs that extend beyond the length of the specified potential route are not assigned to a potential route. This is determined by projecting the origin and destination points of and OD pair onto the straight line representation of a corridor (even if corridors are defined as nonlinear, a straight line representation from end to end can be constructed). If the projections do not fall onto this linear representation, the OD pairs are outside of the potential corridor.

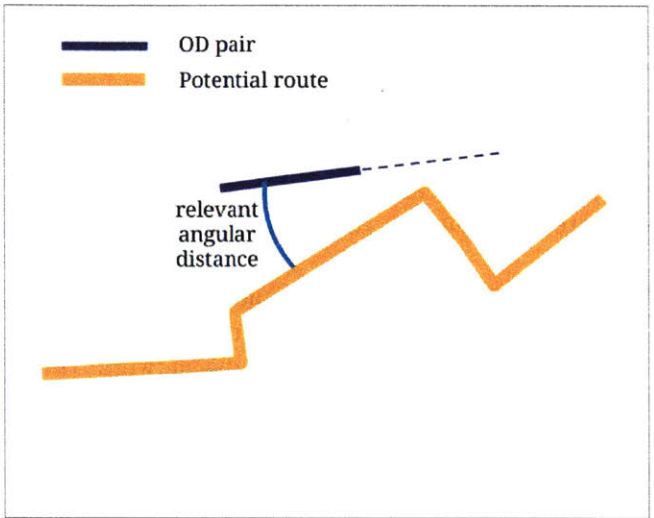

Figure 6-10: Relevant angular distance

6.3.3 Evaluating potential corridor flow

Corridors identified should have sufficient expected demand to justify adding bus service. Corridor-level demand can be evaluated using a variety of measures including journeys, journey-miles, and journey-minutes. The total number of journeys is often too simplistic a measure, as it does not account for journey length. For public transport planning, vehicle load and route flow are also commonly used. Information on peak loads (flows) is critical for frequency setting and generally routes with evenly distributed flows are desirable.

One suitable method for testing potential corridors to determine if they have sufficient demand is to enforce a minimum flow constraint. Flow on a potential corridor can be estimated as the expected hourly flow averaged along the corridor. This measure does not fully capture the shape of demand, such as spatial peaking, but it does account for flow relative to the length of the route, and it is easy to understand. The minimum flow parameter can be interpreted as the number of bus passengers that will pass a point on the route in one hour, averaged across all points along the route. It is estimated as the sum of journey-miles per hour on all OD pairs assigned to a corridor, divided by the length of the corridor.

In some potential corridors, there may be a gap in flow. If geographic barriers were not specifically accounted for in the definition of potential routes, gaps in flow may be the result of these barriers, which can make travel between two points time-consuming, resulting in low levels of current demand and expected demand. Regardless of cause, gaps in flow are not desirable in corridors as candidates for potential bus service. If there is a gap in flow, there is no reason to provide service along the entire corridor. Instead, separate routes can serve demand on each side of the gap.

To eliminate potential corridors with gaps in flow, flow can be estimated at the segment level for potential corridors. If any segment does not meet the minimum flow requirement, the potential corridor is eliminated. Given that many overlapping potential routes are defined, eliminating those with gaps in flow helps ensure that other potential corridors with more consistent flow are identified. This would require the specification of the segment size for which flow is evaluated. Particularly for small segments, estimating flow by segment is computationally more expensive than at the corridor level.

The estimated flow on a potential corridor uses the estimates of expected demand from Step 2 in the framework. These estimates should include journeys that shift from existing services and new journeys attracted to the network. These estimates are highly dependent on the assumptions and methods used to estimate each of these demand sources. When setting the minimum flow parameter, one must be cognizant of the possible over or under-estimation of demand.

There are several reasons that demand may be under or over-estimated. First, if a simple demand estimation procedure such as the elasticity method is applied, estimates may not account for variation in latent demand throughout the network. In particular, OD pairs with few or no current public transport journeys were excluded from the analysis. These OD pairs may have latent demand that is not captured. Second, demand estimates depend on the definition of expected travel time and target distance that were used to determine if OD pairs are improvable or not and to estimate potential travel time savings. Particularly if current demand is assigned using all-or-nothing assignment, demand estimates are highly dependent on these definitions.

Another source of potential demand is intrazonal journeys. First, intrazonal journeys were excluded from the OD-level analysis. The benefits of new services to these journeys is uncertain, dependent on the specific within-zone origins and destinations. Therefore the expected benefits for these journeys are not estimated. However, it is likely that some of these journeys will shift to a new bus service. Under this assumption, a portion of intrazonal journeys can be assigned to potential routes. Information about the bus mode share for short journeys within the network can be used to inform the portion of intrazonal journeys expected to use the new service. Or, the estimation of the demand impact of intrazonal journeys can left to post-analysis.

Finally, a new service may benefit one stage of a multi-stage journey. As described in Chapter 5, improvable stages that are part of non-improvable OD pairs have already been included, but improvable OD pairs that require multi-stage service were not split into stages. This prevents possible double-counting, as discussed in Chapter 5. However, if these OD pairs are not assigned to any corridor, and hence no service is introduced to serve the full journeys, journeys stages that benefit from a potential new service are likely to shift to it. To ensure the same potential benefits are not counted on multiple corridors, this demand and the accompanying potential benefits can be evaluated in post-analysis. If an improvable OD pair is not assigned to any corridor, the stages can be distributed to the corridors identified. The only way this could be incorporated within the corridor identification algorithm would be to adapt an iterative approach to corridor identification. This type of extension is discussed in Chapter 8.

Overall, depending on methods for estimating current and induced demand, and for including OD pairs in the corridor identification algorithm, corridor flows estimated within the algorithm can vary. The biases contributing to under or over-estimation of flow should be accounted for when setting the minimum flow parameter. The minimum flow parameter can also be informed by planning standards and flows on existing routes.

6.3.4 An OD-chaining approach

The previous sections defined a sequential approach for the definition of potential corridors. First, a discrete set of potential routes are define. Then, OD pairs are assigned to potential routes and the flow requirement is tested. In the absence of strict design constraints, such as a fixed set of possible bus termini, it is impossible to generate an exhaustive set of potential corridors. Any actual algorithm will only approximate an exhaustive set, as discussed in Section 6.3.1. The challenge therefore is to ensure that the most promising potential routes are defined. Promising routes serve sufficient flow and have significant expected benefits.

While the approach described above begins with the definition of a potential route and then assigns OD pairs, an alternative approach could define potential corridors by chaining together OD pairs. Such a method draws on density-based scan approaches, in which clusters are expanded as long as density persists. However, in this case, clusters would only be allowed to expand into a shape that is appropriate for bus service. Given this constraint, there are methodological questions about how clusters should be expanded, as different methods may result in different corridors. Density-based scan methods define a neighborhood around each seed. Given the shape requirements, defining a neighborhood is challenging. The same maximum distance and maximum angle parameters and measures described in Section 6.3.1 could be used as part of the neighborhood definition. But the neighborhood must also account for corridor length. With OD pairs of different lengths, incremental expansion of clusters as corridors are defined and grown is not straightforward. The development of such a methodology and the assessment of whether it is superior to the methodology described here is left to future work and discussed in Chapter 8.

6.4 An algorithm implementation

Section 6.2 defined a methodological structure for the identification of corridors, and sections 6.3.1, 6.3.2, and 6.3.3 described many methodological alternatives within that structure. For the application of this methodology to the London network, described in Chapter 7, an algorithm following this structure was conceived and implemented.

The algorithm implemented uses each OD pair as an anchor with the midpoint of the OD positioned at the midpoint of potential routes. The algorithm requires the specification of a minimum and maximum route length, and routes of three different lengths are considered (a route of the maximum length, and routes of two intermediary lengths, evenly spaced between the minimum and maximum lengths). Potential routes are represented as straight lines, and consideration of barriers and constraints on termini location are left to post-analysis.

OD pairs are represented as straight lines and assigned to potential routes if they meet the maximum distance and maximum angle constraints. Maximum distance is measured as the euclidean distance from the origin (destination) point to the closest point along the potential route. The minimum flow requirement is assessed base on average expected flow on a potential corridor.

Finally, corridors are placed in a priority queue, ordered by potential benefits. Potential benefits are defined as the total potential travel time savings for existing passengers (t, in journey-minutes) and the total new journeys expected (d, also in journey-minutes). The two quantities can be combined with planner-specified weights, as in Equation 6.2.

$$b = \beta_t t + \beta_d d \tag{6.2}$$

Final corridors are established based on the order of potential corridors in the queue, using the approach described in Section 6.2. Once an OD pair is assigned to a final corridor, it is removed from any other potential corridors, the flow requirement is re-assessed for all altered corridors, and the potential benefits are re-estimated. Chapter 7 demonstrates the application of this algorithm in the London case study and reflects on the potential impacts of different methodological choices and possible improvements.

6.5 Step 3 outputs and post-analysis

This final step in the framework produces a set of corridors that meet the definitions of shape, length, and flow according to the parameters specified. Each corridor consists of a set of improvable OD pairs and every improvable OD pair identified in Step 2 is assigned to (at most) 1 corridor. Expected demand and potential expected benefits are estimated for the entire corridor and for each OD pair within the corridor.

As described in Section 6.3, there are many factors that influence the suitability of a corridor for bus service, including geographic barriers, constraints on termini location, and intricacies of the road network. Section 6.3 proposed methods for addressing these issues within the corridor identification methodology. However, if they are not adequately addressed, these issues can also be considered in post-analysis of the corridors identified. The post-analysis should include detailed analysis of the flow along each corridor, modifications of corridors to account for constraints on termini locations and geographic barriers, and consideration of the road network to determine if it can accommodate a bus route that will serve all OD pairs assigned to the corridor well.

Another important component of post-analysis is developing an understanding of the potential impacts of new services on existing services. In Step 2 of the framework, the expected shift in demand from current public transport services to a new route was estimated. These estimates can be augmented with more detailed analysis, which can be used to make decisions about the existing service. As an extension to the framework presented here, an iterative application may be able to make recommendations about the removal of routes. The requirements and possibilities for such an approach are discussed in Chapter 8.

Some improvable OD pairs likely will not be assigned to any corridor. While these OD pairs could be improved with more direct bus service, they are not located along a corridor with other improvable OD pairs. It does not make sense to improve service on these OD pairs by introducing an entire new bus route. However, there may be other types of improvements that can benefit these OD pairs. This is discussed further in Chapter 8.

Chapter 7

Applying the framework to London's bus network

This chapter summarizes the process and results of applying the bus network sketch planning methodology to the London public transport network. Section 7.1 describes the data used. Section 7.2 explains the definition of zones (Step 1), including the selection of zone size. Sections 7.3 and 7.4 illustrate how parameters used in steps 2 and 3 of the framework are informed by data on bus travel in London.

Section 7.5 describes the sensitivity of results of the framework to parameter selection. Section 7.6 delves deeper into the results, presenting more detailed analysis of corridors identified by the methodology, and Section 7.7 discusses the contributions of the application to the London network, as well as some shortcomings and possible extensions.

7.1 Data

The inputs to the bus network sketch planning framework fall into two categories. The first is data on public transport journeys in London and the second is origin-destination (OD) level shortest path distance and travel time data.

7.1.1 Public transport journey data

The fare card data used for the case study consists of 10 weekdays of Oyster (London's smart card) transactions from October 2012 and February 2013. In that period, Oyster cards were used by approximately 80% of rail passengers and 90% of bus passengers (Gordon et al., 2013). The data was processed using ODX, a methodology developed by Gordon et al. (2013), which infers origin bus stops, alighting bus stops, and links stages of multi-stage journeys using automatic vehicle location data and geographic and time-based thresholds. For simplification, only journeys of up to three stages were included (with stages here defined as being initiated with a smart card tap), which represent 99.4% of all ODX-inferred journeys.

The ODX methodology inferred the starting and ending stop (station) for 81% of journeys made using an Oyster card in the two weeks analyzed. Of the Oyster journeys that did not have a starting and ending stop inferred, 69% were single-stage bus journeys with an origin stop and bus route but no destination stop inferred. This is because in the London network, passengers must "tap out" of the rail network, but do not "tap off" after bus journeys. The

ODX methodology infers the alighting stop for bus journeys based on the next stage of the journey or the next (same day) journey, meaning that destination stops for bus stages without a continuation or another journey that day cannot be inferred.

Because the proposed analysis is OD-based, journeys without destination stops cannot be included, meaning that the disproportionate inference rate for single-stage bus journeys would result in under-representation of bus journeys. To correct for this under-counting, each of the single-stage bus journeys with no inferred destination stop was assigned an alighting stop by the following methodology: For each boarding stop and route, a destination stop distribution was constructed, consisting of the frequency of occurrence of all inferred downstream destination stops for single-stage bus journeys originating at that stop and route. Then, for each journey beginning at that boarding stop which did not have an inferred destination, a destination stop was selected at random from this distribution. This methodology assumes that single-stage journeys with uninferred destinations have the same destination distribution as single-stage journeys with inferred destinations.

Through the inference methodologies, origin and destination stops are inferred for 94% of journeys paid for with an Oyster card, or 90,306,224 journeys in the 10-day period. Assuming Oyster cards were used to pay for approximately 85% of journeys in the period (Gordon et al., 2013), this data represents about 80% of all journeys in the period.

7.1.2 Shortest path data

Shortest path distances and travel times for OD pairs were queried using the Google Maps Distance Matrix API. The API provides estimates of point-to-point car distance and travel time for a specified time of day, based on historical data. This represents the shortest path distance and time estimates accounting for expected traffic conditions.

Ideally, paths that include roads that buses cannot travel on, such as very high speed roads and very narrow roads, should be excluded. The Google Maps Distance Matrix API does not allow the user to exclude specific links. However, users can select estimates that "avoid highways", limiting the influence of high-speed roads on travel time and distance estimates.

As inputs, the Google Maps Distance Matrix API takes addresses, coordinates, or points of interest to define the OD pair. The coordinates of the bus stop closest to the centroid of the origin (destination) zone, on a bus route serving the OD pair was selected as the origin (destination) point. This ensures that the origin and destination points are on roads in the appropriate direction of travel. Compared to using the centroid coordinates, using the bus stops avoids extraneous travel time or distance, particularly in cases of dual carriageways.

7.2 Defining zones (Step 1)[1]

The first step in the framework consists of defining a set of zones. As discussed in Chapter 4, the k-means algorithm is recommended for the definition of zones. Using stop and station coordinates as clustering variables, the k-means algorithm identifies clusters of stops and stations in close proximity. The k-means algorithm takes the number of clusters as an input. Because k-means tends to identify clusters of similar size, this parameter influences the size of zones. For application to a particular network, information about typical access and egress distances can inform the choice of "k". In addition to varying the number of

[1]This section includes some parts heavily drawn from Viggiano et al. (2015)

Table 7.1: Average zone size by number of zones

Number of Zones	750	1,000	1,250
Average zone size (square miles)	0.82	0.62	0.49
Approximate average radius (miles)	0.5	0.45	0.4

zones, different methods can be used to weight stops and stations within the clustering process. Zones will be defined such that higher weighted stops or stations are more likely to be located close to the center of zones.

The following sections demonstrate the zone definition process for the London network. Section 7.2.1 presents information on access distances in London and explains how this informs the selected number of zones. The application of the silhouette score, a metric that evaluates cluster tightness, is discussed in Section Section 7.2.2. Then, Section 7.2.3 shows how smart card registration address data is used to evaluate different weighting methods.

7.2.1 Considering access distance distributions

Access and egress distances vary depending on many factors including individual characteristics, trip purpose, and mode. Data from the London Travel Demand Survey (LTDS), an in-person household level travel diary survey, suggests the average straight line distance from a journey's origin to the boarding bus stop or rail station is approximately 0.4 miles. Access distances to bus stops tend to be shorter than access distances for rail stations: 95% of bus access distances are shorter than 1 mile, according to the survey, while 95% of rail access distances are less than 1.25 miles.

Given the focus on bus network design, the access distance distribution preceding bus stages is considered in more detail. Figure 7-1 shows the distribution of access distances for bus stages using true origins inferred from Oyster card registration data. In London, Oyster card registration includes providing a home postcode. Assuming that the first journey of the day began at this home address, the registration postcode can be linked to the Oyster card journeys analyzed to infer the true origin for these journeys. Overall, the access distance distributions inferred using this linkage methodology were found to be similar to data from the London Travel Demand Survey (Viggiano et al., 2016). Not all Oyster cards are registered, and some registration postcodes were unrealistically far from the initial boarding stop or station, suggesting the individual either moved, did not begin the first journey from home, or registered a non-home address. Using postcodes within 1 mile of the boarding bus stop and within 1.25 miles of the boarding rail station, origin postcodes can be inferred for 343,775 journeys.

Figure 7-1 shows that most bus stages are preceded by very short access segments. In fact, 90% of access distances in the sample are less than 0.5 miles. The shape of the distribution, with a high concentration of very short access distances and a long tail, highlights the challenges in determining an appropriate zone size. If zones are too small, access distances in the tail of the distribution will extend into neighboring zones. However, larger zones will include unreasonable walking distances, given that access distances, particularly preceding bus stages, tend to be very short.

Table 7.1 shows the resulting average zone sizes for 750, 1,000, and 1,250 zones. The table also presents the approximate average diameter of zones for these values of k, assuming circular zones.

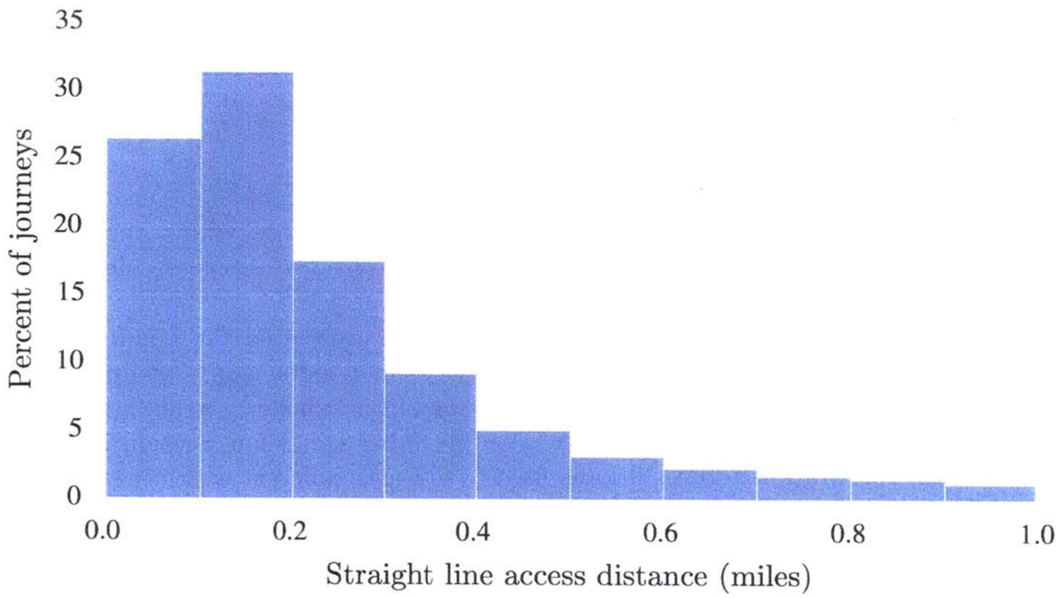

Figure 7-1: Distribution of access distance to first bus stop

For 750 zones, the average zone radius is 0.5 miles. Having fewer than 750 zones is likely to produce zones that include walking access distances that many individuals are unwilling to undertake. Defining more than 1,250 zones is also unreasonable. With 1,250 zones, the average zone radius is about 0.4 miles. This means that access distances in the tail of the distribution, measured from the center of a zone, will extend into neighboring zones. With more zones, it will become increasingly likely that individuals access stops or stations in a given zone from neighboring zones.

7.2.2 Silhouette score

The average silhouette score, is estimated for 750 through 1,250 clusters. The silhouette score evaluates cluster tightness (the closeness of points within a cluster) and cluster separation (the distance between clusters) (Rousseeuw, 1987). Cluster tightness, denoted by $a(i)$ is the average distance from a point to all other points in the same cluster. Cluster separation, $b(i)$, is measured as the average distance from the point to all points in the closest neighboring cluster (based on the distance between point i and all other cluster centroids). The score for each point is calculated as:

$$s(i) = \frac{b(i) - a(i)}{max\{a(i), b(i)\}} \tag{7.1}$$

The resulting score is between -1 and 1. Scores that are lower than 0 indicate that the point is not well-clustered, and is a better fit with another cluster. A high score suggests the point is well-clustered in that it is much closer to other points in its own cluster than to the closest neighboring cluster. The silhouette score can be aggregated to the cluster level by taking the average $s(i)$ for all points in a cluster, or it can be aggregated to the entire

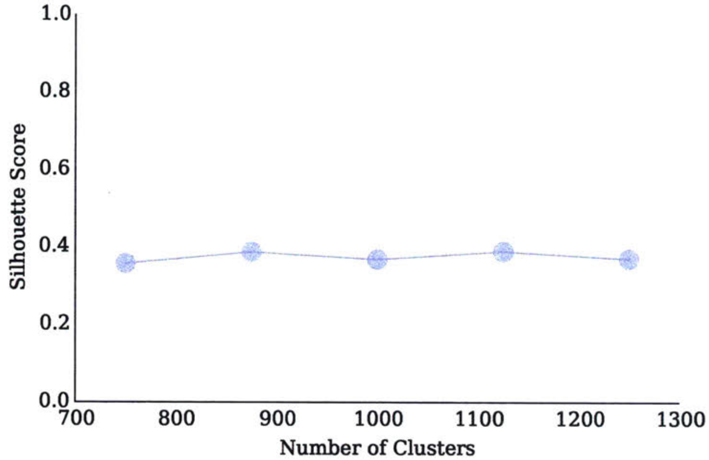

Figure 7-2: Silhouette scores for different numbers of clusters

data set by taking the average $s(i)$ for all points. The score for the full data set can be used to compare different clustering results, and to select the appropriate number of clusters. The best choice is considered to be the number of clusters that produces the largest overall average silhouette score (Rousseeuw, 1987).

Figure 7-2 shows the results. For the silhouette score, a peak indicates the optimal clustering. However, for the stop and station data, the silhouette score remains flat for 750 through 1,250 zones.

The lack of insight from the metric suggests that the stop and station data may not have a strong underlying cluster form. This is not unexpected, as public transport planners typically design stops to be (roughly) evenly spaced to provide good coverage. The London network has dense, even coverage and while there are clusters of stops and stations at activity centers and intersections, not all stops and stations belong to a clear cluster.

Ultimately, given the lack of insight from the silhouette score, the number of clusters was selected to produce zones that match the zone size dictated by the access distance distributions. At the center of the range of values considered, 1,000 zones produces zones with an average radius of approximately 0.45 miles, which is a reasonable zone size, given the access distance distribution.

7.2.3 Weighting stops and stations in the clustering algorithm

The true origins inferred from the Oyster registration data can be used to evaluate different stop and station weighting methods. The percent of journeys for which the zone to which the boarding stop (station) was assigned matches the zone in which the postcode is located (postcodes in London are small, representing on average 15 properties and are geocoded to a single point) is used as an evaluation metric for three different weighting methods.

In the first case (unweighted), each stop and station is treated equally. For the second case (journey-weighted), stops and stations are weighted by number of boardings. Weighting stops and stations by boardings will draw zonal centroids closer to high ridership stops and stations, ensuring they are more centrally located within each zone. The third case is a simplification of the second case. Instead of weighting each stop and station by ridership,

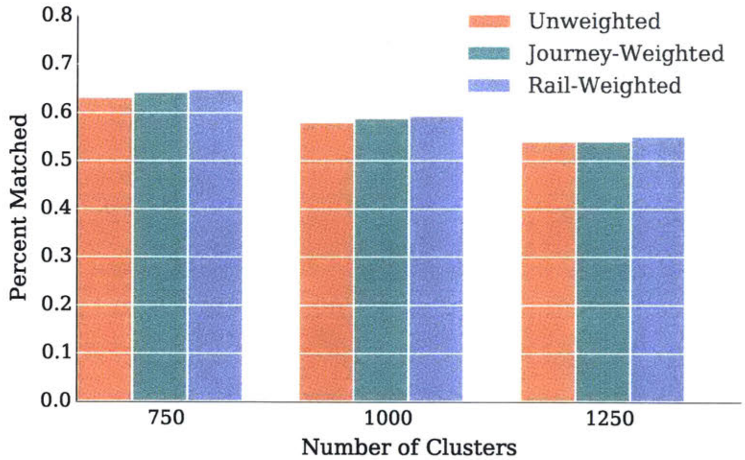

Figure 7-3: Postcode and stop (station) match rate by weighting method

a flat weight of 10 is applied to all rail stations (rail-weighted) to reflect the fact that rail stations typically have many more boardings than bus stops.

Figure 7-3 shows the percent of journeys with postcodes assigned to the same zone as the boarding stop or station for the three weighting methodologies with 750, 1,000, and 1,250 clusters. The matching rate is expected to decline as the number of clusters increases, due to the resulting decline in zone size. The rail-weighted method consistently provides the highest rate of stop and postcode matching. While the differences are small, a t-test reveals that the differences between the rail-weighted and unweighted methods are statistically significant. Furthermore, the differences between the rail-weighted and journey-weighted methods are also statistically significant. The fact that the rail-weighted method outperforms the journey-weighted method is not intuitive, as it would be expected that the journey-weighted method would ensure that high-ridership stops and stations are in the center of zones, increasing the chance that they are accessed from postcodes within the zone. However, given the wide variation in boardings per stop and station, the journey-weighted method results in geographically small clusters surrounding high ridership stations which reduces the chance of the postcode matching the boarding stop or station zone. Given that the rail-weighted method is simple to apply and produced the best results, the rail-weighted method was selected.

Of note is the fact that the matching percentage will decline as the zone size decreases, therefore this metric is not useful for selecting the number of zones.

7.2.4 Stop and station clustering results

Because stops and stations are more heavily concentrated in Central London, zones are smaller at the center and larger at the periphery. The number of stops and stations per zone also varies, as shown in Figure 7-4. Figure 7-5 displays the zones generated for a portion of Central London. Compared to existing zonal schemes, shown in Figure 7-6, it is clear that zones generated using k-means clustering have more rail stations and bus stops at the center of zones. Figure 7-7 shows all 1,000 zones.

Applying the match rate indicator described in Section 7.2.3 confirms the superiority of

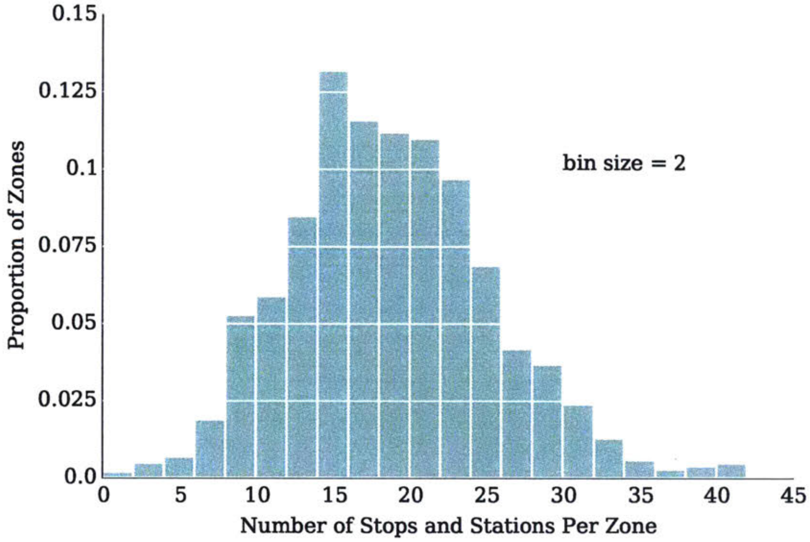

Figure 7-4: Distribution of rail stations and bus stops per zone

the zones defined with the k-means clustering method. The London Transportation Studies (LTS) model defines 949 zones in Greater London, and the percent of linked postcodes located in the same zone as the boarding stop is just 52.8%. This significantly underperforms all the k-means clustering results presented in Figure 7-3. The match rate for 1,000 zones using the rail-weighted method is 59.4%. Journeys are assigned to zones according to their initial and final stop (station). As expected, the zonal OD matrix is sparse; 48% of the OD cells are empty.

7.3 OD analysis (Step 2)

Step 2 in the bus network sketch planning framework consists of analysis at the zonal OD pair level to identify which OD pairs might be improved by the addition of a new bus route and to estimate the benefits a new route might produce.

7.3.1 OD pair constraints

In this application, a few restrictions are made. First, the analysis is constrained to AM peak travel, and travel to and from Central London is excluded. In addition, a minimum sample size constraint is applied to ensure estimates of travel time are reliable. Finally, intrazonal journeys are excluded. The basis for, and effects of, each of these restrictions are explained in the following sections.

Time of day

Most bus routes provide service throughout the day, but planners often focus on peak periods, when demand is greatest. Here, the analysis is performed for the AM peak period. If planners are interested in other periods, similar analyses can be performed.

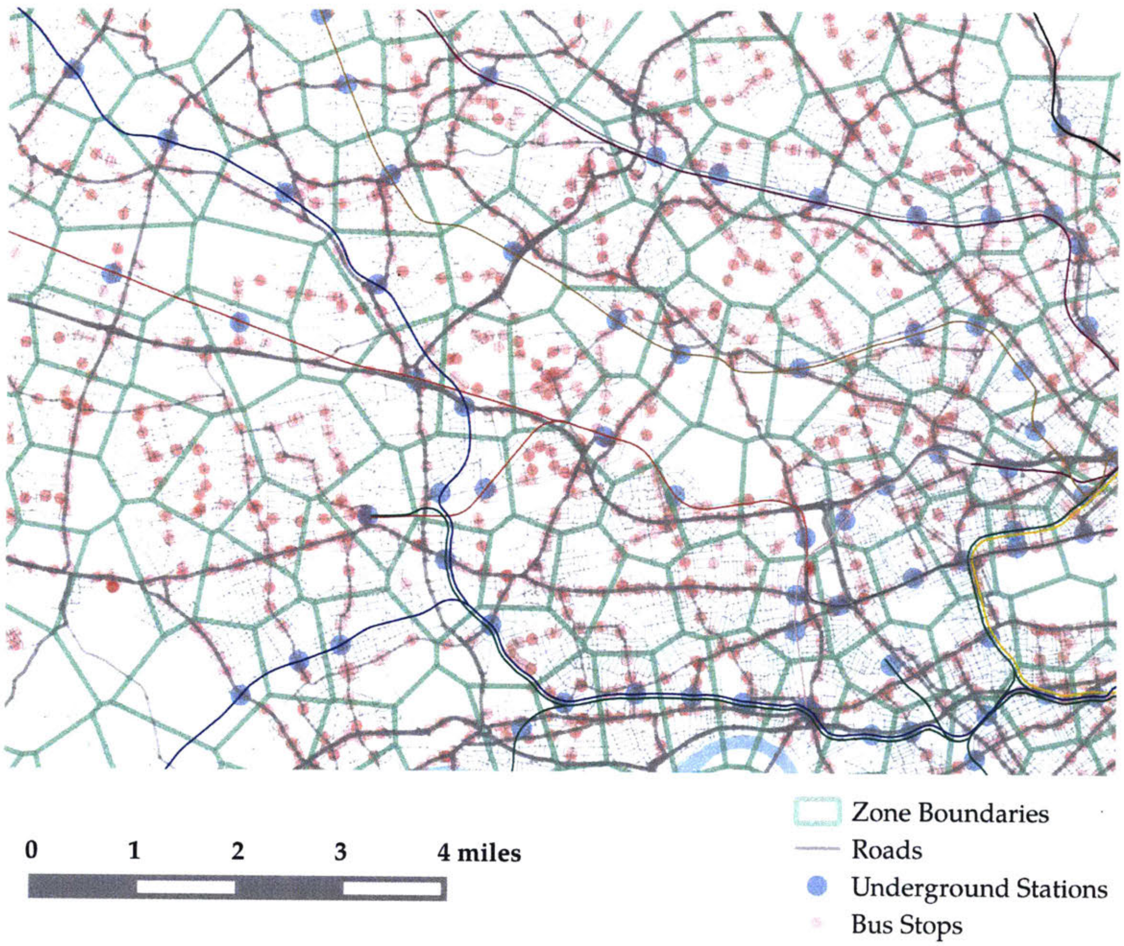

Figure 7-5: Central London zones

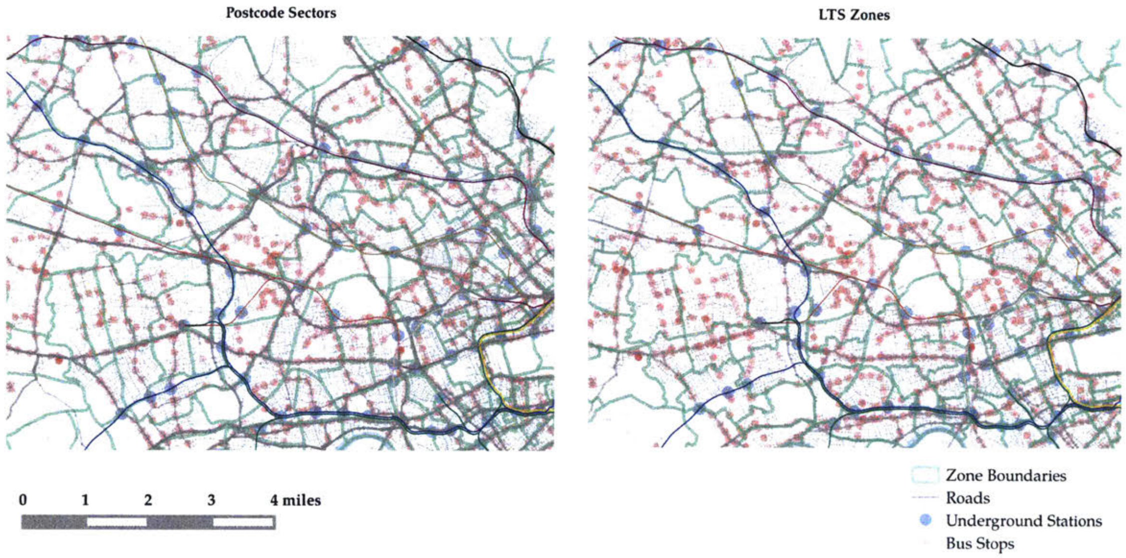

Figure 7-6: Existing zonal schemes

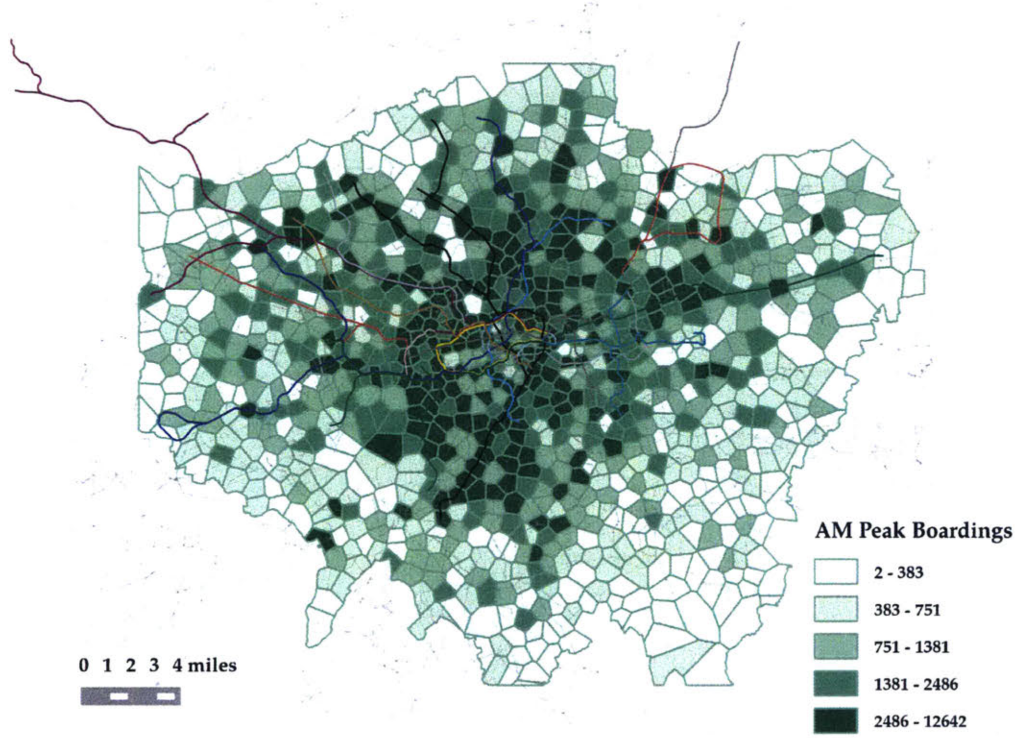

Figure 7-7: Map of all zones displaying total AM peak boardings per zone

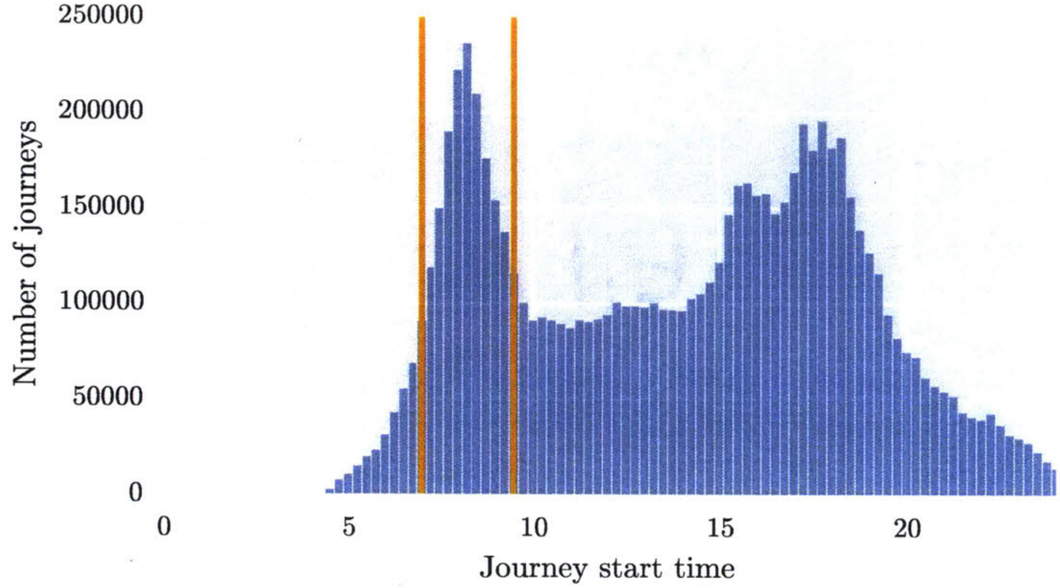

Figure 7-8: Histogram of journey start times

Figure 7-8 shows journey start times, defined as the boarding or tap-in time for the first stage. The period from 7:00 to 9:30 AM is used to define the AM peak in this analysis. In the 10 days analyzed, approximately 20% of journeys began in this period.

Exclusion of Central London

As in many cities, the center of London is distinct from the rest of the city in several respects. Here, Central London is defined as Fare Zone 1 in the Transport for London network. Figure 7-9 shows the zones defined in Step 1 of this analysis that correspond to TfL's Fare Zone 1 highlighted in yellow. Of the 1,000 zones defined in Step 1 of the methodology, 50 are in Fare Zone 1. Central London is the dominant employment center and the destination of a large number of AM peak journeys. Central London has a high density of bus stations and bus routes and is also very well-served by the London Underground. Because it is well-served and because street space is highly constrained, there are few opportunities for new bus services in Central London.

In order to focus on the peripheral region where there are likely to be more opportunities for new services, OD pairs that originate or end in Fare Zone 1 are excluded from the analysis. This exclusions leaves 950 zones and 902,500 zonal OD pairs. Many have no current journeys. 243,076 OD pairs have at least one journey in the 10 days of AM peak data analyzed.

Minimum journey constraint

As shown in Figure 7-10 many OD pairs have few journeys over the 10 AM peak periods analyzed. The figure uses a log scale to reflect the large variation in journeys per OD pair. The x-axis in the figure is also truncated. The maximum number of journeys on a single

Figure 7-9: Map highlighting Central London zones

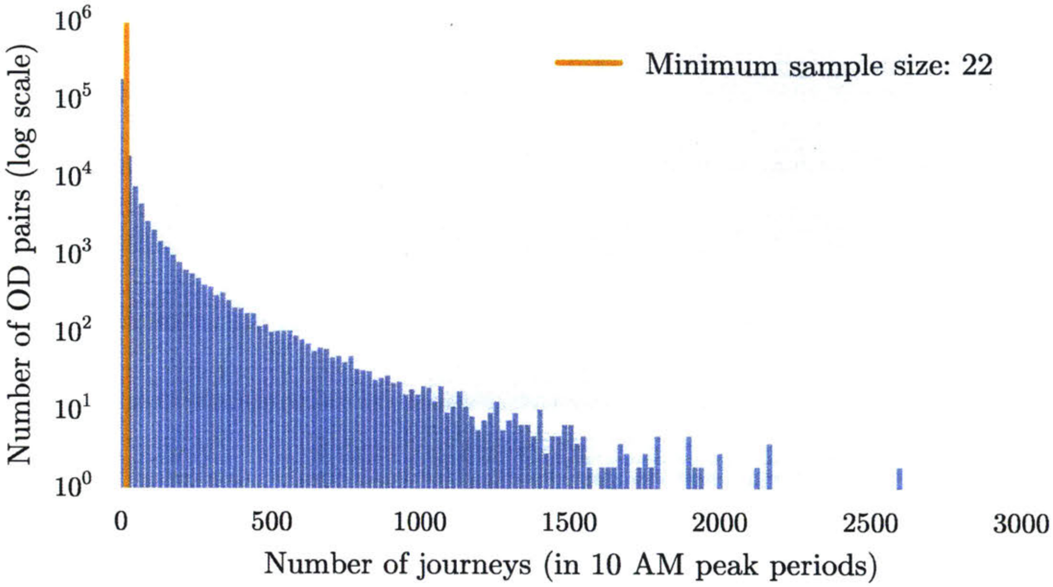

Figure 7-10: Distribution of journeys per OD pair

OD pair is 6,167, which is approximately 247 journeys per AM peak hour.

The minimum sample size, given by Equation 5.1, described in Chapter 5, is applied to exclude OD pairs that do not have sufficient journeys to reliably estimate travel times.

The population standard deviation is unknown and may vary by OD pair, depending on factors such as number of paths available, and congestion. As an estimate, the average sample standard deviation for OD pairs with a large number of journeys (more than 50), is used. For these OD pairs, the average sample standard deviation is 4.9 minutes. The allowable error is set as 2 minutes of deviation from the mean estimate. This yields a minimum sample size of 23 journeys over the 10 AM peaks analyzed, or just under 1 journey per AM peak hour.

Only 45,916 OD pairs have at least 23 journeys, further emphasizing the sparsity of the OD matrix. However, the impact of this constraint on the percent of journeys included in the analysis is small, removing just 16% of all journeys.

Exclusion of intrazonal travel

Some journeys start and end in the same zone. Of the OD pairs that meet the sample size constraint, 762 (about 2%) have the same origin and destination zone. This excludes an additional 4% of the peripheral AM peak journeys.

Effects of constraints on inputs

Figure 7-11 displays the distribution of AM Peak journeys by mode in London. Underground, Overground, and National Rail are grouped as rail for this analysis. In the Underground, passengers can interchange between lines without tapping their card. Therefore, the path an individual takes through the rail network is not always known. If an individual's starting

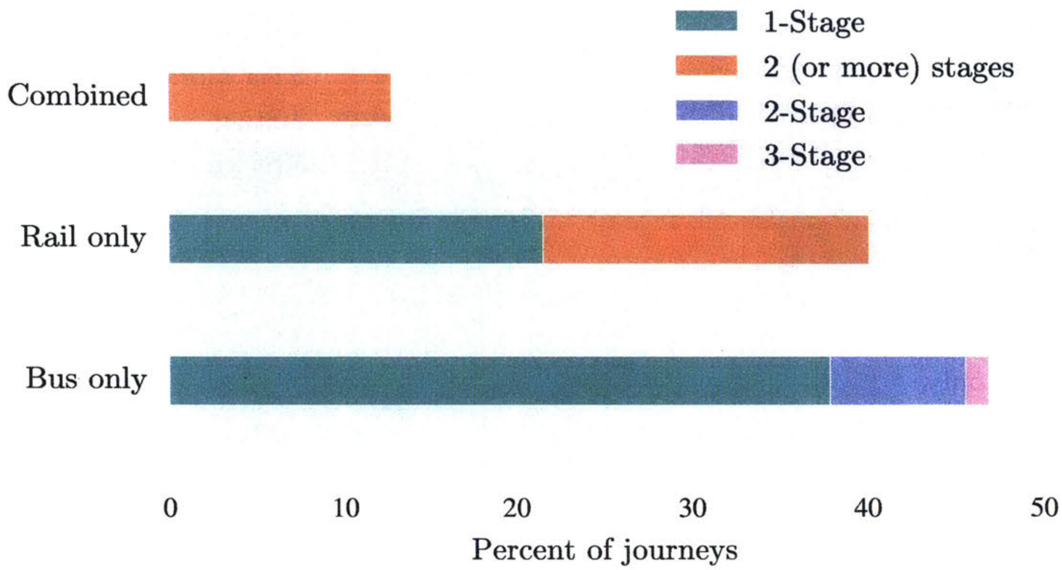

Figure 7-11: All AM peak journeys by mode and stages

and ending station are served by the same line (i.e. the Victoria line), it is assumed to be a single-stage. Otherwise the number of stages is designated as two or more. Passengers must tap their Oyster card at the beginning of each bus stage so bus journeys can be accurately characterized as one, two, or three stages.

Eliminating journeys originating or ending in Central London, and enforcing the minimum journey and minimum distance constraints results in a higher percentage of bus journeys on the OD pairs analyzed. Figure 7-12 shows the distribution of journeys and stages for the OD pairs analyzed. Given the heavily radial nature of London's Underground, combined and rail journeys are more likely to start or end in Central London.

7.3.2 Filtering OD pairs that are well-served

OD pairs that are already well-served by the existing network are filtered out using the decision tree described in Chapter 5. Out of 45,154 OD pairs, 17,942 OD pairs are not served by single-stage rail or single-stage bus.

Of the remaining OD pairs, 19,752 are not served by single-stage rail but are served by single-stage bus. These OD pairs are further evaluated. OD pairs that are served by single-stage bus can be improved if a potential new service can provide more direct service than the existing service. Determining if this is the case requires defining the current bus distance and target bus distance for each OD pair.

Defining current bus distance

The current bus distance should reflect the distance on the shortest path for an OD pair using the existing bus routes. This analysis is limited to those OD pairs served by single-stage bus, as OD pairs that are served only by multi-stage alternatives have already been

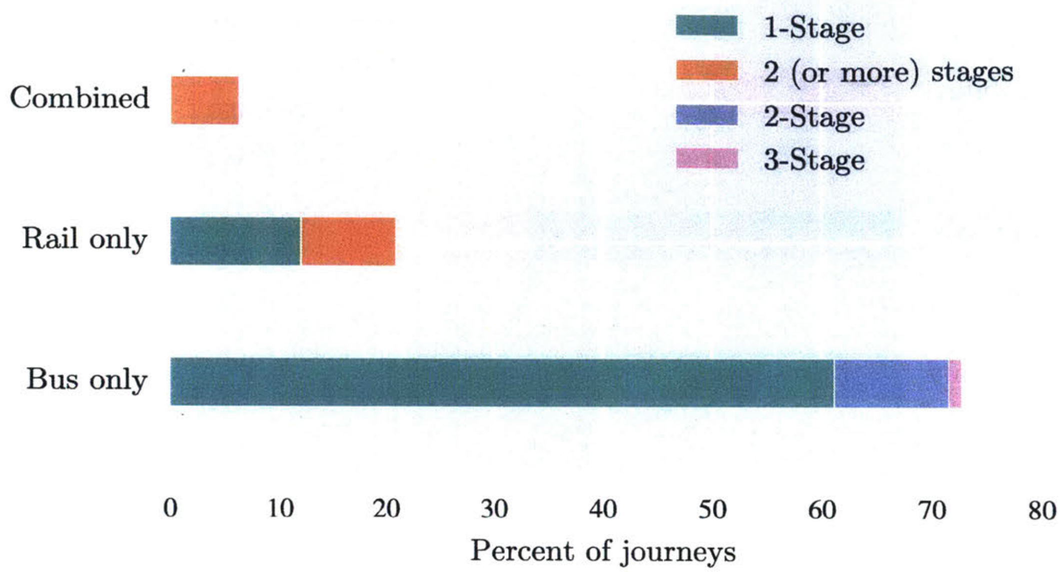

Figure 7-12: AM peak journeys by mode and stages, excluding Central London and including sample size and distance constraints

filtered out. An OD pair may be served by multiple bus routes providing single-stage service. Additionally, each bus route may have more than one bus stop in the origin (destination) zone resulting in variation in in-vehicle distance for a single route.

In order to account for both factors, the current bus distance for the London network is estimated as the minimum of the median in-vehicle distances for each bus route serving the OD pair. The median in-vehicle distance for each route is determined based on all journeys on the OD pair using a bus route, and therefore is influenced by the distribution of journeys within the origin (destination) zone, if journeys are well-distributed it will tend to select the in-vehicle distance for the bus stops closest to the centroid of the zone for each route. Using the minimum of the median values for each route selects the in-vehicle distance on the existing route that provides the most direct service.

One possible improvement over this method would be to measure the distance on each route serving a zonal OD pair from the bus stop closest to the origin (destination) zone centroid, ignoring other stops in the zones. Once the most direct (shortest) route is selected based on these distances, the centroid-adjacent bus stops on the direct route are established as the origin (destination) points for the zonal OD pair. Then, the shortest path distance can be queried based on these points, ensuring the shortest path and current bus distances are directly comparable and eliminating within zone variation in the two distances.

Defining target distance

The shortest path between two points can be queried using the Google Maps Distance Matrix API (see Section 7.1.2). The origin (destination) point is the bus stop serving the OD pair that is closest to the zone centroid. As discussed, this could be improved by selecting the specific bus stop that is on the route that provides the current shortest path.

Because the actual path of a potential new bus route is not expected to follow the shortest path exactly, the shortest path estimates will be adjusted to define the target distance using Equation 5.2 described in Chapter 5. Comparing the current distance and shortest path distance for OD pairs served by single-stage bus in London's existing network can inform appropriate choices of m and a.

Figure 7-13 shows the distribution of the ratio of the current in-vehicle distance estimate to the shortest path estimate for OD pairs served by single-stage bus. This directness ratio is shown in conjunction with the distribution of the shortest path distance for the OD pairs. The ratio of current in-vehicle distance to the shortest path distance is larger and more variable for short trips. The variation in ratio decreases with distance. This suggests that using a single ratio to scale up shortest path distance to expected distance will be inaccurate.

In part, the difference in ratios for shorter and longer OD pairs is because even a small deviation from the shortest path represents a significant percentage increase for short pairs. Secondly, despite efforts to select in-vehicle distances and shortest path distances for trips originating (ending) near the centroid of the zone, some in-zone variation persists. If existing bus demand is unbalanced, skewed towards one side of the zone, this can skew in-vehicle distances relative to the shortest path distance. This effect may be especially significant in zones that have barriers such as railroad tracks or highways that impact the distance required to reach different points within the zone.

There are examples of short OD pairs where the distance ratio exceeds 1.1 even though direct service is offered along the shortest path. This is because demand is skewed toward longer trips within the zonal OD pair. Figure 7-14 shows one such example. The buses travel on the shortest path available. However, the estimated ratio of bus distance to shortest path distance is 1.1, because the origin and destinations points used to estimate the current distance and shortest path distance do not match.

Variation in origin (destination) points within the zone also accounts for some OD pairs for which the current bus distance is less than the estimated shortest path distance. In other cases this is the result of fast but indirect paths. The Google Maps Distance Matrix API seeks the path with the shortest travel time, not the shortest distance. Therefore it sometimes selects a path that is indirect but avoids congestion.

Overall, it is expected that the effects of unbalanced demand on estimated current bus distance will balance out over the large set of OD pairs analyzed. For approximately 25% of OD pairs, the ratio is less than one. The median ratio is 1.1 and the 75th percentile is 1.3. This is a relatively narrow range, and the ratios are quite close to 1, indicating that overall the OD pairs served by single-stage bus provide quite direct service.

An alternative way to compare the current bus distance to the shortest path distance is to take the difference. Figure 7-15 shows the distribution of the difference in current bus distance and the shortest path distance with the distribution of shortest path distance for the same OD pairs. The distribution of the difference in distances is tightest for short OD pairs. Surprisingly, it does not continue to spread for the longest OD pairs analyzed. Instead, the distribution narrows for OD pairs longer than 4 miles. This may reflect the fact that for long OD pairs, only those with direct bus service have sufficient existing demand to meet the sample size constraint. The median difference in distance is 0.2 miles. The 75th percentile is 0.5 miles and the 90th percentile is 1.1 miles.

Given the distributions in Figures 7-13 and 7-15, selecting a large (conservative) value for the multiplier term, m, will identify mostly short OD pairs for improvement. These pairs are more subject to intrazonal variation effects, and may only have potential for moderate improvement. Therefore, only moderate values of m from 1.05 and 1.1 are considered.

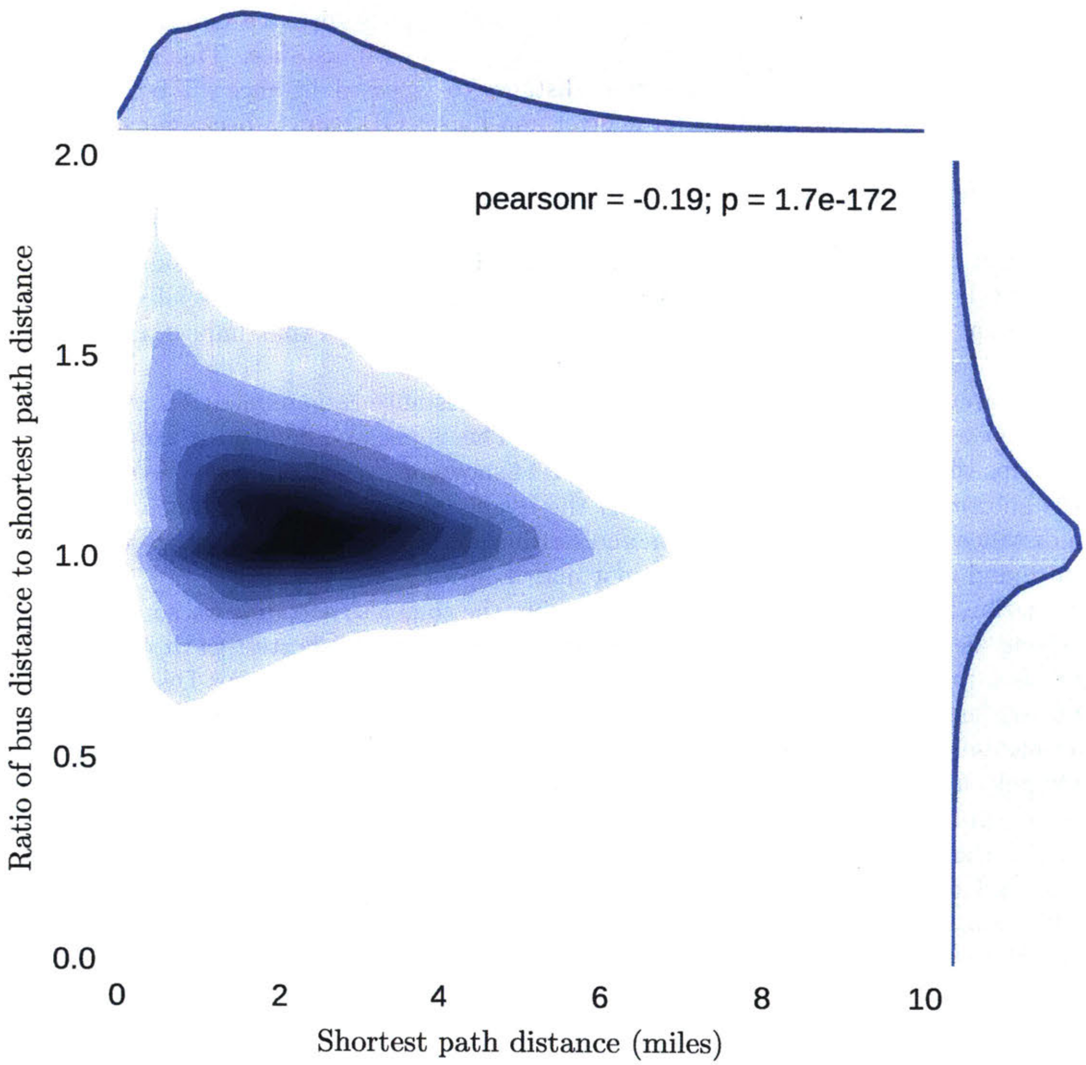

Figure 7-13: Joint distribution of distance and bus directness ratio

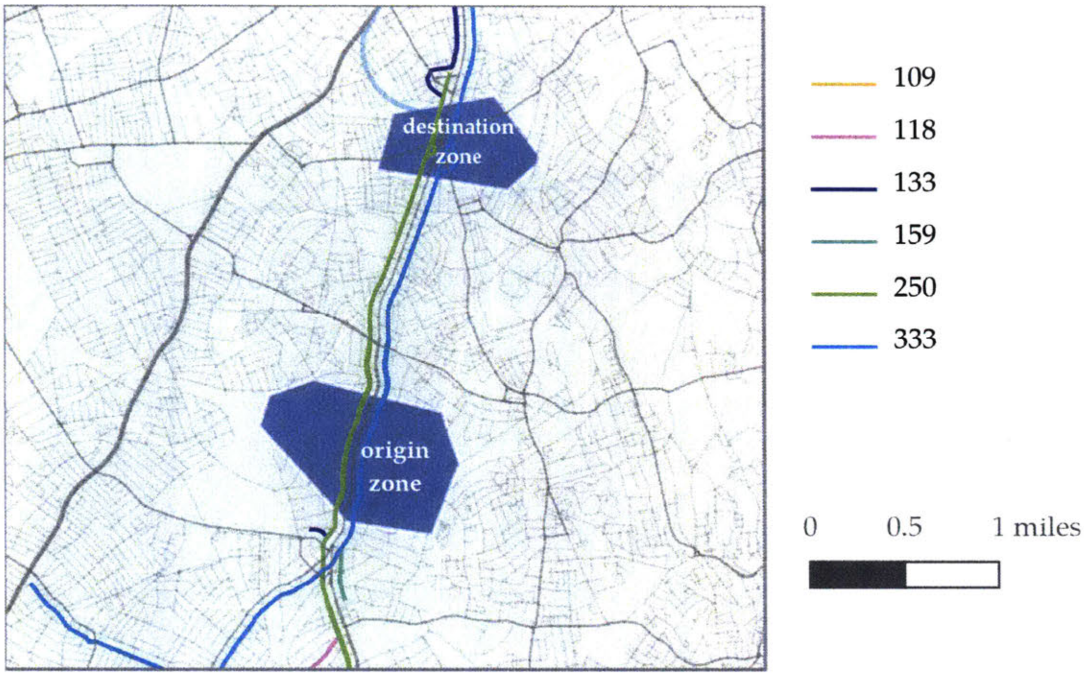

Figure 7-14: An OD pair that is misrepresented by the bus directness ratio

Adding a fixed amount, a, to the shortest path distance identifies a less-biased set of OD lengths. A conservative value for a will eliminate very short OD pairs, of less than one mile.

For these short journeys, access and egress stages typically make up a significant portion of the journey. In addition, short distance OD pairs are walkable for many (though not all) individuals. As a result, short distance OD pairs may have different demand behavior, responding differently to new services compared to longer OD pairs. Finally, short distance OD pairs are unlikely to be improved substantially by structural improvements, particularly the high level network structure changes that are the focus of this analysis. Therefore, eliminating short OD pairs is desirable.

Because OD pairs in London's network served by single-stage bus appear to be overall well-served, conservative choices for a are applied to ensure that only OD pairs that can be improved are selected. Values of 0.6 miles and 0.9 miles are considered. There are 19,762 OD pairs served by single-stage bus. Those that exceed the target distance are identified as OD pairs that are circuitous and not well-served, according to the definition in Chapter 5. Table 7.2 shows the percent of OD pairs served by single-stage bus that are identified as candidates for improvement with different values of a and m. Applying the most conservative values for a and m (1.1 and 0.9 miles) identifies 1,737 OD pairs that are served by single-stage bus for improvement.

The filtering process identifies the set of OD pairs that are not well-served by existing services. However, some multi-stage or circuitous OD pairs may have good performance (in terms of travel time) despite the fact that they are not well-served. The next part of the OD-level analysis further refines the set of OD pairs that are candidates for improvement based on expected travel time savings.

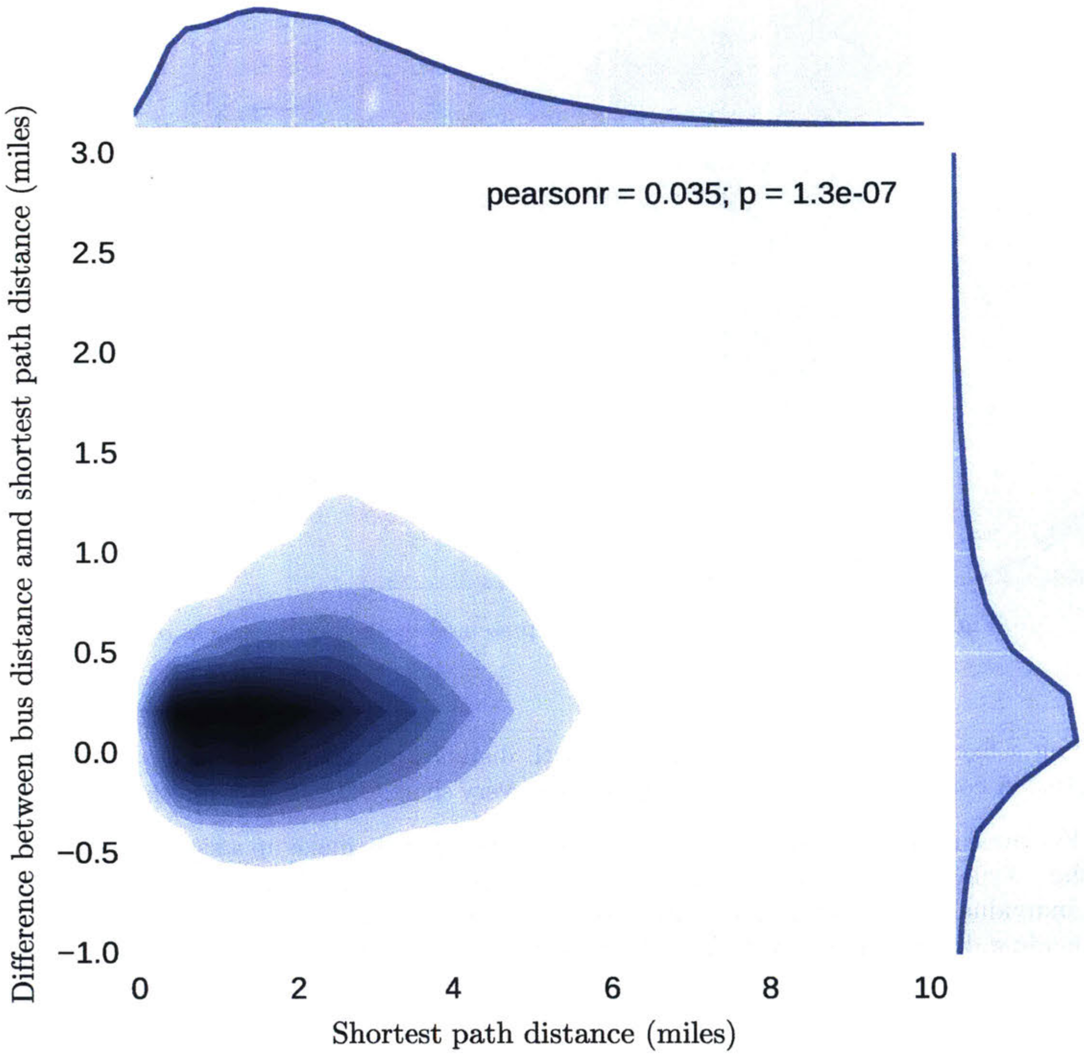

Figure 7-15: Joint distribution of distance and the difference between bus distance and shortest path distance

Table 7.2: OD pairs classified as circuitous for different m and a

m	a	Circuitous OD pairs
1.1	0.9	9%
1.05	0.9	11%
1.1	0.6	13%
1.05	0.6	17%

7.3.3 Estimating expected travel time

The expected travel time is estimated using Equation 5.3 from Chapter 5. The shortest path travel time, t_{short}, is queried from the Google Maps Distance Matrix API, which accounts for expected congestion. Transport for London states that planners aim for a stop spacing of approximately 400 meters (0.25 miles) (Bus Priority Team, Transport for London, 2006), which is used as $d_{interstop}$.

Dwell time is highly variable depending on demand as well as boarding technology. Often it is estimated using a fixed stopping penalty and an additional penalty per boarding and alighting passenger. Levinson (1983) suggested 5 seconds plus 2.75 seconds per boarding or alighting passenger is reasonable in most communities. York (1993) estimated dwell time for nine routes in London and Exeter, UK. He estimated a dead time of 2.38 to 8.26 seconds per stop. In addition, he found dwell time increased by 0.99 to 2.94 seconds for each alighting passenger and 1.84 to 5.49 seconds for each boarding passenger paying with a pass. Each passenger boarding paying cash added 2.74 to 8.87 seconds.

In addition to dwell time, there is added travel time due to the acceleration and deceleration of the bus. Levinson (1983) reported acceleration and deceleration time to range from 11 to 23 seconds per stop, depending on stop spacing and other factors. Robinson (2013) estimated acceleration and deceleration times based on AVL data for Transport for London's Route 45. He found the average time was 11.6 seconds, but with significant variability. The 10th percentile was 3.0 seconds while the 90th percentile was 41.8 seconds. In short, the time required to make a stop is highly variable. Some stops will be skipped entirely if there are no boarding or alighting passengers. As a result, some stops will not add any time and others may add more than a minute.

The demand on any new route is unknown, but assuming that each bus trip serves 100 journeys, and has approximately 36 stops, there will be an average of approximately 3 boardings and 3 alightings per stop. Using the middle of the ranges proposed by York (1993), and the average acceleration/deceleration time found by Robinson (2013), a stop including 3 boardings and 3 alightings would add approximately 35 seconds. Allowing for an occasional stop (approximately 1 in 16) to be skipped, this value is reduced to 33 seconds. Using a value of 1.1 for m' produces values similar to the median travel time for OD pairs with existing single-stage service.

7.3.4 Identifying improvable OD pairs

Any OD pairs with current travel times less than the expected travel time are excluded from the corridor aggregation in Step 3. These OD pairs are not expected to be amenable to improvement through new services.

Of the 1,737 OD pairs served by single-stage bus identified as not well-served, 1,621 have travel times greater than the expected travel time for the OD pair. This is reasonable: most OD pairs served by circuitous bus service are improvable (in terms of travel time). For OD pairs served only by multi-stage service, the travel time filter has a much greater impact. Of 17,942 OD pairs served only by multi-stage service, 5,856 OD pairs have travel times that exceed the pair's expected travel time. This is also unsurprising, as many of these OD pairs are served by multi-stage rail, which is often faster than bus.

7.3.5 Estimating potential travel time savings

The potential travel time savings for an OD pair is defined as the current travel time less the expected travel time for the OD pair multiplied by the number of existing journeys expected to switch to the new service. Chapter 5 discussed alternative approaches for estimating the existing public transport journeys that are expected to shift to a new service and proposed developing and applying a within-public transport assignment model. Here, a simpler approach is taken: all current public transport journeys on an OD pair identified as improvable are expected to shift to the potential new service. This method may exaggerate expected benefits in some cases, but is deemed sufficient for sketch planning. Estimates of expected benefits can be refined in post-analysis.

7.3.6 Estimating expected demand

OD level demand is an input to the corridor identification algorithm in Step 3. This includes current demand and expected new demand. As discussed, in this application current demand is defined as all existing public transport demand on OD pairs identified as improvable. Estimating expected new demand is a topic that is the subject of a large body of research.

Some methods attempt to predict corridor level mode share. To estimate such a model, OD-level mode share and demand estimates are required. Transport for London's London Transportation Studies (LTS) Model uses zones of similar size to those defined in Step 1 of this analysis, and estimates zone-to-zone demand by mode. However, the estimates of public transport demand are not consistent with the demand observed based on Oyster transactions. Figure 7-16 shows the LTS estimates of zone-to-zone public transport demand compared to the ODX-processed Oyster demand assigned to the same LTS zonal pairs. LTS estimates are the average hourly demand for the AM peak period, with the peak period defined as 7:00 to 10:00 AM, and ODX estimates are the average hourly demand based on the 10 AM peaks analyzed, with AM peak defined as 7:00 to 9:30 AM.

There is little correlation between the public transport demand estimated by LTS and the demand inferred from Oyster transactions using ODX. This misalignment raises questions about the reliability of LTS model demand estimates at this microscopic level. While demand data was not available to compare to the LTS model estimates for other modes, the mismatch suggests that using LTS model mode share data to estimate a corridor mode share model is likely to be futile.

An alternate method is to use ODX-inferred public transport demand estimates and zonal OD pair level explanatory variables to estimate a model that predicts public transport demand directly. Such "direct demand" models are often used at the sketch planning level, though most commonly for intercity travel (Domencich and McFadden, 1975).

Given available data, a direct demand model was estimated using a log-log form to predict OD-level bus journeys based on bus travel time, income, zonal population and employment, car ownership, car travel time, and zonal bus stop density. The following variables were used in the best-fitting model, which is summarized in Table 7.3:

β_T bus travel time

β_I income

β_{PT} origin zone population * destination zone employment

β_S bus stop density in origin zone

β_E bus stop density in destination zone

All variables have the expected signs. However, the coefficient for bus travel time (-1.0),

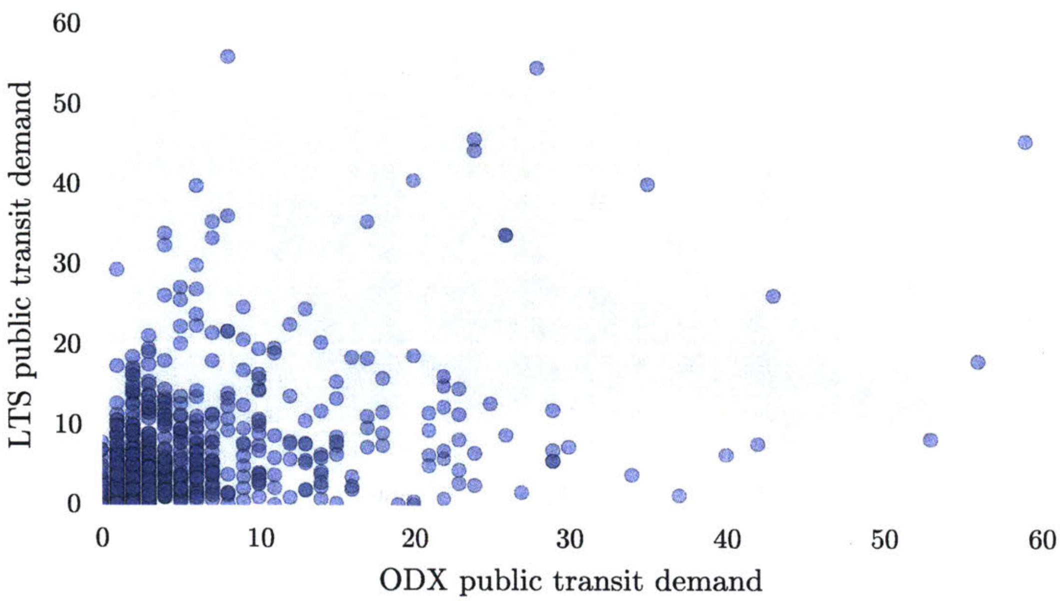

Figure 7-16: Comparison of ODX-inferred and LTS-modeled public transport demand

Table 7.3: Direct demand model regression results

Dep. Variable:	y	**R-squared:**	0.378
Model:	OLS	**Adj. R-squared:**	0.373
Method:	Least Squares	**F-statistic:**	77.39
Prob (F-statistic):	2.23e-63	**Log-Likelihood:**	-684.87
No. Observations:	643	**AIC:**	1382.
Df Residuals:	637	**BIC:**	1409.
Df Model:	5		

	coef	**std err**	**t**	**P> |t|**	**[95.0% Conf. Int.]**
β_T	-1.0076	0.068	-14.750	0.000	-1.142 -0.873
β_I	-0.2353	0.152	-1.545	0.123	-0.535 0.064
β_{PT}	0.2923	0.038	7.767	0.000	0.218 0.366
β_S	0.0959	0.065	1.485	0.138	-0.031 0.223
β_E	0.4301	0.068	6.323	0.000	0.297 0.564
C	9.5065	1.408	6.753	0.000	6.742 12.271

Omnibus:	39.127	**Durbin-Watson:**	1.308
Prob(Omnibus):	0.000	**Jarque-Bera (JB):**	46.377
Skew:	-0.577	**Prob(JB):**	8.50e-11
Kurtosis:	3.633	**Cond. No.**	1.19e+03

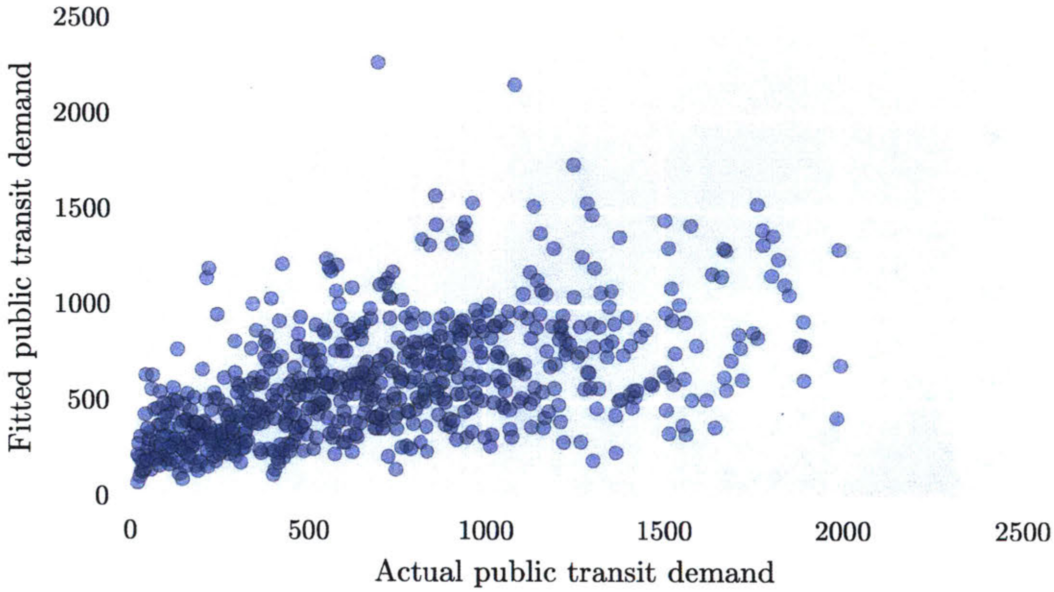

Figure 7-17: Actual demand compared to fitted demand

which can be interpreted as the elasticity of demand with respect to travel time, is greater than would be expected based on the literature. Furthermore, the direct demand model estimates deviate significantly from the true values for the data on which it was estimated (See Figure 7-17).

Given the problems with the direct demand model, an alternative is to use simple estimates of elasticity of demand with respect to bus in-vehicle travel time from the literature. In a survey of UK-based studies Wardman (2012) states that the best estimates range from -0.4 to -0.6. Elasticity of demand with respect to in-vehicle time is expected to vary based on many factors including trip length. For shorter trips, in-vehicle time makes up a smaller percentage of the total journey time, assuming access, egress, and waiting time are uncorrelated with in-vehicle time. As a result, the demand response to in-vehicle time changes for shorter journeys is expected to be smaller. Figure 7-18 shows in-vehicle time as a fraction of total journey time for different in-vehicle time values, based on the assumption that access, egress, and waiting time account for a total of 15 minutes. In fact, the inflection points are similarly located for values of 5 to 25 minutes of access, egress, and waiting time. The vertical lines mark inflection points used to assign different elasticities to OD pairs based on in-vehicle time.

The elasticity values of -0.4 (for in-vehicle times up to 20 minutes) to -0.6 (for in-vehicle times of 40 minutes or more) from the literature will produce conservative estimates of new demand. The estimated direct demand model, discussed previously, suggests that more aggressive values up to -1.0 may be appropriate. The sensitivity of results to different elasticity values is explored in Section 7.5.

Elasticity is a simple tool for estimating new demand that can capture responses to time savings. However, demand is complex. Ridership on a new service can include journeys that shift from existing services, journeys that shift from other modes, and entirely new journeys.

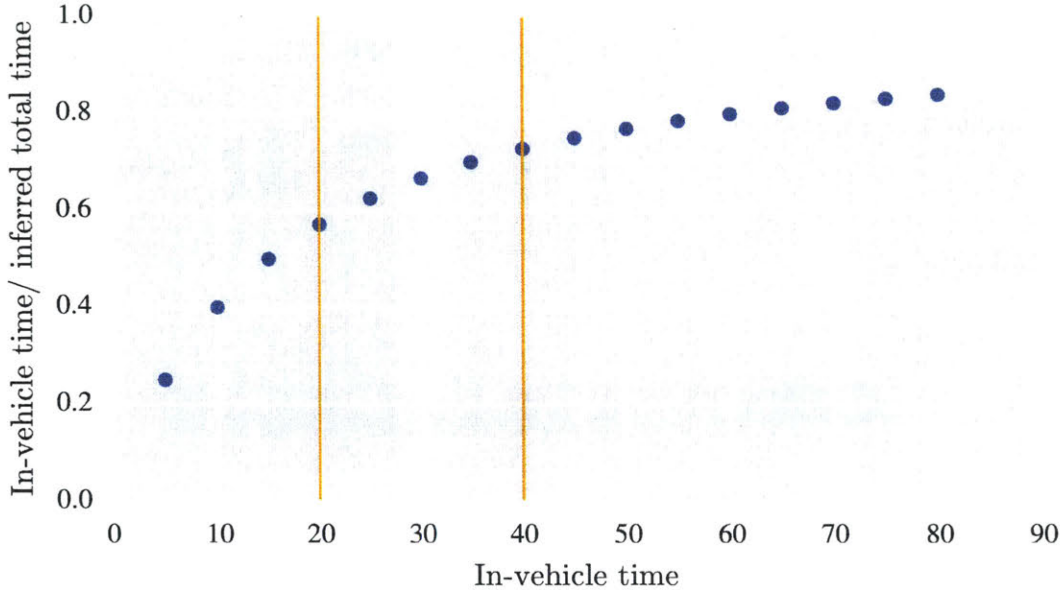

Figure 7-18: In-vehicle time as a portion of total trip time

New journeys, in particular, may be induced over longer time horizons. The demand response to travel time savings is expected to vary depending on characteristics of individuals and of land uses in origin and destination zones. Characteristics of the journey, such as the number of stages, are also expected to influence demand. Developing and applying a model that can account for these nuances should produce more accurate estimates of new demand, leading to improved results from the overall methodology.

In addition to changes in travel time, there are many external factors that change in cities over time. In London, there has been sustained population growth and accompanying land use changes. In order to use this framework to best make recommendations to serve future demand, it is important to consider the effects of these changes on travel demand patterns.

In the application presented here, new demand is estimated based on in-vehicle time elasticities, as estimating demand changes due to other factors is beyond the scope of this research. However, other information about demand changes could easily be incorporated by adjusting the expected OD zonal demand.

7.3.7 Summary of OD pairs identified for improvement

Conservative values for m (1.1) and a (0.9 miles) are applied to ensure that the OD pairs identified are likely to be amenable to improvements. This may come at a cost of missing some OD pairs that could have moderate improvements. However, this ensures that the focus of new bus routes is on OD pairs with significant potential for improvement.

Of 45,154 OD pairs that satisfy the sample size requirement, start and end outside Central London, and are interzonal, 7,477 (about 17%) are identified as candidates for improvement. There are 435,457 current public transport journeys on these OD pairs in the 10 AM peak periods analyzed, or about 17,418 journeys per AM peak hour. Of all the

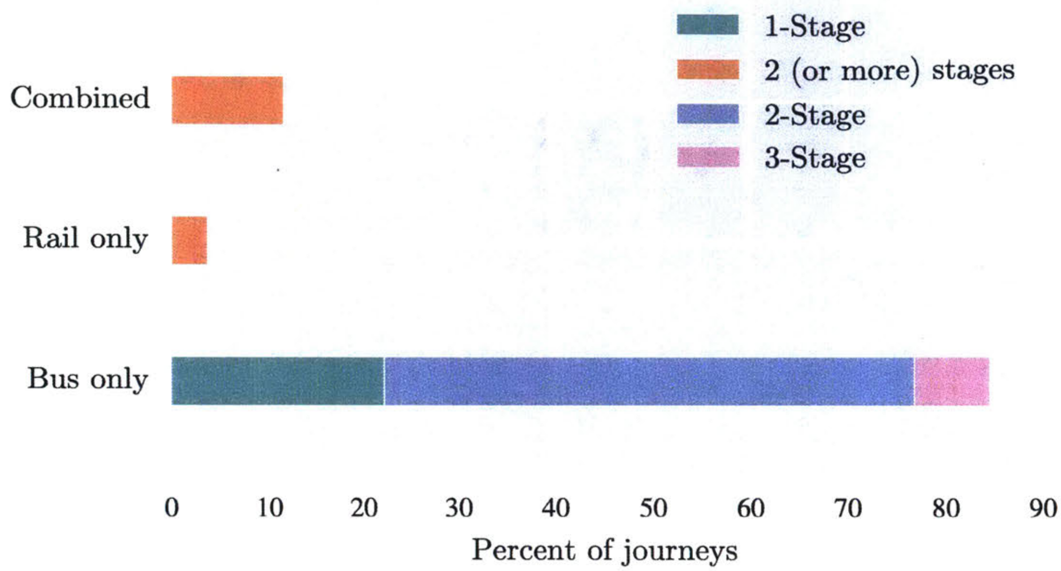

Figure 7-19: Modes and stages of journeys on OD pairs identified for improvement

journeys on the 45,154 OD pairs analyzed, just 7% are on OD pairs that are candidates for improvement. This is expected, as the poorly served OD pairs likely draw less demand than well-served pairs. Using elasticities of -0.4 to -0.6, a total of 2,235 new journeys per AM peak hour would be expected on these OD pairs, if they were all improved.

Figure 7-19 shows the modes and stages for journeys on the OD pairs identified as candidates for improvement. About 63% are multi-stage bus journeys. Just 3% are multi-stage rail and about 22% are single-stage bus. The remaining journeys use a combination of bus and rail.

7.4 Identifying corridors (Step 3)

In order to identify corridors, the corridor identification algorithm uses five parameters. The first three define the shape of a corridor, which should match the expected shape of the service area for a new bus route. These parameters are the maximum distance, the corridor length, and the maximum angular. An additional parameter specifies the minimum demand, measured as average flow, required for a viable corridor. Finally, an optional set of parameters can specify the weights that are assigned to different benefits in prioritizing corridors. Characteristics specific to a given property can inform maximum distance, length, and minimum flow, as discussed in the following sections.

7.4.1 Maximum distance

In the implementation of the corridor identification methodology, potential routes are represented as straight lines. The maximum distance parameter constrains the euclidean distance from OD pairs assigned to a potential corridor to this line. It is measured from the centroid of the origin (destination) zone. The maximum distance specified should correspond to a reasonable access distance to the new service and can be informed by information about the distribution of access distances. The distribution of access distance shown (previously) in Figure 7-1, used to determine zone size, is also helpful in defining maximum distance. The median access distance is 0.2 miles, the 75th percentile access distance is 0.3 miles, and the 90th percentile is 0.5 miles. This suggests that the maximum distance should be at most about 0.5 miles from a potential route.

In Step 1 of the framework, zone size is set to be walkable, with an average radius of approximately 0.45 miles. A corridor that is approximately one zone wide is therefore desirable. In the corridor identification methodology implemented, distance is measured from the potential route to the zone centroid. Therefore, the edge of a zone may be farther than the maximum distance from the potential route. Setting a maximum distance parameter of more than 0.45 miles is very likely to produce corridors that are more than one zone wide. However, as is discussed in more detail in Section 7.5.3, even maximum distances of less than 0.45 miles produce corridors that are more than one zone wide in some places, due to variation in zone size. On the other hand, reducing the maximum distance significantly below 0.45 quickly eliminates corridors, as it requires zone centroids to be linearly aligned.

7.4.2 Corridor length

Corridors identified represent potential areas for new bus services, and therefore should be of an appropriate length to be served by a single route. Different cities have bus routes of varying lengths, reflecting preferences about network structure, route operation, geographic characteristics, and demand patterns. Analyzing the lengths of existing routes can suggest appropriate lengths for new routes.

Figure 7-20 shows the distribution of bus route lengths in Transport for London's bus network. The corridor identification algorithm requires the specification of a minimum and maximum route lengths. Using the 25th to 75th percentile of existing routes suggests a range of 6.6 to 10.6 miles. More flexibly, the 10th to 90th percentiles are 5.1 miles and 13.2 miles respectively.

For use in the algorithm, these values must be adjusted. The algorithm measures length as the straight line connecting the ends of the corridor. However, a bus route is constrained to the road network and may be further constrained by other factors as discussed in Section 7.3.2. Therefore a corridor with a straight line length of 5 miles may result in a somewhat longer bus route. Section 7.5 explores the results of minimum corridor length parameters of 4 to 6 miles (corresponding to actual route lengths of about 5 to 7 miles) and maximum corridor length parameters of 8 to 11 miles (corresponding to actual lengths of 10 to 13 miles).

7.4.3 Corridor demand

The corridor identification algorithm enforces a minimum flow requirement on all corridors identified. The average hourly flow for existing routes can be estimated to inform the selection of this parameter. Figure 7-21 shows the distribution of average AM peak hourly flow for existing routes in one direction. Flows have been scaled up linearly based on route level inference rates to account for stages without inferred destinations. These flows do not include journeys that were not paid for with an Oyster card (an estimated 10% of bus passengers did not use Oyster cards during the analysis period (Gordon et al., 2013)). The

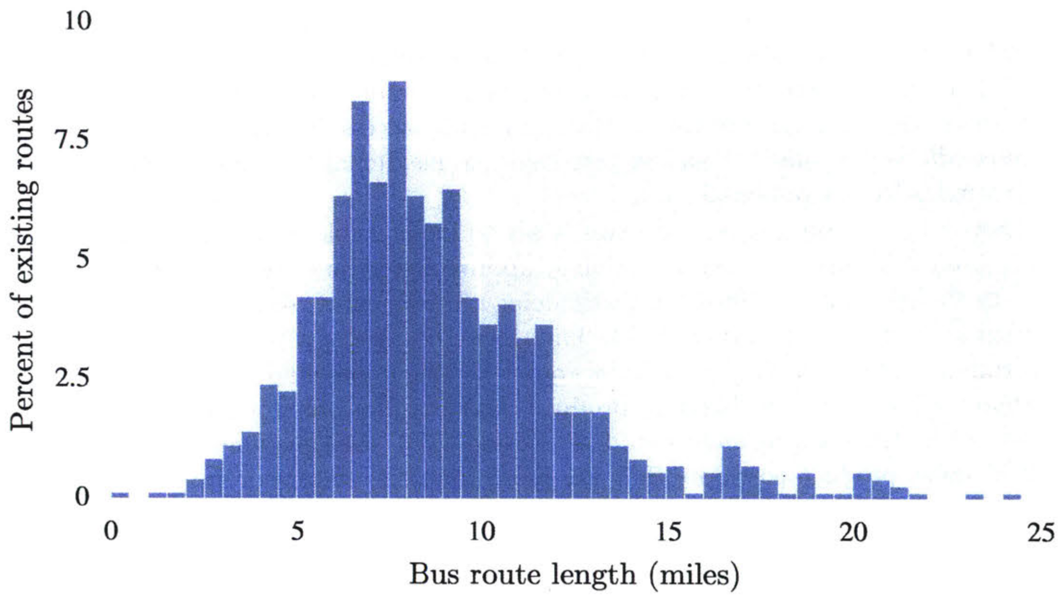

Figure 7-20: Distribution of bus route lengths

25th percentile flow on existing routes is 55 journeys per AM peak hour, and the median is 92.

When corridors are evaluated to determine if they meet the minimum flow parameter, total demand, consisting of current and expected new journeys on OD pairs identified as candidates for improvement are summed. Because this application uses elasticities to estimate new journeys, the estimates may not account for variation in latent demand. In particular, the impacts of population growth and development on demand are not accounted for.

As discussed in Chapter 6, there are some additional expected sources of flow on potential corridors that are not included when corridor flow is estimated within the corridor identification methodology. Demand on OD pairs that are physically located within a potential corridor but are determined to be well-served or have good performance is excluded. The assumption that none of this demand will shift to a new service is likely too extreme.

In this application, only 9% of OD pairs served by single-stage bus were identified as circuitous. This conservative choice means that some of the OD pairs defined as well-served may actually see modest improvements from a new bus route. Secondly, some demand is expected to shift even if a new service provides similar quality of service to existing routes.

In addition, the estimates of new journeys based on elasticities may underestimate induced demand, as the new bus route will do more than improve in-vehicle time for new journeys. It will also increase capacity, and may also reduce access and egress times for some journeys, depending on the final route alignment and stop locations.

Also excluded from the flow estimate is intrazonal demand. As noted in Section 7.3.1 approximately 4% of AM peak peripheral public transport demand is intra-zonal. The additive factor in the distance threshold estimation also excludes many short OD pairs where performance and potential for improvement is more difficult to estimate. 25% of public transport journeys in the analysis area are shorter than 1 mile. For short OD pairs,

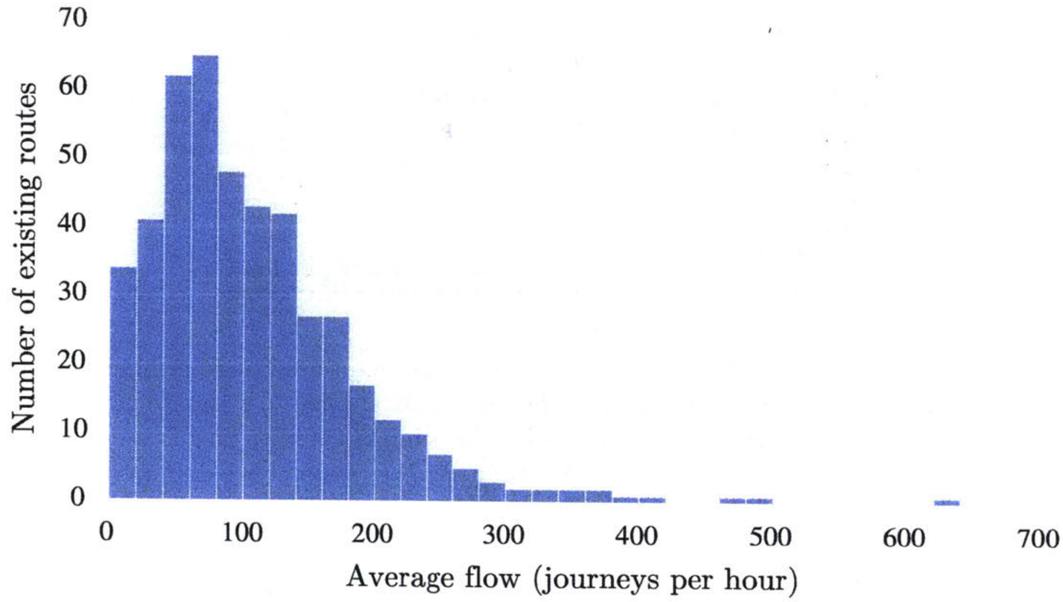

Figure 7-21: Distribution of average hourly flow for existing routes

there are also likely a significant number of walking trips. These trips may shift to bus if it serves the OD pair well. This shift may not be fully accounted for using the elasticity approach to estimate new journeys.

Finally, improvable stages of journeys that were deemed well-served overall have been added to the demand on OD pairs identified for improvement (See Section 5.2.2). However, there may be cases where an OD pair that was identified for improvement was not assigned to any corridor, but some stages of the journeys on this OD pair lie within an identified corridor, and could be improved. In general, the methodology clusters OD pairs served by multi-stage services based on their origin and destination zones, ignoring interchange zones, as the assumption is that more benefits will be realized by serving the full OD pair with single-stage service than by improving just part of the journey. Including multi-stage journeys as well as their individual stages when assessing OD-level demand results in double-counting. Therefore, this additional demand can only be assessed after corridors have been defined.

Due to these unaccounted for sources of demand, the estimated corridor flow likely underestimates actual potential flow, and therefore the minimum flow parameter should not be set too high. Section 7.5 discusses how different minimum flow parameters impact the identification of corridors, and Section 7.6 explores expected additional demand on the corridors identified.

7.5 Sensitivity to parameter selection

The zonal scheme defined in Step 1 and the OD-level analysis in Step 2 determine the inputs to Step 3 of the methodology. Parameter choices in each step impact the set of corridors identified. Here, the sensitivity of the results to seven parameters are evaluated.

Table 7.4: Base scenario parameters

Distance (time) multiplier	1.1
Distance additive	0.9 miles
Elasticities	−0.4 to −0.6
Maximum distance	0.4 miles
Length	5 to 9 miles
Angle	22.5 °
Flow	50 passengers per hour

The parameters explored are: the multiplier and additive factor used to determine the target distance standard and expected travel time, the elasticities applied to estimate new demand, the maximum distance, length, and angle, and the minimum corridor flow.

A base scenario is defined, using values determined to be reasonable based on the analysis described in Section 7.3 and Section 7.4. These values are summarized in Table 7.4. In this analysis, the same values are used for m, the distance multiplier, and m', the travel time multiplier. The value selected for the minimum flow parameter, 50 journeys per hour, is quite low, particularly because corridors are bidirectional: 25 journeys per hour is approximately the 10th percentile of the unidirectional flow on existing routes. However, as discussed in Section 7.4.3, new bus services are likely to draw additional demand not accounted for in the corridor identification algorithm. This is confirmed in post-analysis discussed in Section 7.6, in which some of the additional sources of demand are quantified.

The sensitivity analysis considers the impact of changes to these parameters on the corridors identified by the algorithm. In the base case, 11 corridors are identified (See Figure 7-22). Each corridor identified is prioritized by the sum of potential journeys-minutes saved (for existing passengers) and new journey-minutes induced.

7.5.1 Distance and time multiplier and additive factors

Changing the multiplier (m and m') and additive (a) factors impacts the definition of the target distance standard and the expected travel time. The target distance standard and the expected travel time estimates dictate which OD pairs are identified as candidates for improvement. The expected travel time estimate also impacts the estimation of travel time savings and hence new journeys. Table 7.2 showed the percentage of OD pairs served by single-stage bus identified as circuitous given different values of m, m' and a.

In the base case, conservative values for m, m', and a were selected, and 7,477 OD pairs were identified for improvement. If m and m' are reduced to 1.05, 8,346 OD pairs are identified for improvement, and 13 corridors are identified. Figure 7-23 shows the corridors identified and their expected benefits. The expected benefits are not strictly comparable to the benefits in the base case, because the reduced expected travel time results in an increase in the expected benefits for each OD pair. The 11 corridors identified in the base scenario are included (in some cases with extensions) along with two additional corridors in Southwest London. The order of priority remains similar.

In the scenario in which m and m' are 1.1, but a is reduced to 0.6 miles, 8,278 OD pairs are identified for improvement. The 16 corridors identified are shown in Figure 7-24. In addition to the 11 corridors identified in the base case, 5 new corridors are identified in Southwest London. It is interesting that reducing either m (m') or a identifies additional

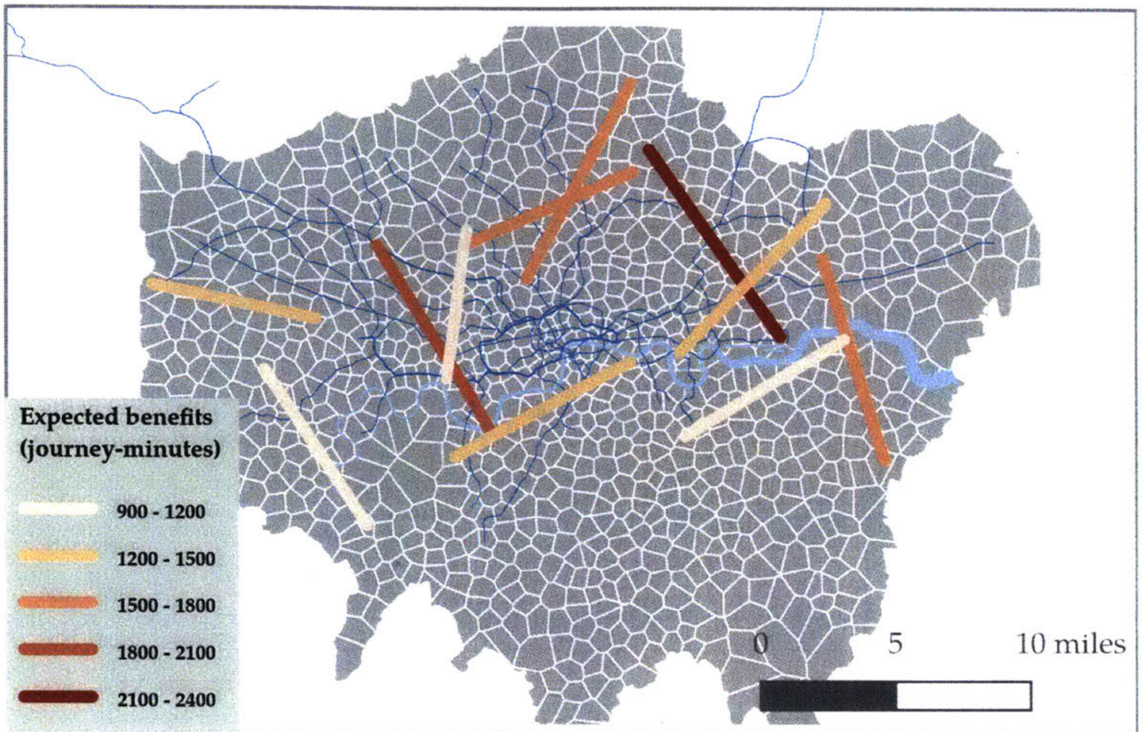

Figure 7-22: Corridors identified in the base case

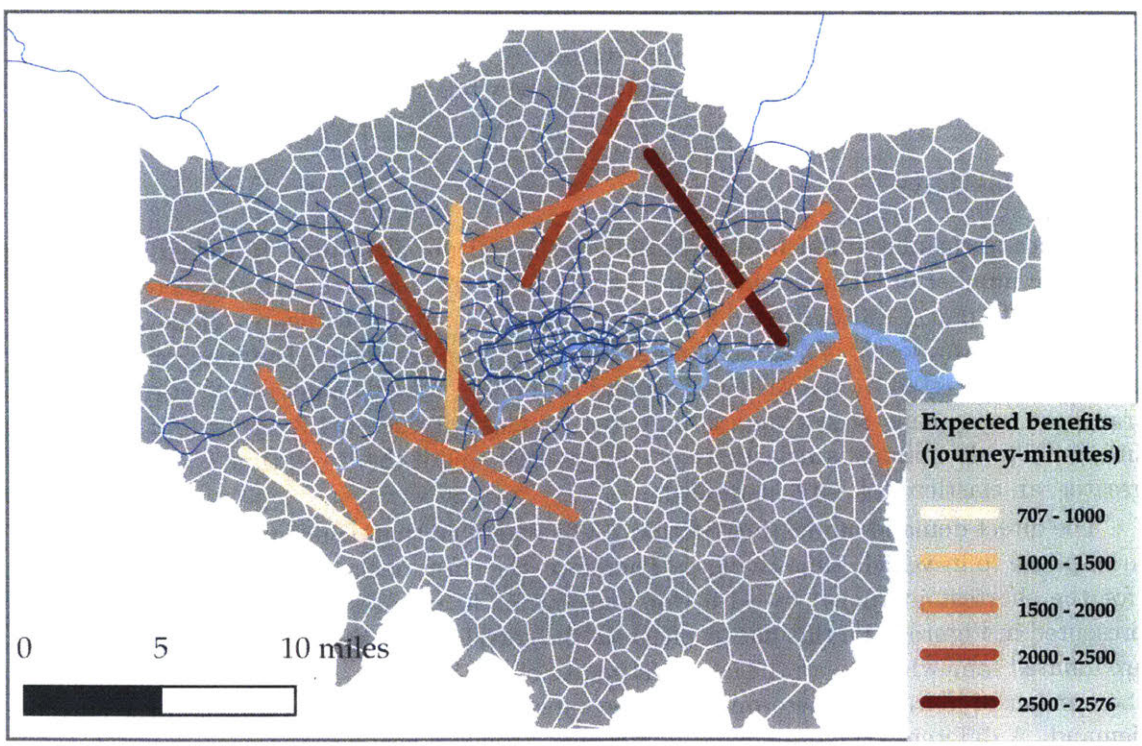

Figure 7-23: Corridors identified with m and m' reduced to 1.05

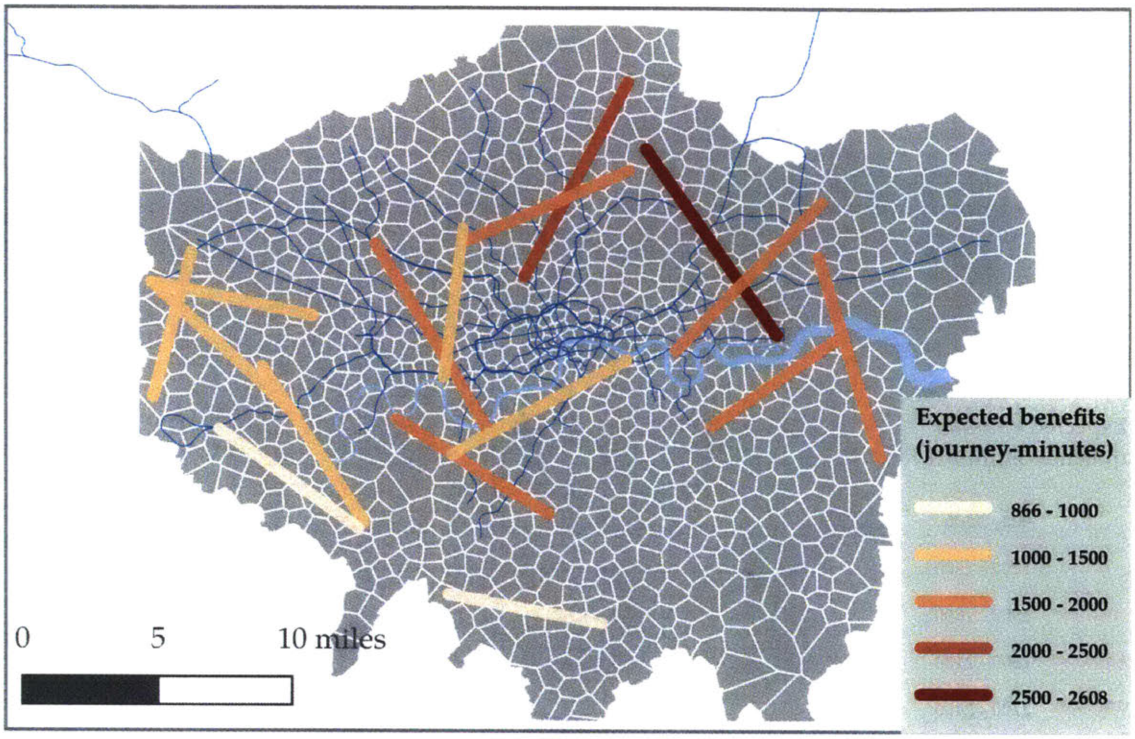

Figure 7-24: Corridors identified with a reduced to 0.6 miles

corridors in Southwest London, exclusively. This suggests that the area does not have a high concentration of OD pairs that are very poorly served, but has OD pairs that may see moderate improvements through the addition of new bus services.

7.5.2 Elasticity

The choice of elasticity influences the inputs to the corridor algorithm. However, it does not affect the number of OD pairs identified for improvement, but rather the expected demand on each OD pair. Elasticities closer to zero produce smaller estimates of new demand, resulting in lower levels of OD demand available to form corridors. The base case uses the elasticities suggested from a review of studies in the UK. For OD pairs with a current public transport travel time of less than 20 minutes, an elasticity of -0.4 is applied. If the travel time is 20 to 40 minutes an elasticity of -0.5 is used, and if the travel time is 40 minutes or greater, an elasticity of -0.6 is applied.

The direct demand model (See Section 7.3.6) suggested that demand may be more elastic in response to in-vehicle time changes. Therefore, the effect of elasticities ranging from -0.8 (for the shortest journeys) to -1.0 (for journeys of 40 minutes or more) is explored. As presented in Figure 7-25, this has no effect on the corridors identified. The expected benefits are inflated somewhat due to the larger estimates of new journeys. Using an elasticity model, the number of journeys is determined by the expected change in travel time and the current demand. A different type of demand model that produces new demand patterns that differ more significantly from existing patterns may have a more substantive effect on the results of the methodology.

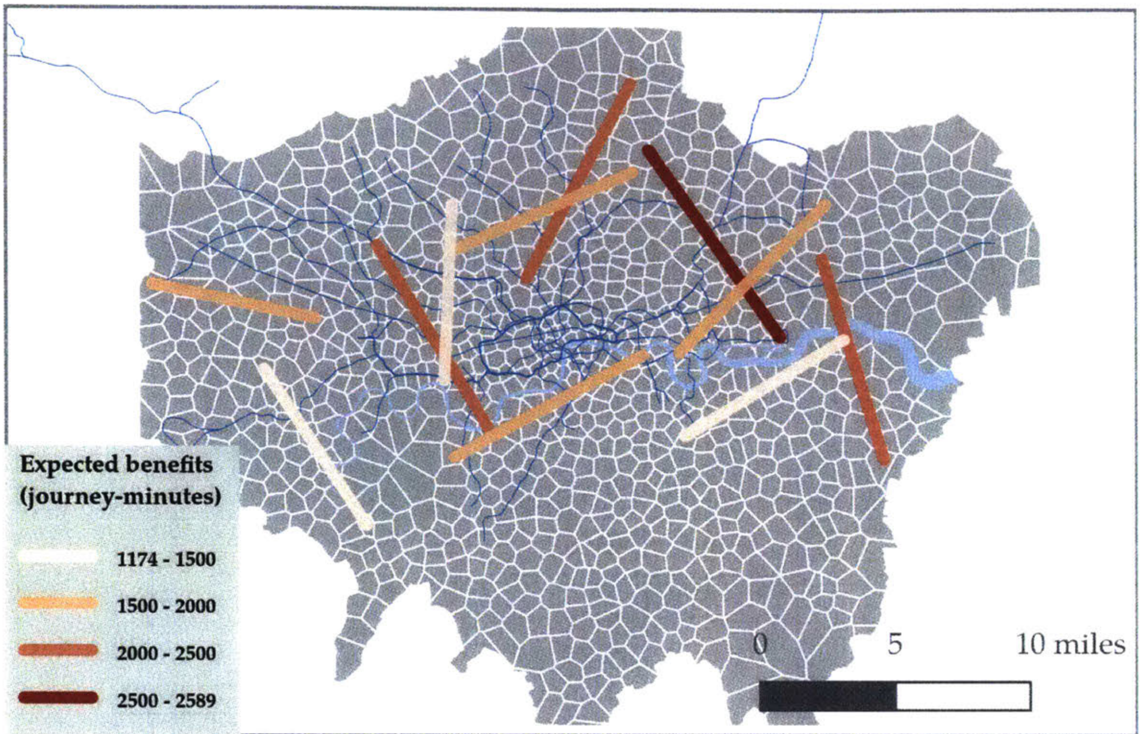

Figure 7-25: Corridors identified with elasticity increased to −0.8 to −1.0

7.5.3 Corridor shape: distance, angle and length

The maximum distance, angle, and length parameters determine the shape of the corridors identified. Maximum distance defines as a buffer extending orthogonally from the linear representation of the potential route. Only OD pairs with origins and destinations within the buffer are included. However, origins and destination are defined by the zone centroid. The zone itself may therefore extend beyond the buffer. Because zones are defined so as to be walkable based on data on access distances, identifying corridors that are approximately one zone wide is desirable. In the base case, the maximum distance parameter is set to 0.4 miles. The zones defined have an average radius of about 0.45 miles, though they vary somewhat in shape and size. Some of the corridors identified in the base case have sections that are two zones wide, and may be difficult to serve with a single route.

Reducing the maximum distance parameter just slightly, to 0.35 miles, has a significant impact, eliminating 6 of the corridors identified in the base case. The five corridors identified are shown in Figure 7-26. In theory, these corridors should have fewer instances of side-by-side zones and be easier to serve well with a single route. In fact, instances of side-by-side zones persist. Section 7.6 shows example corridors identified in the base case and restricted distance scenarios in detail, highlighting the challenges of identifying corridors that can be served by a single route.

The maximum angle parameter eliminates short journeys that are within the maximum distance, but have a significantly different direction compared to the rest of the corridor. Restricting the maximum angle parameter to 15 degrees (with a maximum distance of 0.4 miles) eliminates just three of the corridors identified in the base case. Figure 7-27 shows the 8 corridors identified. Again, restricting the angle will in general result in corridors that

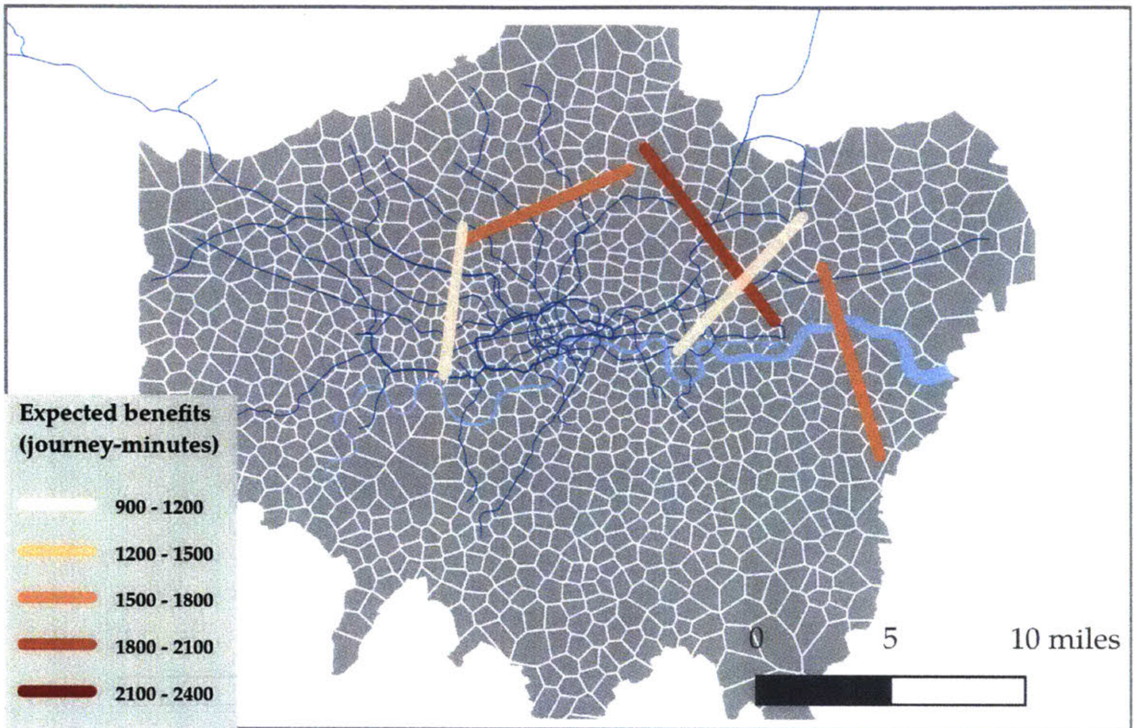

Figure 7-26: Corridors identified with maximum distance reduced to 0.35 miles

can be more easily served by a single route. However, the road network and location of OD pairs in the corridor may present challenges, even with a restricted angle. This issue is explored further in Section 7.6.

Similar to the distance and angle parameters, the length parameter has a physical meaning, and should be set to produce the route lengths desired by planners. Adjusting for the fact that corridor length is measured as a straight line, corridor lengths of 5 to 9 miles are allowed in the base case, corresponding roughly to the 25th and 75th percentile of existing bus route lengths in London (about 6 and 10 miles, respectively).

If the length parameters are relaxed, allowing corridors ranging from 4 to 11 miles (actual expected route lengths of approximately 5 to 13 miles), 11 corridors are identified (See Figure 7-28), however, there are some differences compared to the base case. An additional short corridor is identified in West London. Some corridors are extended, and others are shortened. One east-west corridor south of the Thames is eliminated.

These differences are the result of the implementation of the length and flow constraints in the algorithm. Corridors of three discrete lengths are considered. The discrete intervals evenly divide the distance from the minimum to the maximum length. The algorithm tests the longest length first and determines if the minimum flow is met. If it is not, the shorter lengths are considered. In all three iterations, the minimum length must be met. As a result, depending on the length parameters, different discrete corridor lengths are tested. If there is variable flow along the corridor some intervals may meet the flow constraint while others do not. Highly variable flow is generally not desired. Therefore, the corridors that are only identified with some length parameters and not others may not be good candidates for new service. Ultimately, route length constraints should be set based on planner judgment.

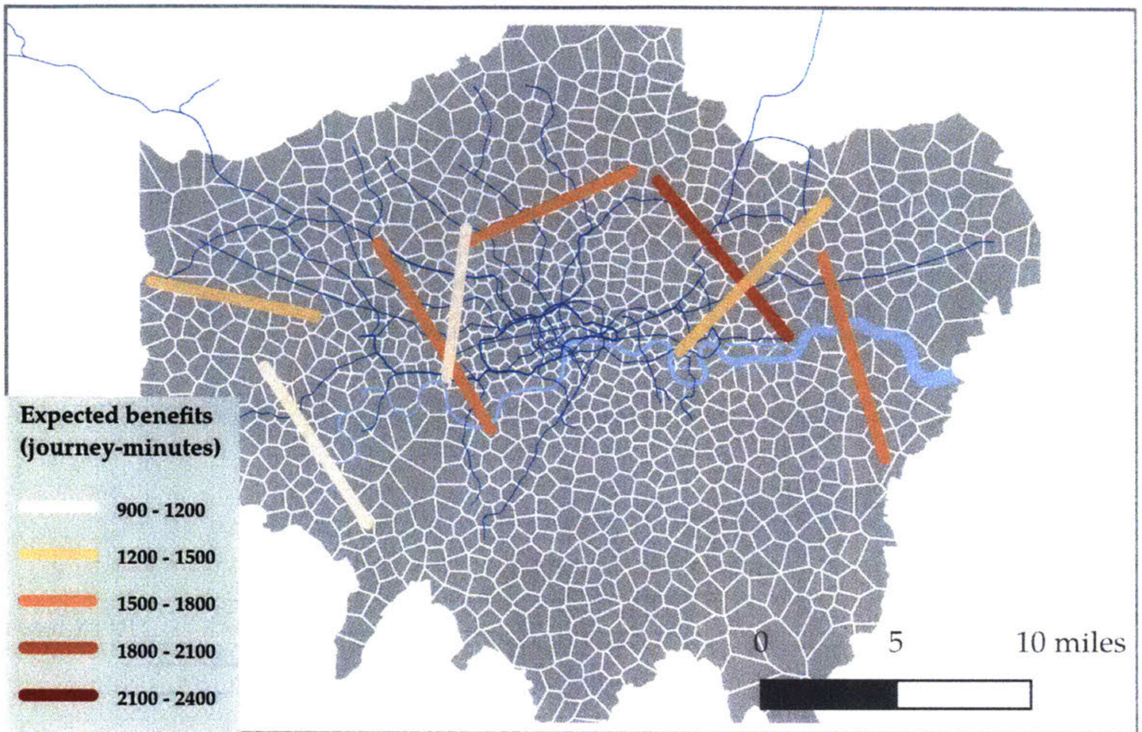

Figure 7-27: Corridors identified with angle restricted to 15 degrees

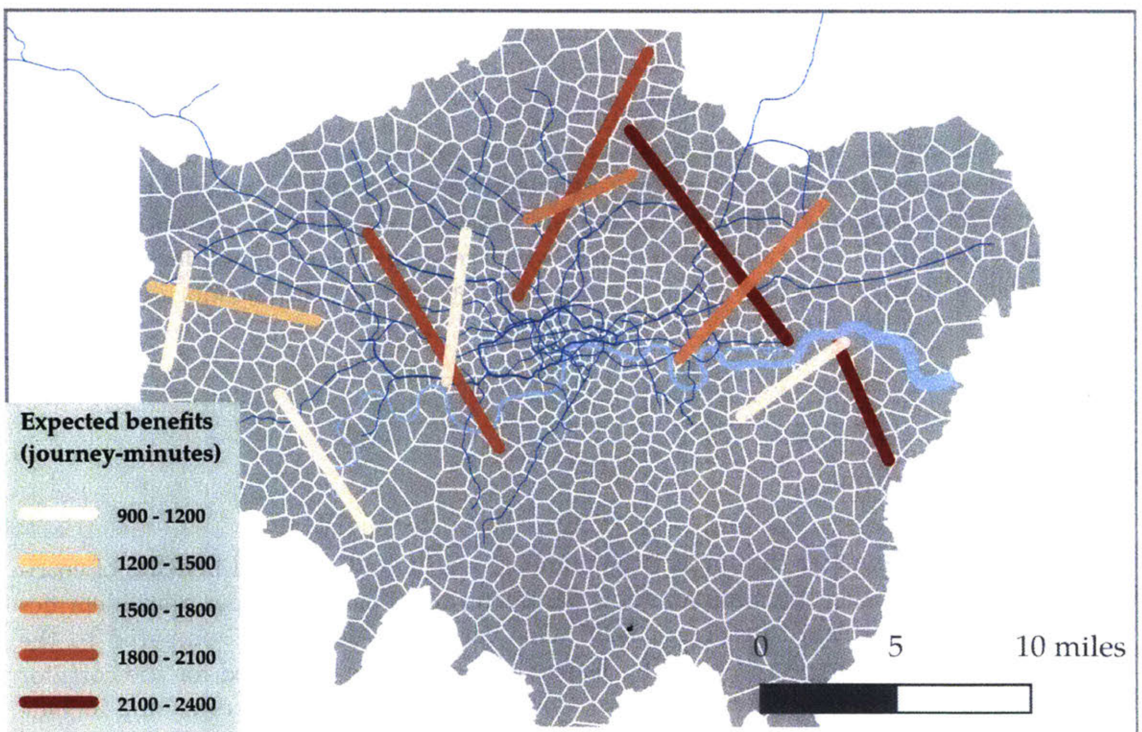

Figure 7-28: Corridors identified with length parameters relaxed

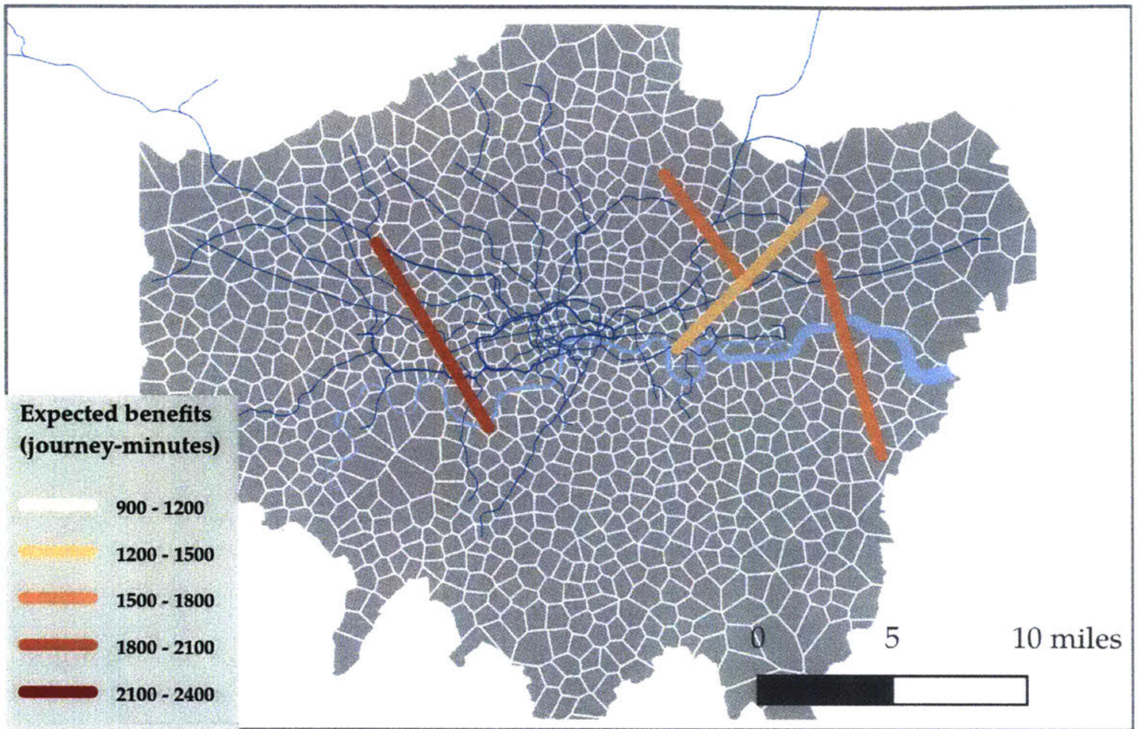

Figure 7-29: Corridors identified with minimum flow parameter increased to 60 passengers per hour

Different lengths can be considered to evaluate how the lengths of identified corridors change and determine which corridors are consistently identified.

Because the inputs are identical in all cases, the expected benefits for corridors in the base case, the restricted distance case, the restricted angle case, and the relaxed length case can all be directly compared. In some cases, restrictions do not affect the expected benefits substantially. Section 7.6.3 explores one case in which reducing the minimum length parameter identifies a shorter, but higher-priority corridor that overlaps the corridor identified in the base case.

7.5.4 Corridor flow

The corridor flow parameter can impact the number, length, and expected benefits of corridors identified. Increasing the minimum flow parameter to 60 passengers per hour identifies just four corridors, displayed in Figure 7-29. One corridor is shortened compared to the base case and has fewer expected benefits. Expected benefits are estimated for the corridor in total and not on a per corridor mile basis. Corridors excluded by this restriction include some with significant expected benefits. Given the sources of additional demand discussed in Section 7.4.3, corridor flow may serve as only a loose indicator of actual expected flow on the corridor. This relationship is explored in more detail in Section 7.6.

7.5.5 Prioritizing corridors

Within the corridor identification algorithm, corridors are prioritized by the sum of the journey-minutes of potential time savings for existing passengers and the expected new passenger journey-minutes. If planners are particularly interested in one of these benefits over the other, the components can be weighted accordingly. The effect of these weights is not shown here because there is no effect on the base case explored. Because new journeys are estimated using elasticities, they are dictated by the current demand and potential travel time savings. In the base case, the corridors identified and the order of priority of corridors is identical whether they are prioritized exclusively by journey-minutes of potential travel time savings for existing passengers or by new passenger journey-minutes.

7.5.6 Summary of sensitivity analysis

In summary, the number of corridors identified is sensitive to all the parameters with the exception of the elasticity values. Relaxing the multiplier and additive factors increases the number of corridors identified but largely preserves the order of priority. Restricting the maximum angle and maximum distance eliminates many of the corridors identified in the base scenario, including some high priority corridors. However, for some corridors, the restrictions have little effect on the expected benefits. The length parameters affect not only the length of corridors identified, but in cases of variable flow along the corridor, can determine whether or not a corridor is identified. This effect is in conjunction with the minimum flow parameter, which also influences the number and length of corridors identified. There are three corridors in East London that are identified in all the scenarios. Two of them are discussed in detail in Section 7.6.

7.6 Scenario analysis

Section 7.5 compared the corridors identified in the base scenario to results with different parameters. Here, the results from two restricted scenarios are considered in more detail. In the first, the maximum distance and maximum angle parameters are restricted. In this scenario, referred to as "restricted angle and distance" the maximum distance is 0.35 miles (compared to 0.4 in the base case) and the maximum angle is 15 degrees (compared to 22.5 degrees in the conservative case). Restricting these parameters should result in corridors that are easier to serve with a single route. However, this effect is explored when the corridors are studied in detail.

The second scenario ("increased flow") uses the base values for maximum distance and angle but increases the minimum flow parameter from 50 journeys per AM peak hour to 60 journeys per hour. This ensures a slightly higher level of flow in the corridors identified. In all scenarios, there is likely to be additional demand that may use new services including latent demand, intrazonal demand, and demand on OD pairs within the corridor that were not identified as candidates for improvement. Post-analysis of corridors estimates some of these additional demand sources for the corridors identified in the two scenarios.

Figure 7-30 shows the corridors identified in these scenarios, with numbers assigned for reference. The "restricted angle and distance" scenario identifies five corridors, numbered 1, 4, 5, and 6; and the "increased flow" scenario identifies four corridors (2, 3, 4, 5, and 6). Three of the corridors: 4, 5, and 6 are similar across these scenarios.

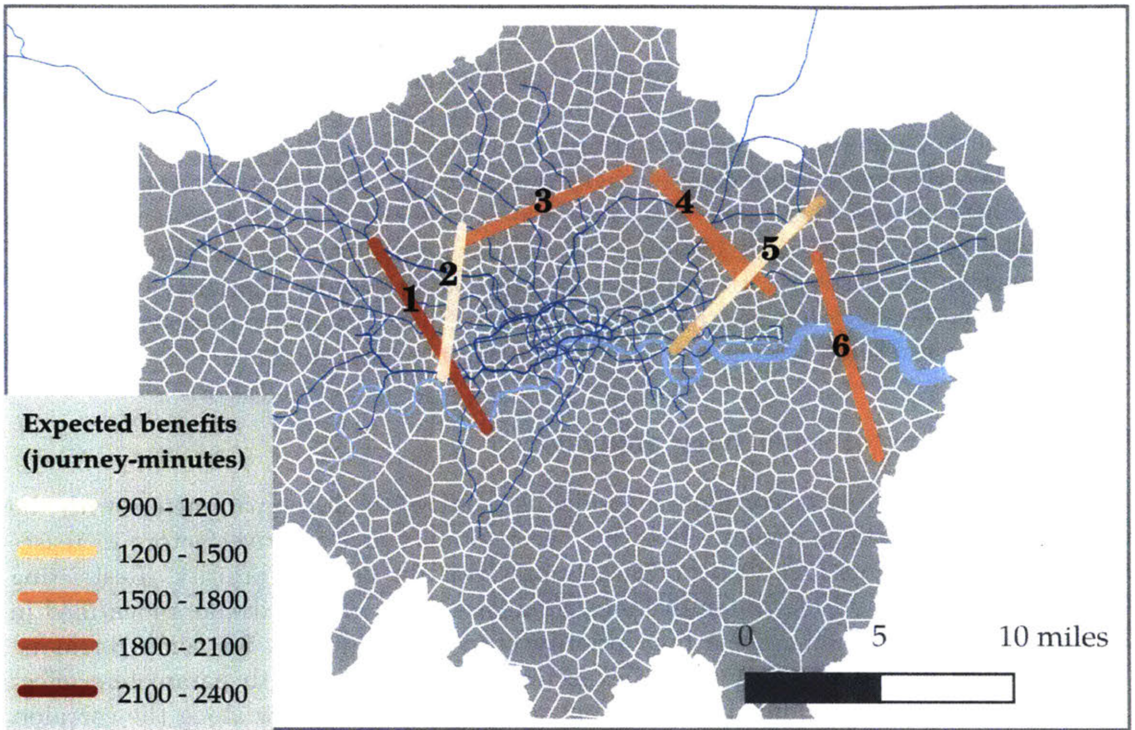

Figure 7-30: Corridors identified in "restricted distance and angle" and "increased flow" scenarios

In both scenarios, only a small percentage of the OD pairs identified as candidates for improvement are assigned to corridors. In the "restricted angle and distance" scenario, 2% of OD pairs and 3% of journeys are assigned. In the "increased flow" scenario, 2% of OD pairs and 4% of journeys are assigned. This reflects the dispersal of the improvable OD pairs across the network. In few cases are the OD pairs concentrated along a corridor with sufficient density to merit a new bus route.

Table 7.5 summarizes the current and expected demand and benefits for each corridor for the average AM peak hour. For corridors 4, 5, and 6, identified in both scenarios, 'ad' refers to the corridor identified in the "restricted angle and distance" scenario and 'f' refers to the "increased flow" scenario. Corridors 2 and 5ad stand out, as they have significantly lower expected benefits, in terms of potential travel time savings and new journey-minutes compared to the other corridors. Compared to Corridor 5ad, Corridor 5f has significantly greater expected benefits. This means that for this corridor, restricting the maximum angle and maximum width removes OD pairs with substantial benefits. In comparison, the expected benefits for corridors 4f and 4ad and for corridors 6f and 6ad are much more similar. In fact, corridors 1, 3, 4, 5f, and 6 all have similar expected benefits, with the expected travel time savings ranging from 1,068 to 1,323 journey-minutes per hour, and expected new journey-minutes from 400 to 490 per hour.

Quantifying additional demand helps determine which of these corridors should have sufficient demand to support a new bus routes, and where there may be additional expected benefits. Some sources of additional demand can be quantified from OD data. Table 7.6 summarizes current intrazonal demand and demand not identified as improvable for each corridor for the average AM peak hour. Intrazonal demand is all current public transport

Table 7.5: Expected corridor demand and benefits per AM peak hour

Corridor number	Current journeys	New journeys	Potential savings (journey-minutes)	New journey-minutes	Average flow (journeys per hour)
1	185	22	1,323	490	61
2	91	12	796	283	50
3	131	20	1,216	407	51
4f	128	20	1,211	432	61
4ad	120	18	1,125	400	51
5f	176	19	1,068	402	62
5ad	117	11	666	253	60
6f	145	20	1,247	451	64
6ad	144	20	1,234	447	67

Table 7.6: Estimates of demand from other sources

Corridor number	Intrazonal journeys	Non-improvable journeys	Low-demand OD pairs
1	355	2,194	108
2	354	1,422	47
3	228	565	45
4f	251	969	22
4ad	321	1,091	22
5f	461	2,294	45
5ad	342	1,543	9
6f	93	394	59
6ad	86	386	43

journeys that start and end within the same zone for all zones assigned to a corridor. Non-improvable demand is the current public transport demand on zonal OD pairs not identified as improvable using the methodology described in Section 7.3, for which both the origin and destination zone have been assigned to the corridor. In reality, a new route may improve or provide similar quality service for some of these OD pairs, and some of these passengers are likely to use a new route, if added.

Table 7.6 also shows the number of zonal OD pairs within origins and destinations assigned to a given corridor that either have no current journeys or fewer than 23 journeys (the minimum sample size used) in the 10 AM peak period analyzed. Because these zonal OD pairs did not meet the sample size constraint for determining the current performance, they were not included in the analysis. However, these OD pairs may have latent demand that would use a new bus service if it was added.

The additional demand varies significantly from corridor to corridor. Corridor 5f has the highest values of intrazonal and non-improvable journeys. Interestingly, Corridor 5ad has far fewer additional journeys. This is likely because the corridor identified in the "restricted angle and distance" scenario is much shorter. Corridors 1, 2, 4, and 5 all have large numbers journeys that were classified as not improvable. If even a small fraction of these journeys switch to a new route, the flow on the route will be significantly more than the value estimated based on the OD pairs assigned to the corridor. Corridors 3 and 6 have the smallest quantities of non-improvable journeys, but even in these cases, these journeys more than double the journeys on the OD pairs identified as candidates for improvement.

In addition to the statistics in Tables 7.5 and 7.6, other information can help prioritize the corridors identified. Factors such as population growth and land-use changes can significantly impact future demand. Ideally, expected new demand due to these factors would be quantified at the OD pair level and included in the corridor identification algorithm. However, if detailed OD matrices are not available, planners can use general knowledge about planned development to prioritize the corridors identified by the algorithm.

Finally, visualizing demand flows and existing routes serving the corridors can identify any potential issues, such as physical barriers, characteristics of the road network, and variations in flow that make some corridors better candidates than others. The following sections describe three of the six corridors identified by the "increased flow" and "restricted angle and distance" scenarios in more detail. Corridor 1 is identified in the "increased flow" scenario, but not in the "restricted angle and distance" scenario. It has the greatest expected benefits of all six corridors, and is discussed in detail in Section 7.6.1. Corridors 4ad and 4f are analyzed in Section 7.6.2, showing the impacts of different parameters on corridor shape. Finally, Corridor 6 is interesting because it crosses the Thames. The impact of this physical barrier is discussed in Section 7.6.3.

7.6.1 Corridor 1

The corridor with the greatest expected benefits extends from Wembley Park to Fulham, including one zone that includes the A217 (Wandsworth Bridge Road) crossing of the Thames. Figure 7-31 shows the mode and stages for the current journeys on the OD pairs included in this corridor. Almost 60% of these journeys are multi-stage bus journeys. Compared to the other corridors, rail plays a more important role in this corridor. Almost 7% of the journeys are multi-stage rail journeys (defined as rail journeys in which the origin and destination stations are not served by the same line or rail journeys with out-of-station transfers) and 15% are combined bus and rail journeys. These journeys are made using a combination of the District and Central lines and the Overground. For these OD pairs, a direct bus link is expected to improve travel times. However, some passengers may have a preference for rail over bus, and therefore may continue to use the existing rail services.

The expected benefits on this corridor, consisting of potential travel time savings for current passengers and anticipated new journey-minutes, are distributed over the 41 zonal OD pairs. Approximately 75% of the benefits, however, are concentrated on just 15 OD pairs. Figure 7-32 shows the zones in the corridor and these 15 OD pairs. None of these OD pairs extend to the zone that straddles the river. In the direction of Fulham Broadway, demand is distributed along the corridor and several long OD pairs are expected to receive significant benefits. The largest expected flows are in the northern part of the corridor, traveling north from the Stonebridge area towards Wembley Park. On these three OD pairs there are approximately 59 journeys expected per AM peak hour. There are 6 to 10 minutes of potential travel time savings for each of these journeys.

Figure 7-33 shows the most important existing bus routes serving the corridor. These routes serve 80% of the demand on the OD pairs shown in Figure 7-32. The three rail services used: the District line, the Central line, and the Overground are also shown.

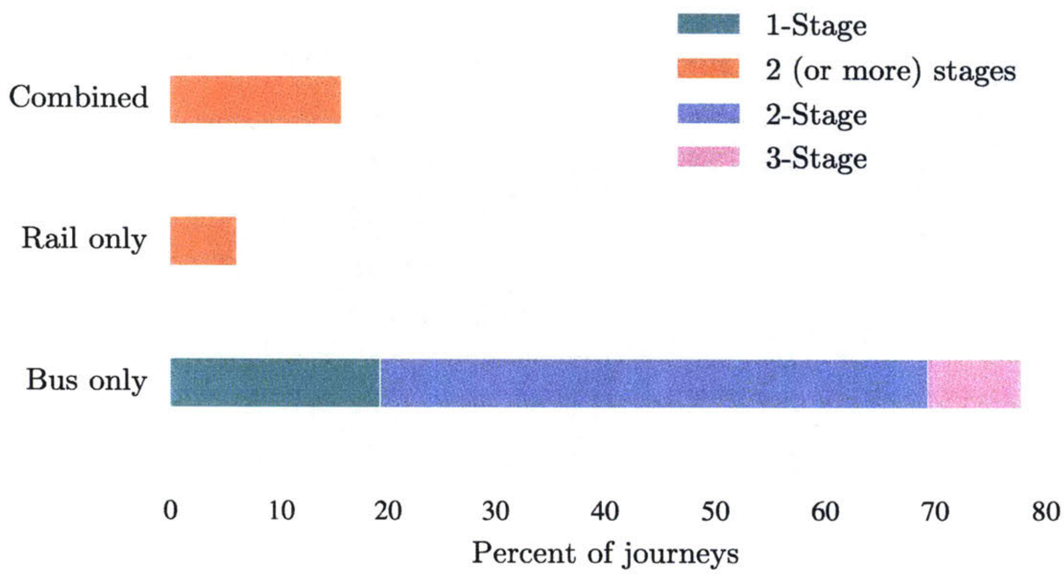

Figure 7-31: Modes and stages for journeys on OD pairs included in Corridor 1

Route 206 traverses the northern part of the corridor, but it takes a circuitous path. There are limited roads running along the corridor in this section due to the River Brent and Chiltern Railways tracks. The shortest path connecting Stonebridge to Wembley Park follows the path of Route 18 and then continues on the path of Route 83 and Route 182. This option, which curves west outside of the corridor, is approximately 2.6 miles from Stonebirdge to Wembley Park. In contrast, the same OD pair along Route 206, which lies within the corridor, is approximately 3.4 miles, as the road network is more circuitous within the corridor. Figure 7-34 shows this OD pair in more detail. This is the highest demand OD pair on the corridor with 29 journeys per AM peak hour, currently. Approximately 54% of current journeys are made on Route 206. The remaining journeys opt for a multi-stage journey, primarily taking Route 18 to either Route 83 or Route 182.

South of Stonebridge, Route 220 serves a large part of the corridor. However, several of the OD pairs included in the corridor begin in Stonebridge, just north of the terminus of Route 220 and extend south, requiring either multi-stage journeys or circuitous travel on Route 266 or Route 260. Extending a single route the length of the corridor would serve these journeys better.

This corridor was not identified in the "restricted angle and distance" scenario. The southern half of the corridor is two zones wide in many places, with a maximum width of about 1.5 miles. The width of the corridor presents a challenge in designing a single route to serve all OD pairs. Figure 7-32 shows that many of the OD pairs that can be improved are in zones next to those served by Route 220 and follow a somewhat different angle than Route 220. While these passengers may be able to walk to Route 220, instead they choose a multi-stage option that presumably reduces their access or egress distance. More detailed data on true origin and destination locations would help determine the optimal alignment for a potential new route.

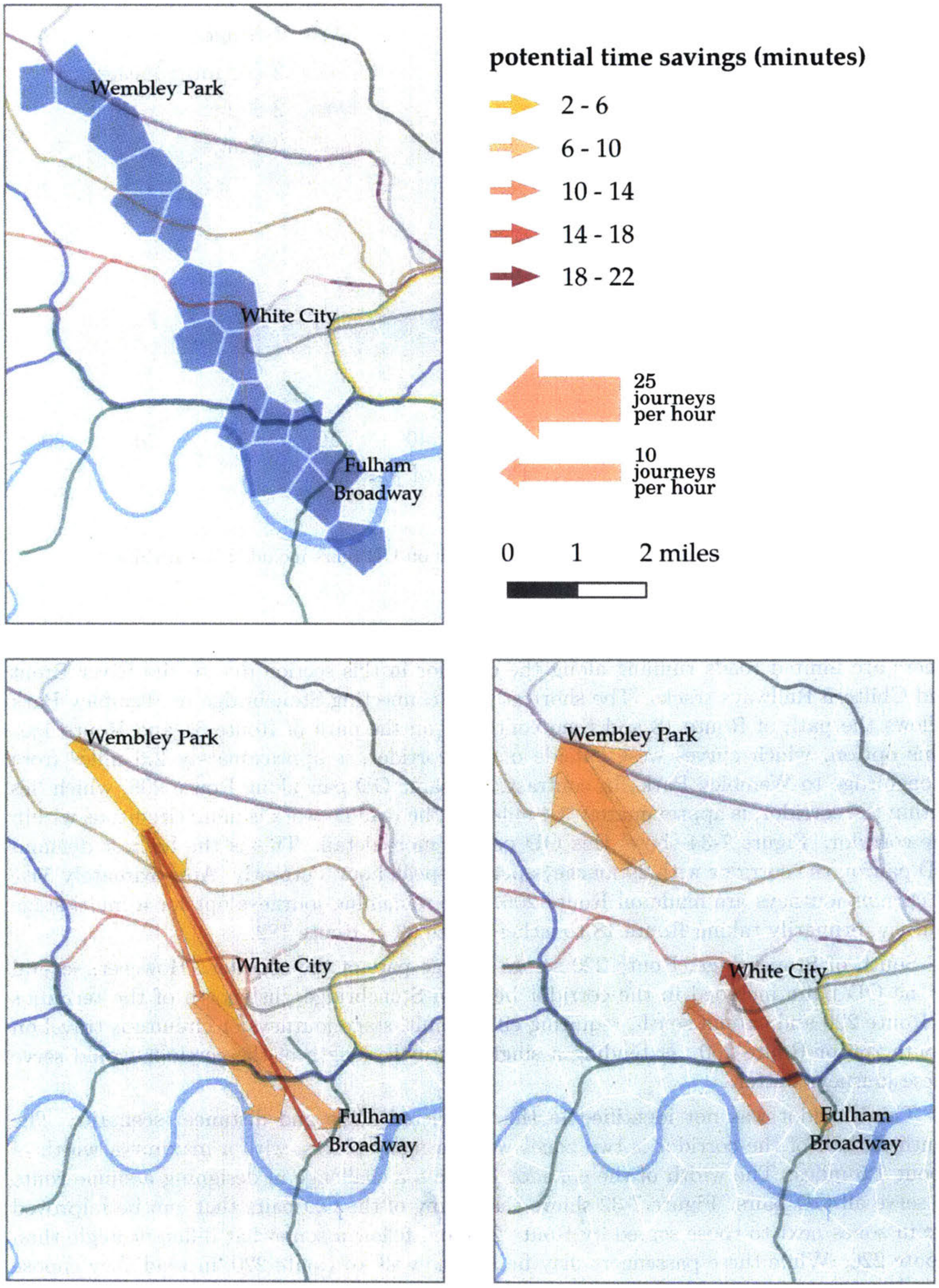

Figure 7-32: Demand and travel time savings on Corridor 1

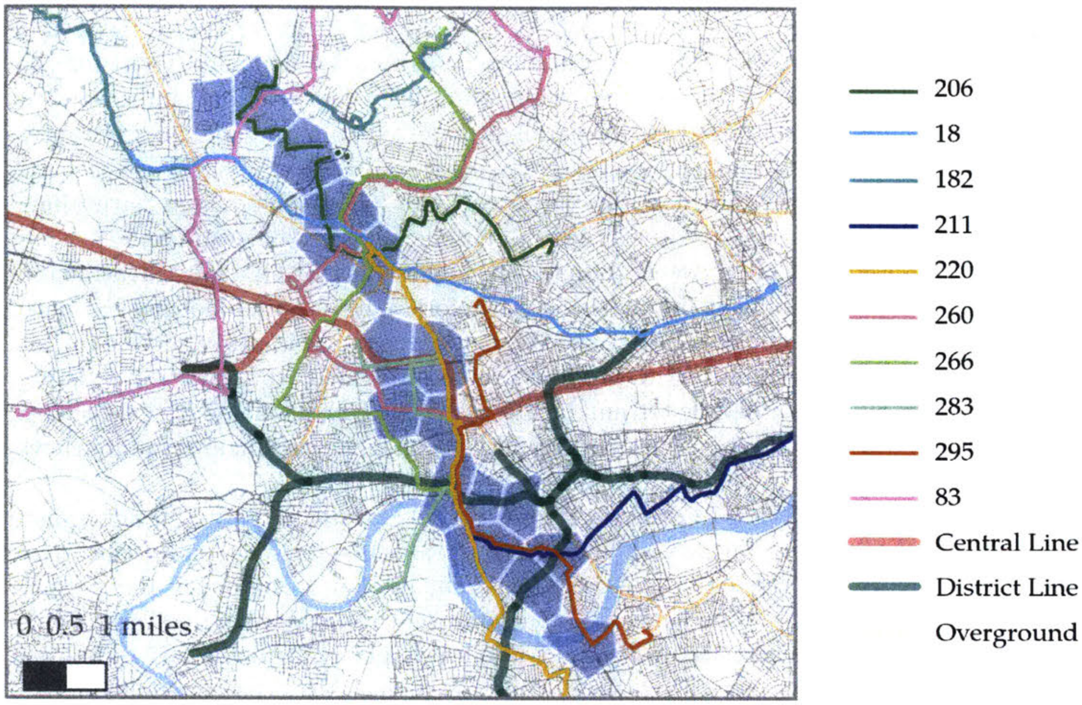

Figure 7-33: Existing routes serving Corridor 1

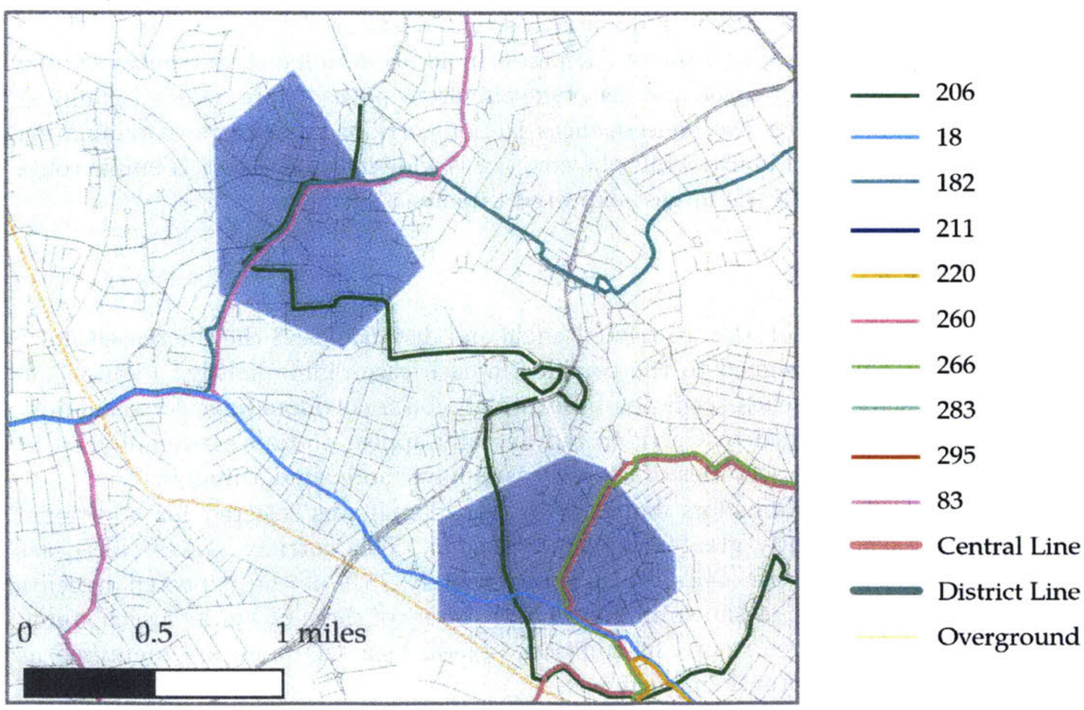

Figure 7-34: Existing alternatives for OD pair on Corridor 1

Given the zonal OD information, Figure 7-35 shows a potential design for a route that could serve the corridor. In the southern part of the corridor it serves the western side of the corridor more so than the east as it has more poorly-served demand. Even for the prominent OD pair southeast of the corridor, this route likely represents an improvement over the existing routes. The proposed route parallels many existing routes including routes 295, 266, 18, and 83.

In the northern part of the corridor, the proposed route follows the curved path along Route 18 and Route 83, and does not serve the middle zone between the origin and destination zones shown in Figure 7-34. However, none of the highest-benefit OD pairs identified in Figure 7-32 originate or end in that middle zone. Therefore, a new route following the path of Route 18 and Route 83 from Stonebridge to Wembley Park would improve service in this part of the corridor. The termini of the proposed route are similar to termini of existing routes, suggesting they are reasonable termini locations. By linking segments of these routes into a single route, many multi-stage journeys could be replaced with single-stage journeys on the new route.

Finally, as was shown in Table 7.6, this corridor has a substantial potential for additional demand including both short intrazonal journeys, and journeys on OD pairs that were classified as not improvable. Some of these passengers may use a new route, if it serves their origin and destination. If 20% of the intrazonal journeys and the demand on non-improvable OD pairs shifts to the new route, this will add 510 journeys per hour to the 207 estimated journeys per hour on the OD pairs identified for improvement. Also, the corridor narrows in some places. In these places, zones within the maximum distance are not included in the corridor, either because they were not classified as improvable or did not meet the minimum number of journeys constraint. These zones may be sources of additional demand on the corridor.

Overall, the corridor has some existing services, but moderate improvements could be made by adding services to better connect OD pairs. In addition, some of the routes serving the corridor, in particular routes 266 and 260, take circuitous paths. More detailed planning for this corridor may consider the consequences to non-corridor passengers of re-aligning some of the existing routes. Finally, while the corridor is wide in some places, a single route can serve most OD pairs with the largest expected benefits.

7.6.2 Corridor 4

Corridor 4 is identified in both the "restricted angle and distance" and the "increased flow" scenarios. The OD pairs assigned to the corridor in each case differ slightly. Figure 7-36 shows the zones assigned in each scenario. Interestingly, both corridors have some side-by-side zones. The zones assigned to Corridor 4ad actually appear to be less well-aligned at the southern end, but the trajectories of the OD pairs have more similar angles.

Given the similarity of Corridors 4ad and 4f, Corridor 4f was selected for additional analysis because it has slightly greater expected benefits. This corridor serves Northeast London and consists of 31 OD pairs, with approximately 75% of the expected benefits concentrated on 14 OD pairs, as shown in Figure 7-37. Many of these OD pairs are clustered in the northern part of the corridor. In addition, several long OD pairs are identified as having significant potential travel time savings.

The journeys on the 31 OD pairs assigned to the corridor are mostly bus journeys (about 90%), with over 50% being single-stage bus journeys. There are also a significant number of combined journeys, making up almost 10% of the journeys on the OD pairs identified for

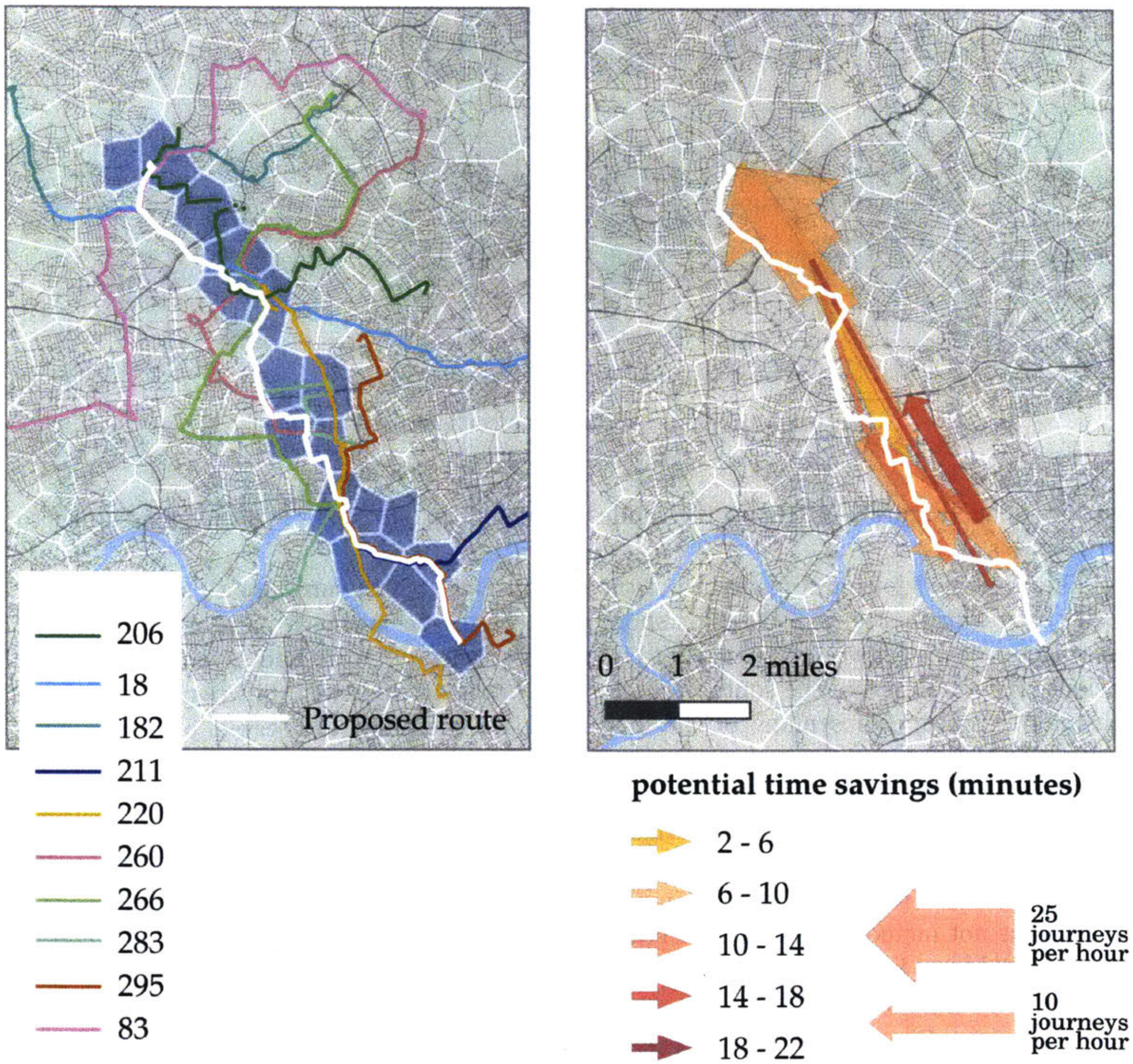

Figure 7-35: Proposed route to serve Corridor 1

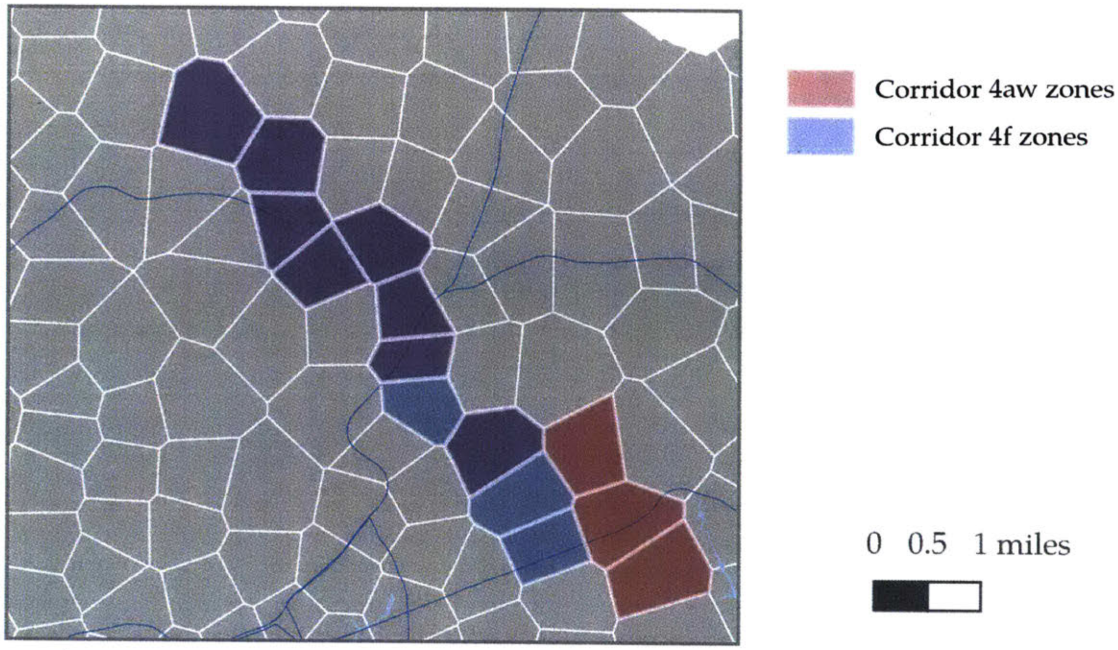

Figure 7-36: Zones assigned to corridors 4ad and 4f

improvement. Figure 7-38 shows the modes and number of stages of journeys assigned to the corridor.

The combined journeys make use of the Overground, which runs along part of the corridor, as shown in Figure 7-39. For short journeys in the northern part of the corridor, passengers choose between a somewhat circuitous journey on W15, and several multi-stage options. Route 58 serves the southernmost zones in the corridor well, but bends away from the middle zones. Thus many of the long OD pairs identified are served primarily with multi-stage options.

The corridor includes a distinct bend. This reflects the fact that only improvable OD pairs are clustered into corridor. In this case, the middle part of the corridor follows the paths of routes W15, 257, and 357, as these routes take circuitous paths and therefore OD pairs along them were identified as improvable. The zone where the Overground crosses Route 97 is not included, as OD pairs starting (ending) in this zone were either classified as not improvable or had insufficient demand.

The bend makes it challenging to design a route that can provide every OD pair with direct service. Figure 7-40 shows a potential route that could serve the corridor, with potential "shortcuts" that provide more direct service for some OD pairs at the expense of not serving others. The primary proposed route offers only a minor improvement over Route W15 in the northern part of the corridor. Providing a non-circuitous route to serve the middle of the corridor is difficult due to both the bend and the A12, which intersects the corridor and can only be crossed at limited locations. The northern terminus of the proposed route corresponds with the terminus of Route W15, indicating that it is a reasonable location for a route terminus. At the southern end, a small extension will reach the terminus of Route 58, although it may also be possible to have a terminus at the location depicted.

As shown in Table 7.6, this corridor has significant potential for additional demand, including 969 journeys per hour on the non-improvable OD pairs within the corridor. Despite

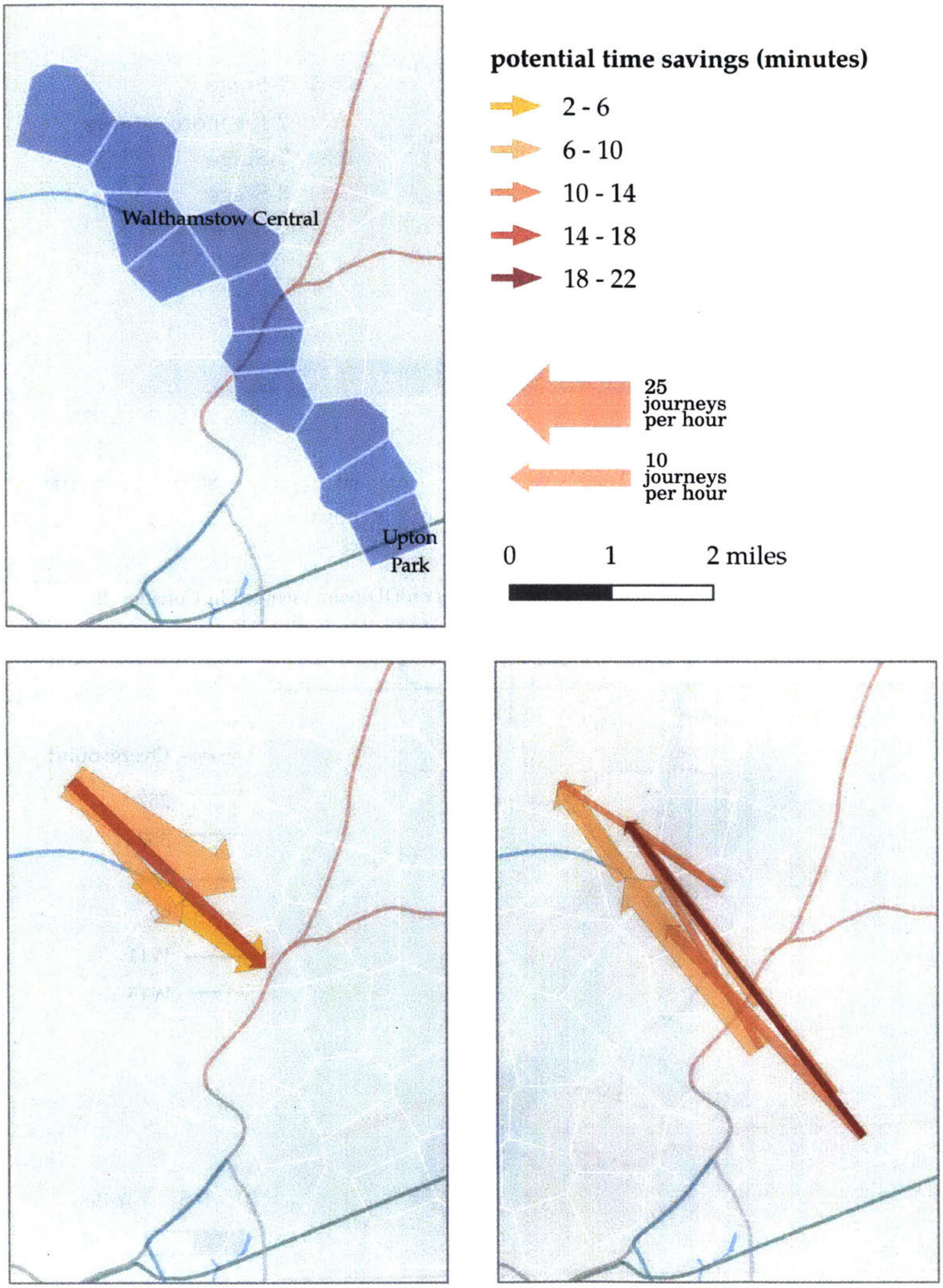

Figure 7-37: Demand and travel time savings on Corridor 4f

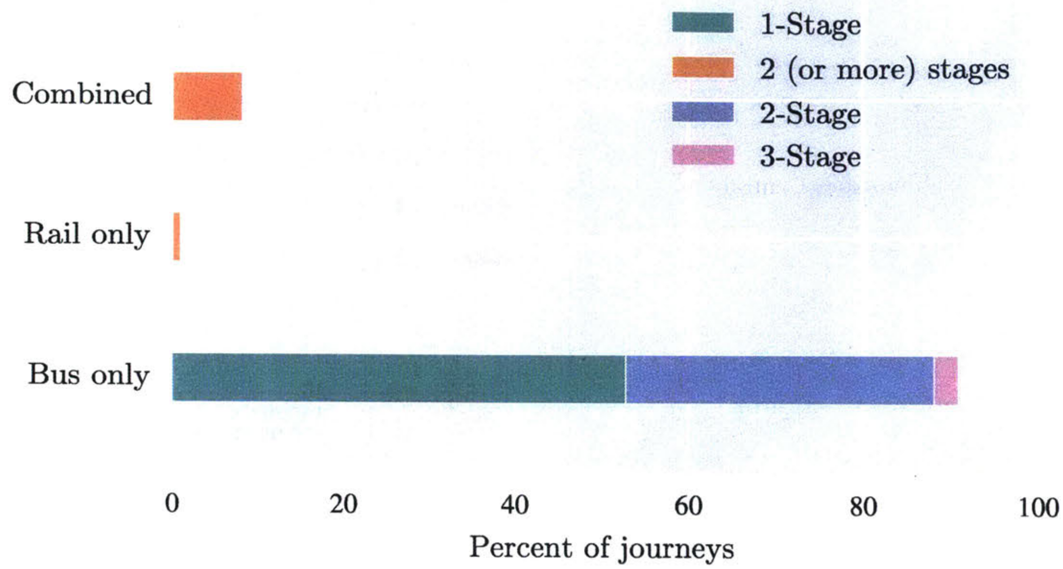

Figure 7-38: Modes and stages for journeys on OD pairs included in Corridor 4f

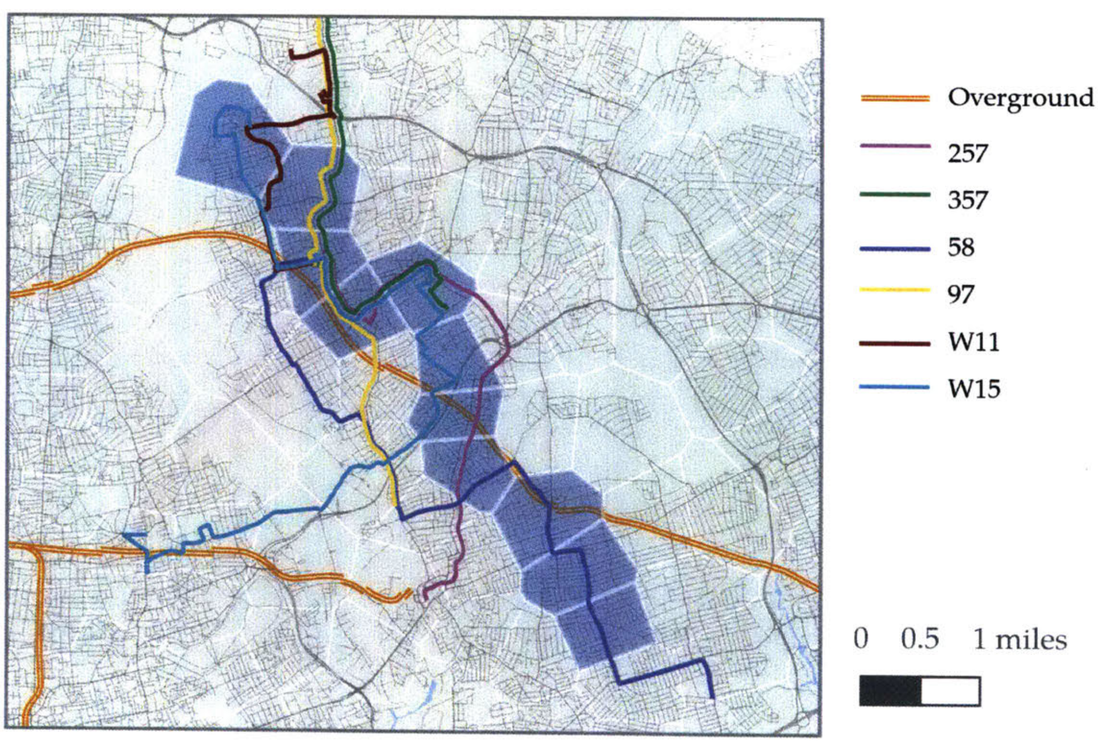

Figure 7-39: Existing routes serving Corridor 4f

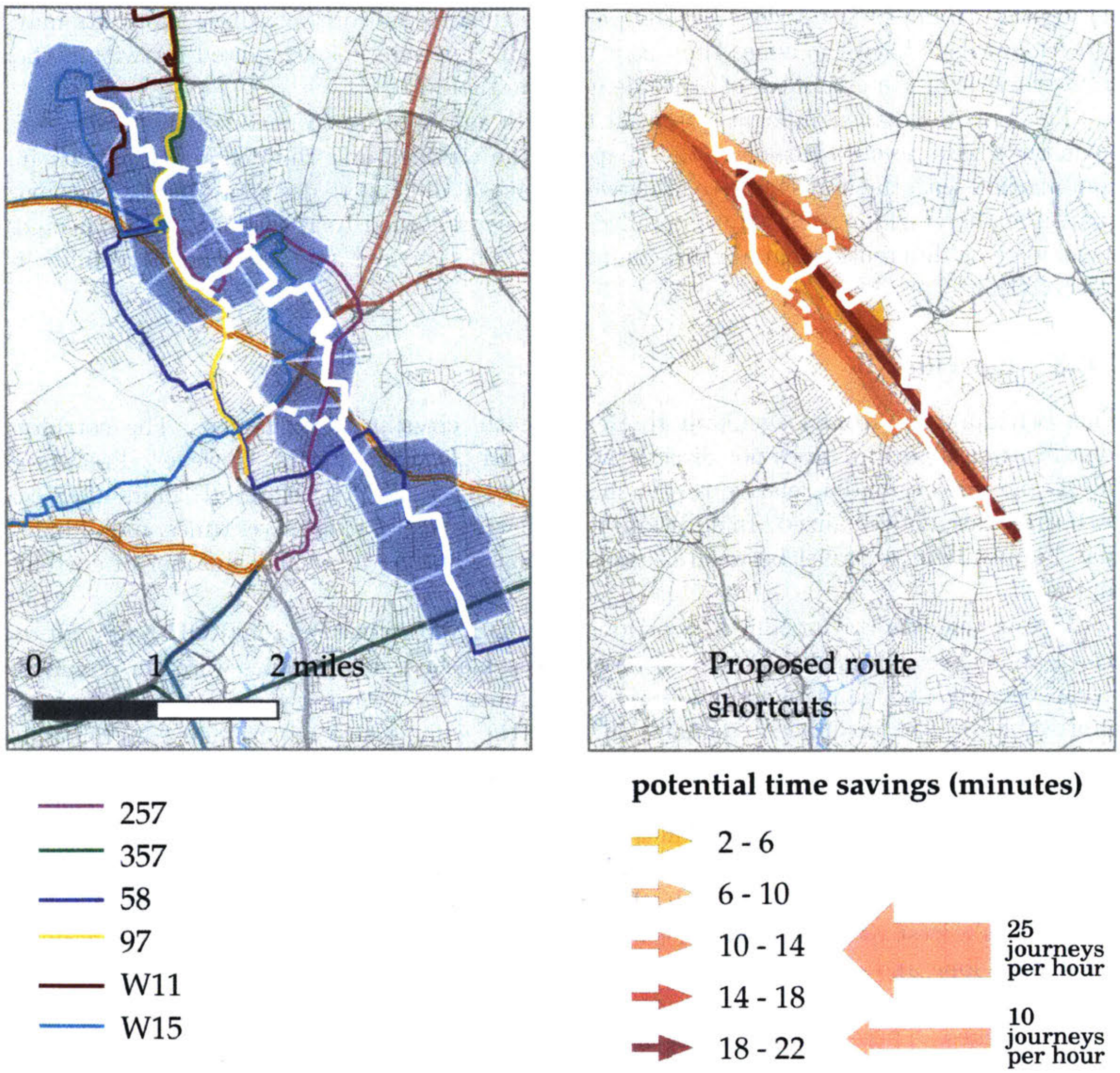

Figure 7-40: Proposed route to serve Corridor 4f

the presence of the Overground along part of the corridor, only 17 of the non-improvable journeys are made using single-stage rail. The remainder are made either by bus or a combination of bus and rail. While the expected travel time for these journeys was not expected to be improved by a new route, depending on the origin (destination) locations of these journeys within each zone, some may see improvement through a new service. It is reasonable to expect some of these journeys to shift to a new route. Also, similar to Corridor 1, the proposed route shortcuts extend outside of the zones assigned to the corridor. In addition to improving service for OD pairs assigned to the corridor, these shortcuts may attract journeys from the zones they pass through that were not accounted for within the framework or in the estimates of other demand in Table 7.6.

This corridor has significant potential benefits and appears to have sufficient demand to merit a new route. However, the shape of the corridor and the demand along it, in combination with the underlying road network makes it difficult to design a single route that can serve the corridor well. More detailed analysis is needed to re-evaluate the expected benefits given different potential route alignments. This corridor appears to be more difficult to serve than Corridor 1.

7.6.3 Corridor 6

This corridor runs from Dagenham to Bexleyheath, crossing the Thames. The corridor identification algorithm does not directly account for physical barriers. However, directness of OD pairs is evaluated relative to the shortest path by car. Therefore OD pairs affected by barriers are not penalized. In addition, when barriers cause long travel times, individuals may be unwilling to travel between these locations, resulting in a small number of current journeys.

This is the case in Corridor 6. Corridor 6 is made up of just 13 zonal OD pairs, all of which are shown in Figure 7-41. The corridor has some gaps, where zones were not included either because OD pairs including these zones along the corridor were not identified as improvable or had insufficient demand to meet the sample size requirement. While the corridor includes both sides of the Thames, none of the OD pairs cross the river. In fact, only 2 journeys were made from any of the three zones to the north of the Thames River to any of the 7 zones south of the Thames in the 10 AM peak periods analyzed. As a result, none of these cross-Thames OD pairs had enough journeys to be evaluated. Crossing the Thames in this corridor requires multiple public transport stages and taking more than an hour. The shortest path available for car travel between the two zones adjacent to the river is 8.5 miles long and includes a ferry.

Therefore, this corridor should be considered as two separate corridors, one on each side of the Thames. There is no rationale for serving the full corridor and no reasonable path available to serve it. Furthermore, the OD pairs north of the Thames account for only a very small share of the total demand and expected benefits on the corridor. Each OD pair north of the river currently has about 1 journey per AM peak hour. The corridor was identified in both the "restricted angle and distance" and "increased flow" scenarios. The only difference in the definition of the corridor in the two cases is the exclusion of one of the OD pairs north of the Thames in the "restricted angle and distance" case. In the truncated section shown in Figure 7-42, the two are identical. In the remainder of the analysis presented here, the corridor is truncated, with only the portion south of the Thames included.

The straight line length of the truncated corridor is 4.7 miles, just under the minimum length parameter applied (5 miles). In fact, the truncated corridor closely mimics a corridor

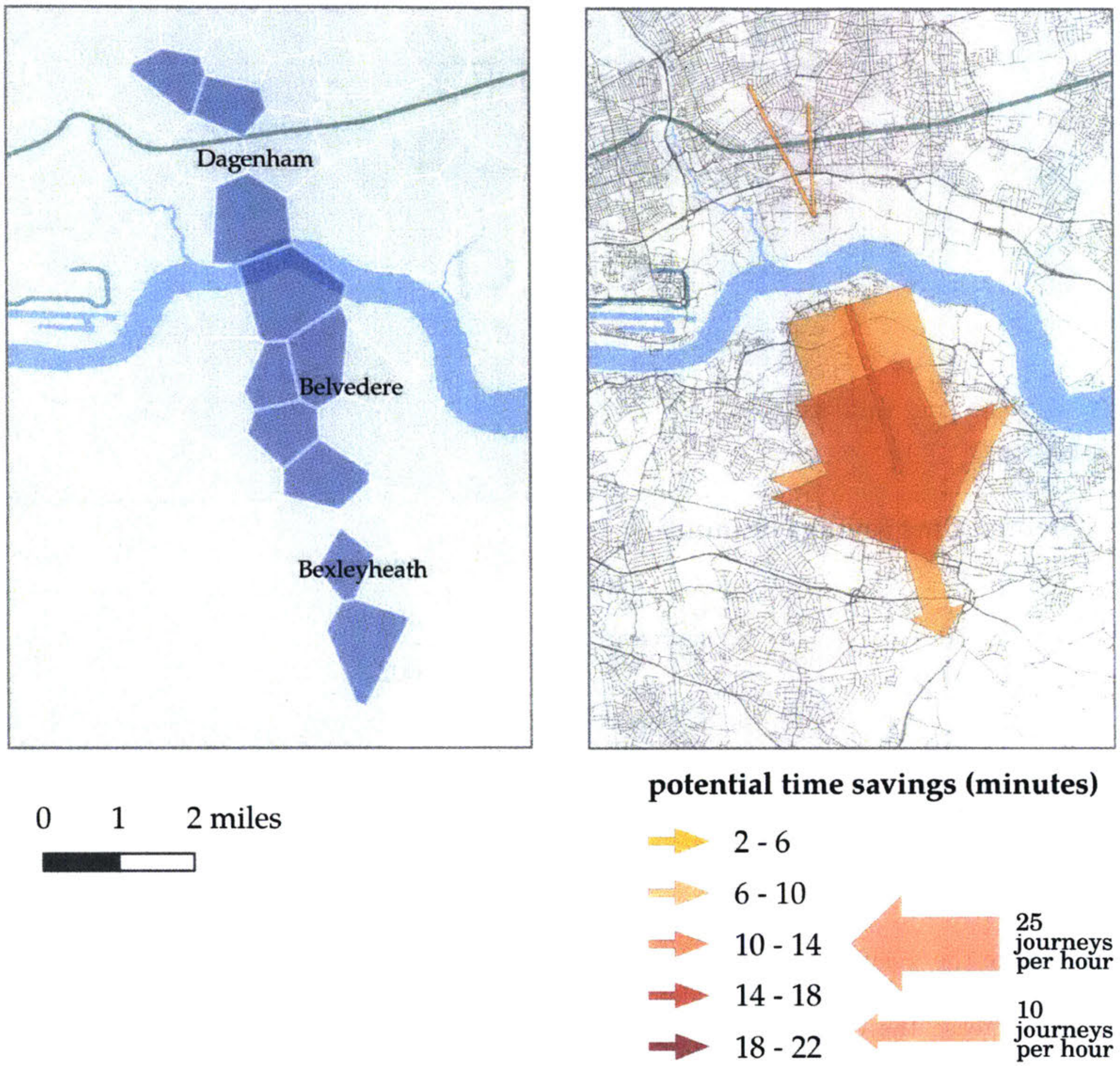

Figure 7-41: Demand and travel time savings on Corridor 6

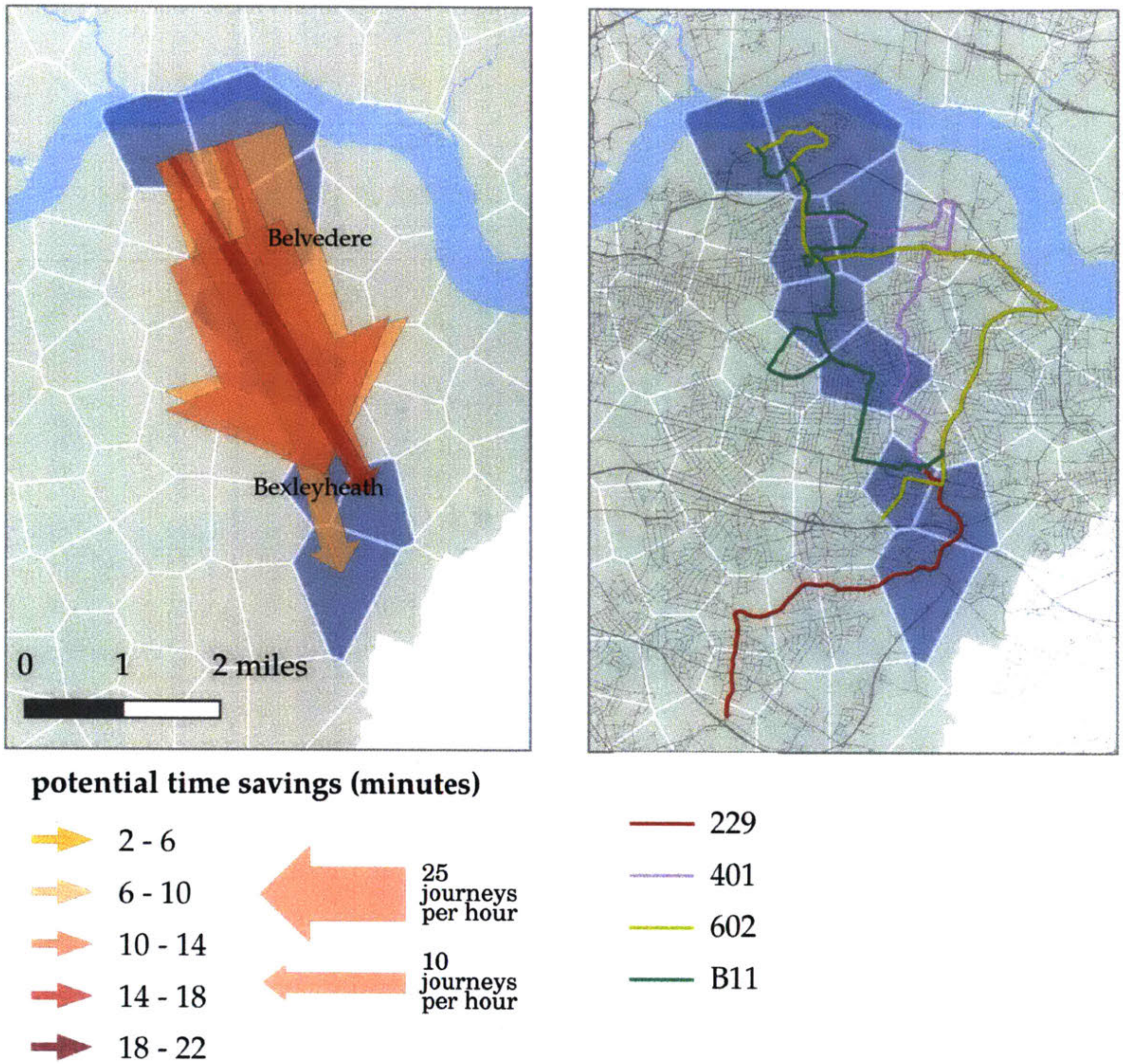

Figure 7-42: Updated Corridor 6 with existing routes

identified in the "relaxed length" scenario discussed in Section 7.5.3. The corridor identified in the "relaxed length" scenario is identical except for the addition of two OD pairs, which further increase the expected benefits for the corridor. Figure 7-42 shows the updated corridor including these additional OD pairs.

The additional OD pairs assigned to the short corridor increase the expected travel time savings to a total of 1,613 journey-minutes per AM peak hour, with 595 expected new journeys-minutes per hour. This means the short corridor is expected to attract more benefits than any of the other corridors identified in the "restricted width and angle" and "increased flow" scenarios. Because the corridor is shortened, the average flow is significantly higher, 132 journeys per hour, dwarfing other corridors analyzed.

Compared to other corridors, this corridor has less intrazonal and non-improvable demand. However, the fact that the improvable current and expected demand generates substantial average flow means that the corridor has sufficient demand to sustain a new route,

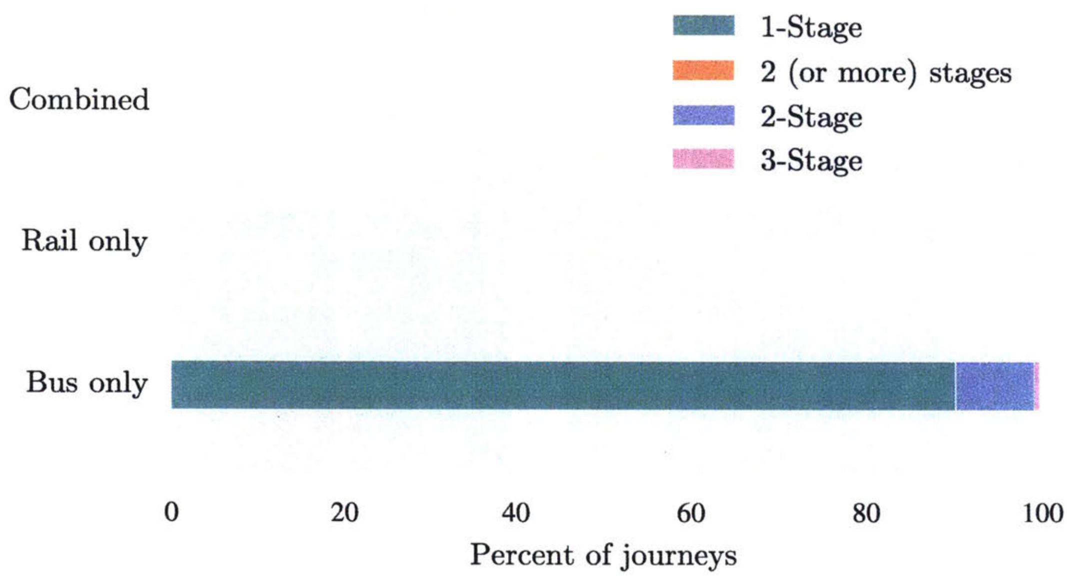

Figure 7-43: Modes and stages for journeys on OD pairs included in Corridor 6

even if only a small number of short and non-improvable journeys use it. Planned development in this area may also generate additional demand.

Currently, the corridor is served only by bus. As shown in Figure 7-43, approximately 90% of current journeys are single-stage. However, the routes serving the corridor are circuitous (See Figure 7-42).

Figure 7-44 shows a potential path for a new route which is less circuitous than the existing routes. Additional analysis is required to determine the consequences if this new route replaces an existing route. The proposed route does not serve the northeastern part of the corridor as well as the northwestern part. This is deliberate, given the demand patterns on the route: the western portion of the corridor has more demand. The northern end of the proposed route is close to the termini of two existing routes, Route 602, and Route B11. At the southern end of the proposed route, more analysis is necessary to determine if the proposed terminus is appropriate.

In short, this corridor is a good candidate for new service, as it has large concentration of poorly-served demand and is the site of new development. The existing bus routes serving the corridor are circuitous. A more direct option, such as the one proposed, may bring substantial benefits to existing passengers and is also likely to attract some new passengers.

7.6.4 Summary

Corridors 2, 3, and 5 are not discussed in detail here. Corridor 2 has limited expected benefits compared to the other corridors identified. Corridor 3 has significant expected benefits, but its flow (based on improvable OD pairs) is low, and the non-improvable and intrazonal demand is less significant than other corridors. Corridor 5f is very promising with significant expected benefits and additional journeys that could contribute to flow on the route. However, restricting the angle and width parameters, as in Corridor 5ad,

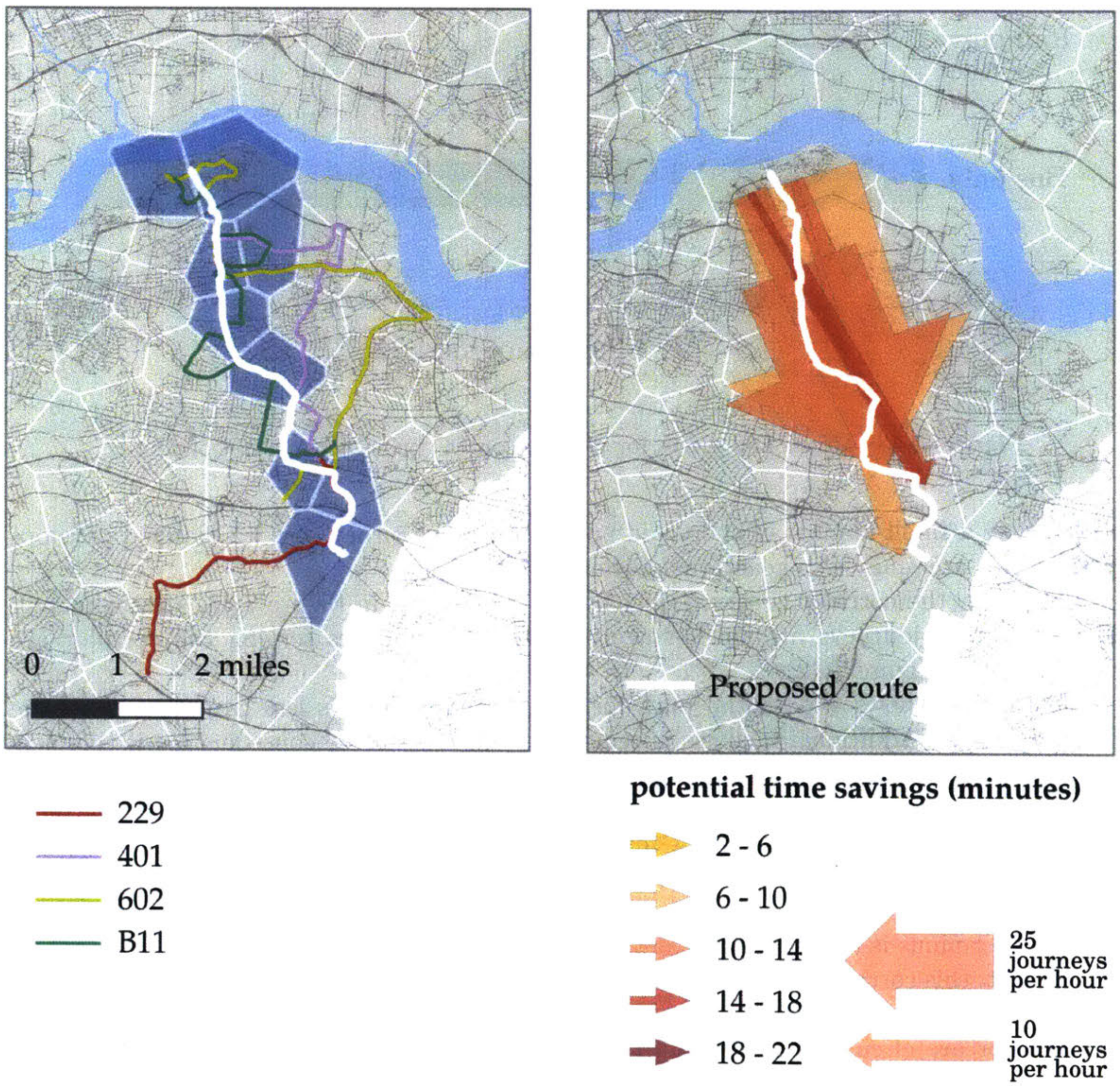

Figure 7-44: Proposed route to serve Corridor 6

significantly reduces the expected benefits on the corridor. While it is unclear that restricting the maximum angle and maximum distance parameters results in a corridor that is easier to serve, the potential of Corridor 5 is highly dependent on whether the corridor shape and road network allows a single route to serve the demand well.

Of the three corridors analyzed in detail, Corridor 6 is the most promising, with the greatest expected benefits and the most existing improvable flow. Corridor 1 appears to be a good candidate for new service, with slightly smaller expected benefits, and more of the expected demand stemming from OD pairs that are already moderately well-served. Corridor 4 is the worst candidate, as the shape of the corridor makes it difficult to serve all the OD pairs well. Interestingly, this issue persisted in the scenario in which maximum angle and maximum distance were restricted.

In all cases, the proposed routes are only approximations. There may be particular locations, such as schools or hospitals that are critical to serve and require deviations from the proposed routes. Planners may also wish to consider a broader set of of potential corridors. Several of the corridors identified in the base and "relaxed length" scenarios have expected benefits of more than 1,200 journey-minutes of potential travel time savings per hour.

Finally, it is important to note that the results discussed here are for the AM peak period. This likely explains the differences in demand by direction in the corridors analyzed in detail. Prior to making decisions about new routes, planners should apply the framework to data from other periods of the day, and compare the recommendations. The final selection of corridors for new services may take into account the variation in demand and expected benefits by time of day.

7.7 Discussion

London has a dense public transport network, and not surprisingly, the majority of the OD pairs defined in Step 1 of the framework are well-served. About 17% of the OD pairs are identified as candidates for improvement. The corridor identification algorithm defined between 4 and 16 corridors based on these OD pairs, depending on the parameters applied. This is a reasonable number of corridors for planners to consider in more detail. Of three corridors evaluated in detail, two appear to be good candidates for new bus service.

Of note, is the fact that only a small percentage of journeys on improvable OD pairs are assigned to corridors (between 3 and 11% in the different scenarios). This is due to geographic dispersal of the OD pairs identified for improvement, and suggests opportunities for services other than standard bus routes to improve service on these OD pairs. Extensions of the framework to identify opportunities for these types of services are discussed in Chapter 8.

The framework is designed for high-level bus network sketch planning. However, it is based on a granular assessment of OD pairs. As a result, it can identify subtle deficiencies in the network. For example, Corridors 1 and 4, assessed in detail in Section 7.6, appear to be generally well-served, with a high density of bus and rail services. However, many OD pairs within the corridor require multi-stage and circuitous paths. While the individuals traveling on these OD pairs may have more direct services available to them in neighboring zones, the analysis shows that many individuals opt instead to use circuitous and multi-stage paths, presumably in order to reduce the access and egress distances required for their journeys.

Step 2 of the framework systematically identifies these opportunities across a large network consisting of thousands of zonal OD pairs. Here, a binary classification is used: OD pairs are either improvable or not. Section 7.7.1 discusses the issues with this method and possible improvements.

Based on the opportunities identified in Step 2, the corridor algorithm finds corridors that are good candidates for new bus routes. The results revealed several promising corridors. However, two challenges also emerged. The first is appropriately estimating corridor flow, discussed in Section 7.7.2, and the second is constraining the shape of the corridor such that it can be served by a single route (See Section 7.7.4).

7.7.1 Classifying OD pairs

The application of the framework demonstrates the importance of the target bus distance and expected travel time. Reducing the target distance and expected travel time, even marginally, resulted in the identification of many additional corridors. However, when these thresholds are reduced, the likelihood that OD pairs identified for improvement are actually improvable decreases. Classifying an OD pair as improvable (or not) is challenging.

Estimates of current distance and travel time are subject to variation in paths and origin (destination) stops and stations. In some cases, current distance or travel time are not directly comparable to the target distance or expected travel time. This can result in misclassifications of OD pairs. Section 7.3.2 proposes an alternative method for defining current and target distance to ensure that origin and destination points are identical and therefore distances are directly comparable. This would eliminate some misclassifications.

Still, the use of thresholds – target distance and expected travel time – to determine whether OD pairs are well-served and improvable presents challenges. The target distance serves as a planner-defined standard for directness. However, planners may struggle in deciding what the definition should be. In some parts of the network it may be easy to design and operate a route that provides very direct service, with a distance similar to the shortest path distance. In others, it may be much more challenging. Defining a single standard may miss opportunities for improvement beyond the standard, and in some cases may select an OD pair for improvement, when improvement is not possible.

These issues are most prominent for OD pairs where the current distance and travel time is very similar to the target distance and expected travel time. OD pairs with current distance and travel times just over the threshold are identified for improvement, while similar-performing OD pairs just under the threshold are excluded from the analysis. This boundary issue can be mitigated with a within-public-transport demand model, as mentioned in Chapter 5. Such a model would assign less demand to a new service if the expected improvements are small, as would be the case for OD pairs very close to the thresholds. As such, these boundary-adjacent OD pairs will have less impact on the results of the corridor identification methodology. Or, a simpler heuristic approach could assign a percentage of current demand to the potential new service, with the percentage increasing as the expected time savings increase.

Chapter 5 also described how a range of values could be estimated for potential travel time savings and new journeys producing a range of values for expected benefits of new services for each OD pair and hence for each potential corridor. These ranges signal the relative levels of uncertainty of the expected benefits for different potential corridors. However, defining ranges presents additional challenges, as the sources of uncertainty in these estimates may not be fully understood.

Despite some issues with the binary classification of OD pairs, it also has benefits.

The definitions of well-served and improvable are clear and easily applied. In addition, the definitions act as service standards, which are common in public transport planning. Analysis of the sensitivity of resulting potential corridors to varying definitions of well-served and improvable can help planners evaluate service quality and opportunities in the network. In the context of the framework, planners can experiment with different definitions of well-served and improvable and the resulting corridors allow them to easily see how opportunities for new service change when different standards are applied.

7.7.2 Understanding corridor flow

In the results presented in Section 7.6, a low minimum flow parameter was used. In conjunction with a conservative definition of improvable, larger values for minimum corridor flow disqualify most of the corridors identified. However, analysis of other potential sources of demand, not accounted for in the corridor flow estimates, suggested that these additional sources are significant for many corridors.

Additional sources of flow include intrazonal journeys, journeys on OD pairs not identified as improvable, and stages of multi-stage journeys on improvable OD pairs. In the case of intrazonal journeys and journeys on OD pairs not identified as improvable, the expected benefits for these passengers are uncertain and expected to be small. As discussed in Chapter 6, including these OD-level flows as inputs to the corridor identification methodology is not recommended. Even if they are included as OD-level demand without expected benefits, they may significantly impact which corridors are identified and result in many corridors being identified that have very small expected benefits. Stages of multi-stage journeys for which the OD pair representing the full journey was identified as improvable are already included as inputs for the full journey OD pair. Including the stages as inputs would result in double-counting. For these reasons, these additional sources of flow are estimated in post-analysis of the corridors identified by the corridor identification algorithm. These estimates help make final decisions about which potential corridors should be prioritized for new service.

In the application of Step 2 for the London network, the estimates of current and new demand for improvable OD pairs were produced with fairly simple methods. For journeys using the current network, it was assumed that all existing journeys on OD pairs that are identified as candidates for improvements will switch to the new service. As discussed, these demand estimates could be improved with a heuristic or more sophisticated assignment model.

New journeys on OD pairs that are candidates for improvement were estimated using elasticities. The elasticities estimate the expected demand change due to changes in in-vehicle time. They do not account for impacts of access, egress, waiting time, and capacity changes that new bus routes will bring. Nor do they account for variability in the existing public transport mode share between OD pairs, population growth or planned development. Incorporating these factors to estimate OD-level demand would improve estimates of corridor flow.

Finally, OD pairs that currently have very few public transport journeys (less than one per AM peak hour) are not included in the analysis. These OD pairs may have low public transport demand because they are poorly served, and therefore *may have significant* latent demand. However, understanding performance and demand on these low-demand OD pairs requires additional data, such as disaggregate demand data for other modes, and population and employment data. The inclusion of estimates of this latent demand and the potential benefits to these new passengers could identify other opportunities for new services in the London network.

In addition to improvements in demand estimation, Chapter 6 discussed a potential improvement to the corridor identification algorithm to account for gaps in flow along the corridor. Using a segment-level minimum flow requirement instead of (or in addition to) a corridor-level requirement will exclude corridors such as Corridor 6, which included zones on both sides of the Thames, which caused a gap in flow. (The elimination of the longer corridor would allow the shorter corridor, south of the Thames to be identified, as the algorithm prioritizes longer corridors over shorter ones). However, this was the only corridor identified that had this problem, and it was easy to adjust in post-analysis. This suggests that this extension of the methodology may not be critical.

7.7.3 Accounting for existing routes

Chapters 1 and 3 introduced and discussed the challenges of planning new bus routes in an existing network. While the framework is designed to identify opportunities for new routes, the impacts on existing routes cannot be ignored. In the application to London's network, existing routes were displayed in the post-analysis to confirm that the corridors were not already served by direct single-stage service.

The OD-level analysis used demand in the existing public transport network to estimate demand for new services. This analysis can also estimate the expected impact of new services on existing route ridership. More sophisticated methods for assigning demand to proposed new services will improve these estimates. Given estimates of ridership reduction on existing routes, planners may decide to accompany the addition of new routes with the removal of existing routes. Here, this process is left to post-analysis on a corridor-by-corridor basis. However, Chapter 8 discusses the possibility of incorporating decisions about the removal of existing routes into the framework using an iterative process.

7.7.4 More realistic representations of corridor shape

Of the corridors identified, some were challenging to serve with a single route. These corridors were two zones wide in some places, requiring a potential new route to take a circuitous path in order to serve all OD pairs assigned to the corridor. The corridor identification algorithm uses a set of parameters to constrain the corridor shape: maximum distance, maximum angle, and minimum and maximum length parameters. Because the corridors are composed of zonal OD pairs, and the zone size is defined to be walkable, corridors are ideally one zone wide. The analysis showed that in the London context, reducing the maximum angle and maximum distance parameters reduced the number of corridors identified, but did not eliminate the identification of corridors wider than one zone. Because the results depend on the alignment of zone centroids, it may not be possible to tune these parameters such that corridor width is never an issue.

Additionally, the maximum distance and maximum angle parameters prevent large deviations from linear, but small wiggles may persist in the corridors identified. In some cases, the OD pairs can be well-served despite these deviations (as in Corridor 1), but in others, it makes it difficult to serve all OD pairs well (as in Corridor 4). Here, this distinction is evaluated after corridors are identified, to further restrict the set of corridors that are good candidates for new service.

Whether or not a corridor that meets the maximum distance and maximum angle constraints can be well-served is highly dependent on the road network. Chapter 3 introduced the challenge of spatial representation in this planning framework. Chapter 6 proposed methods for incorporating the road network within the corridor identification methodology. The implementation discussed in this chapter, used abstract linear representations of potential routes and euclidean distance to assess similarity. Instead, potential routes can be defined as actual paths through the road network, and distance between OD pairs and the potential route can be measured as road (or walking) network distance. These adjustments to the methodology would help ensure that a corridor can be served by a potential route and that OD pairs in the corridor are not separated from the potential route by a barrier or other discontinuity in the road network.

Incorporating geographic barriers and termini restrictions

Chapter 6 also discussed methods for accounting for geographical barriers. In the application to London, the Thames presents a major barrier. Dividing the network into two – one on either side of the Thames – eliminates the possibility of identifying corridors that cross the river. While corridors such as Corridor 6, that cross the Thames in a location where there is no bridge, are problematic, corridors that cross at existing bridges may be of interest. In the case of this application, a segment-level flow constraint, as discussed in Section 7.7.2, would be sufficient to eliminate the problematic version of Corridor 6, with OD pairs on either side of the river, and replace it with the portion of the corridor south of the Thames.

In the case of Corridor 1, smaller barriers including the Brent River and the Chiltern Railways tracks presented issues in terms of possible routing options. Chapter 5 presented two possible methods for accounting for barriers. Using the road network can implicitly account for barriers, which produce gaps in the road network, and can result in two points that are close in terms of euclidean distance being quite far in terms of road network distance. Using road network distance to assign OD pairs to potential corridors will therefore address the issue of these minor barriers.

The other possible method is to spatially define a set of barriers and require linear paths to bend to accommodate these barriers. Allowing these bent corridors may also identify corridors that are not identified in the implementation here, with strictly linear corridors. However, in a large network such as London, deciding what constitutes a barrier and how to spatially represent each one poses a challenge.

Another constraint on route design is the location of termini. For the three corridors analyzed in detail, the proposed route termini corresponded with existing route termini in all but one instance. However, it is possible that the methodology applied here will identify corridors without termini which are practical for bus service. Corridors can be adjusted or eliminated in post-analysis if this is the case. Chapter 6 described an alternative method in which a set of pre-defined termini are used to define the endpoints of provisional corridors from which the corridors with opportunities for new services are selected, based on expected benefits and flow, given a specified maximum corridor width. Such a method is more constrained than the existing method and is likely to identify fewer corridors given the same input OD pairs and flow and width parameters. However, in cases where termini locations are highly constrained, it increases the chances that the corridors identified are good candidates for bus service.

Chapter 8

Conclusion

This dissertation provides several methodological contributions to service planning in existing public transport networks. Broadly, it presents a new approach to sketch planning in public transport networks. It develops a framework for bus network sketch planning in an existing public transport network, and introduces new methodologies for conducting and aggregating origin-destination (OD) analysis. As demonstrated for the London bus network, the framework proposed here successfully identifies opportunities for new bus services. Contrary to ad hoc methods, alternatives are generated and evaluated systematically, which may reveal opportunities that would otherwise be overlooked. Still, there are some limitations to the OD-level analysis and aggregation methods presented here. In addition to summarizing the key contributions, this chapter discusses future work related to this topic including improvements to the proposed framework, related problems to address, and potential applications of the framework to solve other types of planning challenges.

8.1 Contributions

Overall, this dissertation develops a framework for bus network sketch planning that uses OD-level data. The framework is flexible, as parameters can be adjusted to suit a particular network, and it is applicable to complex, multi-modal, existing networks. As illustrated in the application to London's bus network, the framework is effective in identifying corridors for new bus services. It identified a reasonable number of corridors (4-16, depending on the parameters applied) from a vast number of zonal OD pairs and even more disaggregate journey data. Each element of the three-step method provides a distinct contribution. The first step defines zonal OD pairs that are suitable for this type of analysis. The second step identifies the subset of OD pairs that can be improved by a potential new bus route providing more direct services, and quantifies the expected OD-level benefits. The final step aggregates improvable OD pairs into corridors that are appropriate for bus service.

8.1.1 A new zone definition methodology

Transportation analysis and planning often uses zones defined for other purposes. For example, it is common to use postcodes and census tracts. Even zones specifically designed for transportation analysis may be defined based on residential or work locations or based on available demographic information. They often use roads as boundaries. For OD-level public transport analysis, of the kind proposed here, zonal OD pairs should group bus stops and rail stations that represent similar origins (destinations) so that multiple paths can be evaluated together. The methodology developed in this framework clusters bus stops and rail stations based on their coordinates, and can also incorporate ridership patterns by weighting stops and stations based on boardings or alightings. As a result, while many existing zonal schemes have stops or stations at the boundaries of zones, this method clusters stops and stations close to zone centroids. The resulting zones, which reflect the structure of the public transport are useful for analyzing and understanding demand and performance in public transport networks, and may be able to replace traditional zonal schemes for some public transport planning applications. They may be particularly useful for analysis of network resilience, as they group multiple paths that serve the same origin and destination.

8.1.2 OD level analysis of directness

The second step of the proposed framework consists of OD level analysis to assess the directness of OD pairs and the opportunities for improvement. Overall, OD analysis allows for a better understanding of travel in public transport networks, because it considers complete journeys, including multi-stage journeys. The methodology proposed in this dissertation classifies OD pairs as well-served by the existing network or not. This defines a standard for directness of service that is critical for the purposes of this framework but may also be useful outside the framework for evaluating and planning public transport networks. In addition, the methodology defines a threshold to classify OD pairs as improvable or not based on current and expected travel time. The expected travel time estimate, which assumes new direct service, enables the estimation of expected benefits of new service at the OD pair level. Benefits include travel time savings for existing public transport passengers and new journeys attracted to the network. The estimation of benefits is critical for the transition from understanding directness in the network to improving it through additional bus services.

8.1.3 Corridor identification methodology

Completing the transition from understanding to planning is the third step in the framework, the identification of corridors. OD pairs classified as improvable are clustered into corridors that meet the necessary characteristics for new bus services. In the public transport planning process, this is often referred to as generating alternatives. Traditional approaches in planning theory generate alternatives systematically but exhaustively, producing a very large set of alternatives to be evaluated. In practice, more ad hoc methods are likely to be applied, which may miss some opportunities. The corridor algorithm proposed here systematically considers all improvable OD pairs, but only identifies corridors that meet the shape and demand characteristics desirable for bus services. While post-analysis is still required to evaluate these corridors in more detail, the methodology produces corridors with a high likelihood of being appropriate for new services.

8.1.4 Impacts on public transport planning

Applying this framework can help public transport planners improve mobility in existing networks. It identifies corridors where new services can provide travel time savings to current public transport passengers and attract new passengers to the network. Moreover, it does so in a way that can be integrated into public transport planning processes in a straightforward way.

As discussed, traditional approaches to public transport network design use partial optimization methods to identify a set of optimal routes and are rarely used in practice. Partly, this is because planners do not seek full network redesigns as discussed in Chapter 1. However, the approach in this dissertation may be more appealing not only because it is designed for incremental improvements. It also is likely to be more intuitive to public transport planners. The parameters used at each step in the framework have tangible definitions and link to typical public transport analysis and planning considerations. For example, the zone size parameter can be linked to distributions of access distance. The target bus distance and expected travel time are recognizable to planners as service standards, as are parameters regarding bus route length and distance from a potential route to OD pairs that it serves. As shown in the sensitivity analysis in the application to the London network, planners can see how changes in parameters impact the recommended corridors.

While optimization approaches may appear to planners as black box, the framework presented here has straightforward steps. Intermediate results, such as the set of improvable OD pairs or the overlapping candidate corridors can be visualized and analyzed, allowing planners to easily follow the planning process proposed in this framework. Planners are generally familiar with the selection between alternatives and the planning of bus services at the corridor level. Thus, many aspects of the framework will be familiar to planners.

In addition, the framework may enable planners to make frequent, informed adjustments to bus services. In practice, network-level bus planning is attempted only infrequently, if at all, in many networks. Given planner-defined parameters and necessary inputs, the set of methodologies in this framework can be automated. As such, the framework can be used for network monitoring, providing OD and corridor-level evaluation and producing planning recommendations at regular intervals.

8.2 Future work

Throughout the dissertation, potential improvements to the methodologies within the framework were described, and they are reviewed here. In addition, the framework brings to light several related problems. More broadly, the general approach developed to use OD-level data for planning can be applied to other planning challenges. Some examples are described here.

8.2.1 Methodological improvements

Several potential improvements to the framework proposed here were discussed in the description of the OD analysis and corridor identification algorithm and in the context of the London application. In Chapter 3, two sets of methodological challenges were introduced. The first is the choice of spatial representation of the network and travel within it at each step in the framework. The second is the role of the existing network when planning new bus services. Methods for addressing these challenges were discussed throughout chapters 4, 5, 6, with a specific implementation applied to the London network, as described in Chapter 7. Here, opportunities to improve the ways in which these challenges are addressed are described.

Improving spatial representation

In the London application, spatial abstraction was used, particularly for corridor identification. While this implementation demonstrated the potential of the framework and produced some corridor recommendations that appear promising, the challenges related to geographic barriers and characteristics of the road network can be addressed by using more realistic representations of the road network in steps 1 and 3 of the framework.

In Step 1, bus stops and rail stations are clustered based on their coordinates. In the London implementation, euclidean distance was used to assess the similarity of stop and station locations. This ignores barriers and details of the road network. Using road network distance in the clustering of bus stops and stations will reduce the chances that two stops (stations) separated by a geographic barrier or other discontinuity in the road network are assigned to the same zone. This improvement is straightforward, requiring only a representation of the road network and a method for estimating the shortest path between two points through a network, many of which exist.

Chapter 6 discussed several ways to account for spatial properties of networks when defining potential routes and assigning OD pairs to routes, in Step 3 of the framework. In the application to London, potential routes are defined as straight lines, and the distance from a potential route to an OD pair is assessed as the euclidean distance. Similar to in Step 1, this could be improved by using the road network distance, or ideally walking network distance to assign OD pairs to potential routes. To do so requires the estimation of the shortest path from the origin (destination) point to the potential route, through the network provided. This requires the estimation of the shortest path to multiple locations along the route to identify the closest point on the route. Thus, the additional computational burden may be significant. However, the potential benefits are likely also significant; this approach will prevent OD pairs from being assigned to potential routes from which they are separated by barriers or gaps in the road network.

The potential routes themselves can also be represented as actual paths through the road network rather than as straight lines. This redefines the shape of potential corridors to one that is actually possible given the road network. Not only does this help ensure that potential corridors are good candidates for new service, but it may help identify corridors that are missed by defining corridors as strictly linear. Particularly when there are geographic barriers, some OD pairs cannot be served by a linear (or pseudo-linear path). By defining potential routes as actual paths through the road network, these non-linear routes will be generated. Work is left be done to define a method for generating potential routes using the road network. The set of potential routes should be sufficiently exhaustive to ensure that promising potential corridors are identified, but cannot be too large or the assignment of OD pairs to potential routes and selection of final routes will require massive computation.

As an alternative to using a full road network representation, Chapter 6 discussed the explicit modeling of barriers. Potential routes and the distance from OD pairs to potential routes would be modeled as a collection of line segments, angling to avoid barriers. This requires less computation than modeling the entire road network, but requires the definition of barriers, which may pose challenges. The definition of what constitutes a barrier is not unambiguous, particularly for minor barriers. Some barriers can be traversed on foot but not in a bus. In defining the network of barriers, it may be prudent to simplify those that have complex shapes. Determining whether to model the entire road network or to model barriers may depend on the data and computational means available. Exploring these alternatives can be the subject of future research.

Chapter 6 also discussed using an OD-chaining approach to define potential routes and corridors, instead of generating the route structure first and then assigning OD pairs. While there are opportunities to draw on density-based scan algorithms and on trajectory clustering work in particular, the distinct constraints on bus corridor shape and the requirement that full trajectories are clustered present distinct challenges. Thus, developing such an approach requires more research. If developed, it may represent a more efficient way to define an exhaustive set of promising potential corridors. That is, it may require fewer potential corridors be generated in order to generate all promising potential corridors, because potential corridors are limited to chains of improvable OD pairs.

Chapter 6 also discussed limiting potential routes to originate (end) at a predefined set of termini. This constrains the number of potential routes, which may make the incorporation of computationally-intense elements, such as the incorporation of the road network to define route shape and measure distance, more tractable. Finally, another straightforward improvement is to replace the corridor-level flow requirement with a segment level requirement. Not only does this guarantee a minimum level of flow along the entire corridor, but it may eliminate corridors that cross geographic barriers, as they often result in gaps in flow.

Overall, incorporating spatial complexity requires more inputs (such as the road network) and more computation. However, it can help ensure that the most promising corridors are identified. Alternatively, spatial details can be analyzed after corridors are identified. This post-analysis can be used to eliminate or adjust corridors that cannot be well-served by a single bus route due to spatial characteristics. However, leaving this assessment to post-analysis may result in corridors that are good candidates being missed. Also, it is important to note that even if abstract corridor shapes and euclidean distance are used for the definition of zones and corriodors, information about the road network is required in Step 2 for the definition of target bus distance and expected travel time. For the OD-level analysis of directness and improvability, using abstract (straight line) path definitions is insufficient.

Improving treatment of the existing network

In the framework, zonal OD pairs are defined based on the existing public transport network, and OD-level analysis uses information about available paths and travel times in the existing network. However, the impact of potential new routes on the existing network is not fully accounted for within the framework. Instead, decisions about existing services are left to post-analysis.

In the OD-level estimation of demand for new services, assumptions are made about demand on existing services. In the application to the London network, expected OD-level demand for a new service was estimated as current demand adjusted with an elasticity factor to reflect additional demand generated by a reduction in travel time. This reflects a simplifying assumption that all existing demand on an improvable OD pair will shift to the new service, and that the application of the elasticity factor estimates new journeys that previously were not made in the public transport network. At the OD level, improved demand assignment and prediction models will improve estimates of expected benefits and also result in more reliable flow estimates at the corridor level. While demand modeling is the subject of a large body of research, developing good demand models, particular at the zone-to-zone level, is a persistent challenge.

These improved estimates would also allow for better accounting of the impact of new routes on existing routes within the planning framework. The framework could potentially be extended to make recommendations about both the addition and the removal of bus routes. Given good OD level estimates of the expected shift of journeys from existing to new services, a very simple method for removing routes could take the following form.

1. Assume routes are added to recommended corridors with x benefits or more.
2. Given expected demand shits, eliminate or curtail existing routes where expected flow is below y

The outputs of the framework would be a set of corridors that are good candidates for new routes and a set of existing routes (or parts of routes) that are recommended for elimination. The process can be stopped after this single iteration, or steps 2 and 3 of the framework can be applied again assuming the network changes have been made. Research is needed to determine under what conditions such an iterative approach would reach an equilibrium, at which point no additional corridors for new service are identified. The methodology may borrow from optimization approaches, such as the ones used to address the Transit Network Design Problem. When developing a methodology for the removal of routes, it may be important to include certain restrictions, as some routes serve a critical population or serve an important role in the network regardless of flow.

8.2.2 Applying the framework to other planning challenges

The analysis in this dissertation, in particular the application to London's network suggests a need for additional services to improve OD pairs that are not well-served and are improvable as defined in the framework. In London's network, a small percentage of improvable OD pairs were assigned to the corridors identified. Overall, the improvable OD pairs were distributed throughout the network and therefore most were not along a corridor with sufficient density to merit a new bus route.

Often, sparser demand is served with on-demand services. On-demand services do not follow a fixed route of schedule. Recently, there has been renewed attention to on-demand services with the rise of transportation network companies and the development of autonomous vehicles. One possible extension of the framework developed here for bus sketch planning is to adapt it to be used to plan other modes.

The planning challenges for on-demand services are distinct from bus network planning. Significant research is focused on ride-sharing. This requires the grouping of OD flows, which is similar to Step 3 in the framework presented here. However, spatial aggregation for ride-sharing purposes must also account for the time of demand and generally must be accomplished in real-time.

Other planning questions for on-demand services may relate to which areas of a city are the best candidates for on-demand services. Or, opportunities for on-demand services to feed a public transport network can be analyzed. Unlike bus services, on-demand service generally serves passengers' true origins and destinations. Services can be flexible, but the level of service provided, in terms of waiting time and travel time depends on the fleet availability and operation.

Some aspects of the framework presented in this dissertation are applicable: the processing of OD level data to identify and quantify opportunities for improvement, and the spatial aggregation of travel information. Similar to predicting demand for bus service, demand for on-demand service may come from journeys currently served by public transport, or may include journeys that shift from other modes as well as latent demand. For spatial aggregation in the context of on-demand service planning, clusters may represent a region of the city that could be served by a fleet of on-demand vehicles. Or, they could consist of a set of OD pairs with a common destination, such as a rail station, where feeder bus service can be replaced with on-demand service.

In addition to adapting the framework for different modes, it could be adapted to address different planning challenges. One example is deciding where to add bus priority corridors. Priority measures can improve both speed and reliability, so this would require developing metrics to assess each of these. The aggregation would depend on whether planners were considering adding new routes or adding priority to existing routes. In the first case, a similar corridor aggregation methodology to the one proposed in this dissertation could be applied. In the latter case, the aggregation would need to account for existing routes.

A similar application is identifying opportunities for express service. The particular demand patterns that are generally associated with this type of service, such a long trip lengths, and high levels of concentrated OD-level demand would need to be identified through the OD analysis and spatial aggregation. Again, corridors could be identified from OD-level information to find groups of OD pairs that could be improved by a new express bus route.

The adaptable approach developed in the context of this research is an important contribution to public transport planning at this time. Disaggregate OD-level data is becoming more plentiful and is already supporting an improved understanding of travel behavior and system performance. The strategic definition of OD pairs, development of appropriate metrics, and spatial aggregation of disaggregate data allows planners to take the next step and use the improved understanding to improve planning.

Bibliography

Russell L Ackoff. From data to wisdom. *Journal of applied systems analysis*, 16(1):3–9, 1989.

Gennady Andrienko, Natalia Andrienko, Jason Dykes, Sara Irina Fabrikant, and Monica Wachowicz. Geovisualization of dynamics, movement and change: key issues and developing approaches in visualization research. *Information Visualization*, 7(3-4):173–180, 2008.

David Arthur and Sergei Vassilvitskii. k-means++: The advantages of careful seeding. *Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms*, pages 1027–1035, 2007.

M Hadi Baaj and Hani S Mahmassani. Hybrid route generation heuristic algorithm for the design of transit networks. *Transportation Research Part C: Emerging Technologies*, 3(1): 31–50, 1995.

Hugo Badia, Miquel Estrada, and Francesc Robusté. Bus network structure and mobility pattern: A monocentric analytical approach on a grid street layout. *Transportation Research Part B: Methodological*, 93, Part A:37 – 56, 2016.

Mousumi Bagchi and PR White. The potential of public transport smart card data. *Transport Policy*, 12(5):464–474, 2005.

Saeed Asadi Bagloee and Avishai Avi Ceder. Transit-network design methodology for actual-size road networks. *Transportation Research Part B: Methodological*, 45(10):1787–1804, 2011.

Kinan Bahbouh and Catherine Morency. Encapsulating and visualizing disaggregated origin-destination desire lines to identify demand corridors. *Transportation Research Record: Journal of the Transportation Research Board*, (2430):162–169, 2014.

Kinan Bahbouh, James R Wagner, Catherine Morency, and Chantal Berdier. Traclus-dl: A desire line clustering framework to identify demand corridors. In *Transportation Research Board 94th Annual Meeting*, number 15-3508, 2015.

Richard Balcombe, Roger Mackett, Neil Paulley, John Preston, Jeremy Shires, Helena Titheridge, Mark Wardman, and Peter White. The demand for public transport: a practical guide. Technical report, Transportation Research Laboratory, 2004.

Saurav Dev Bhatta and Matthew P Drennan. The economic benefits of public investment in transportation: A review of recent literature. *Journal of Planning Education and Research*, 22(3):288–296, 2003.

Maurizio Bielli, Massimiliano Caramia, and Pasquale Carotenuto. Genetic algorithms in bus network optimization. *Transportation Research Part C: Emerging Technologies*, 10 (1):19–34, 2002.

Bus Priority Team, Transport for London. Accessible bus stop design guidance, 2006.

Francesco Calabrese, Francisco C Pereira, Giusy Di Lorenzo, Liang Liu, and Carlo Ratti. The geography of taste: analyzing cell-phone mobility and social events. In *International Conference on Pervasive Computing*, pages 22–37. Springer, 2010.

S Carrese and S Gori. An urban bus network design procedure. In *Transportation planning*, pages 177–195. Springer, 2002.

Ennio Cascetta and Armando Cartenì. A quality-based approach to public transportation planning: theory and a case study. *International Journal of Sustainable Transportation*, 8(1):84–106, 2014.

Avishai Ceder. Operational objective functions in designing public transport routes. *Journal of advanced transportation*, 35(2):125–144, 2001.

Avishai Ceder. *Public Transit Planning and Operation: Theory, Modelling and Practice.* 0xford: Elsevier, 2007.

Avishai Ceder and Nigel HM Wilson. Bus network design. *Transportation Research Part B: Methodological*, 20(4):331–344, 1986.

Alfred Ka Kee Chu and Robert Chapleau. Enriching archived smart card transaction data for transit demand modeling. *Transportation Research Record: Journal of the Transportation Research Board*, 2063(1):63–72, 2008.

F Ciaffi, E Cipriani, and M Petrelli. Feeder bus network design problem: a new metaheuristic procedure and real size applications. *Procedia-Social and Behavioral Sciences*, 54:798–807, 2012.

Ernesto Cipriani, Stefano Gori, and Marco Petrelli. Transit network design: A procedure and an application to a large urban area. *Transportation Research Part C: Emerging Technologies*, 20(1):3–14, 2012.

Daniel Csikos and Graham Currie. Investigating consistency in transit passenger arrivals: insights from longitudinal automated fare collection data. *Transportation Research Record: Journal of the Transportation Research Board*, (2042):12–19, 2008.

Carlos F Daganzo. Structure of competitive transit networks. *Transportation Research Part B: Methodological*, 44(4):434–446, 2010.

Elizabeth Deakin and Songju Kim. Transportation technologies: Implications for planning. Technical report, UC Berkeley: University of California Transportation Center, 2001.

Elizabeth Deakin, Marianne Payne, and Val Menotti. Development of bay area rapid transit system expansion criteria and process. *Transportation Research Record: Journal of the Transportation Research Board*, (1887):18–25, 2004.

Sybil Derrible and Christopher Kennedy. Characterizing metro networks: state, form, and structure. *Transportation*, 37(2):275–297, 2010.

Thomas A Domencich and Daniel McFadden. *Urban travel demand - a behavioral analysis.* North-Holland Publishing Company, 1975.

Martin Ester, Hans-Peter Kriegel, Jörg Sander, Xiaowei Xu, et al. A density-based algorithm for discovering clusters in large spatial databases with noise. In *Proceedings of the second international conference on knowledge discovery and data mining*, volume 96, pages 226–231, 1996.

William L Garrison and Duane F Marble. Factor-analytic study of the connectivity of a transportation network. *Papers in Regional Science*, 12(1):231–238, 1964.

Fabien Girardin, Francesco Calabrese, Filippo Dal Fiore, Carlo Ratti, and Josep Blat. Digital footprinting: Uncovering tourists with user-generated content. *IEEE Pervasive computing*, 7(4), 2008.

Jason Gordon, Harilaos Koutsopoulos, Nigel Wilson, and John Attanucci. Automated inference of linked transit journeys in london using fare-transaction and vehicle location data. *Transportation Research Record: Journal of the Transportation Research Board*, 2343:17–24, 2013.

Steven Ronald Gordon. *Relationships between economies of scale and the shape of transportation networks.* PhD thesis, Massachusetts Institute of Technology, 1974.

Daniel J Graham. Agglomeration, productivity and transport investment. *Journal of transport economics and policy (JTEP)*, 41(3):317–343, 2007.

Greater London Authority. The london plan: The spatial development strategy for london consolidated with alterations since 2011. Technical report, Greater London Authority, 2016.

Greater London Authority. Mayor sets out bold plans to increase bus ridership across the capital. `https://www.london.gov.uk/press-releases/mayoral/mayor-sets-out-plans-to-increase-bus-ridership`, Mar 2017. Accessed: 2017-04-17.

JF Guan, Hai Yang, and Sumedha Chandana Wirasinghe. Simultaneous optimization of transit line configuration and passenger line assignment. *Transportation Research Part B: Methodological*, 40(10):885–902, 2006.

Valérie Guihaire and Jin-Kao Hao. Transit network design and scheduling: A global review. *Transportation Research Part A: Policy and Practice*, 42(10):1251–1273, 2008.

Zhan Guo. *Transfers and path choice in urban public transport systems.* PhD thesis, Massachusetts Institute of Technology, 2008.

Halcrow Fox and Associates, Accent Marketing & Research, and University of Leeds. Review and specifications of model elasticities. Technical report, Department of Transport, 1993.

Dick Hasselstrom. *Public transportation planning: A mathematical programming approach.* PhD dissertation, Univ. of Göteborg, Sweden, Dept. of Business Administration, 1981.

Danny Holten and Jarke J Van Wijk. Force-directed edge bundling for graph visualization. In *Computer graphics forum*, volume 28, pages 983–990. Wiley Online Library, 2009.

Jie Huang and David M Levinson. Circuity in urban transit networks. *Journal of Transport Geography*, 48:145–153, 2015.

Chih-Chieh Hung, Wen-Chih Peng, and Wang-Chien Lee. Clustering and aggregating clues of trajectories for mining trajectory patterns and routes. *The VLDB Journal—The International Journal on Very Large Data Bases*, 24(2):169–192, 2015.

Kjell Jansson and Bosse Ridderstolpe. A method for the route-choice problem in public transport systems. *Transportation science*, 26(3):246–251, 1992.

Jae-Gil Lee, Jiawei Han, and Kyu-Young Whang. Trajectory clustering: a partition-and-group framework. In *Proceedings of the 2007 ACM SIGMOD international conference on Management of data*, pages 593–604. ACM, 2007.

Herbert S Levinson. Analyzing transit travel time performance. *Transportation Research Record: Journal of the Transportation Research Board*, 915:1–6, 1983.

Stuart Lloyd. Least squares quantization in pcm. *Information Theory, IEEE Transactions on*, 28(2):129–137, 1982.

James MacQueen. Some methods for classification and analysis of multivariate observations. In *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability*, volume 1, pages 281–297. Oakland, CA, USA., 1967.

Oded Maimon and Lior Rokach. *Data mining and knowledge discovery handbook*, volume 2. Springer, 2005.

Harvey J Miller, Frank Witlox, and Calvin P Tribby. Developing context-sensitive livability indicators for transportation planning: a measurement framework. *Journal of Transport Geography*, 26:51–64, 2013.

Catherine Morency, Martin Trepanier, and Bruno Agard. Measuring transit use variability with smart-card data. *Transport Policy*, 14(3):193–203, 2007.

Mirco Nanni and Dino Pedreschi. Time-focused clustering of trajectories of moving objects. *Journal of Intelligent Information Systems*, 27(3):267–289, 2006.

Peter Newman and Jeffrey Kenworthy. *Sustainability and cities: overcoming automobile dependence*. Island press, 1999.

ManWo Ng and Hong K. Lo. Robust models for transportation service network design. *Transportation Research Part B: Methodological*, 94:378 – 386, 2016.

Thomas Alexander Sick Nielsen and Henrik Harder Hovgesen. Urban fields in the making: new evidence from a danish context. *Tijdschrift voor Economische en Sociale Geografie*, 96(5):515–528, 2005.

Thomas Alexander Sick Nielsen and Henrik Harder Hovgesen. Exploratory mapping of commuter flows in england and wales. *Journal of Transport Geography*, 16(2):90–99, 2008.

Donald J Patterson, Lin Liao, Dieter Fox, and Henry Kautz. Inferring high-level behavior from low-level sensors. In *International Conference on Ubiquitous Computing*, pages 73–89. Springer, 2003.

Marie-Pier Pelletier, Martin Trépanier, and Catherine Morency. Smart card data use in public transit: A literature review. *Transportation Research Part C: Emerging Technologies*, 19(4):557–568, 2011.

Alasdair Rae. Flow-data analysis with geographical information systems: a visual approach. *Environment and Planning B: Planning and Design*, 38(5):776–794, 2011.

Zong Hao Rao, Ping Zhang, and Jin Ping Guan. Urban traffic corridors determination based on characteristics of od data-nanning city as an example. In *Advanced Materials Research*, volume 403, pages 2445–2448. Trans Tech Publ, 2012.

Steve Robinson. Measuring bus stop dwell time and time lost serving stop with london ibus automatic vehicle location data. *Transportation Research Record: Journal of the Transportation Research Board*, (2352):68–75, 2013.

Peter J Rousseeuw. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *Journal of computational and applied mathematics*, 20:53–65, 1987.

Anita Schöbel. An eigenmodel for iterative line planning, timetabling and vehicle scheduling in public transportation. *Transportation Research Part C: Emerging Technologies*, 74:348 – 365, 2017.

Yousaf Shah, Kevin Manaugh, Madhav Badami, and Ahmed El-Geneidy. Diagnosing transportation: Developing key performance indicators to assess urban transportation systems. *Transportation Research Record: Journal of the Transportation Research Board*, (2357): 1–12, 2013.

Robert J. Shapiro, Kevin A. Hassett, and Frank S. Arnold. Conserving energy and preserving the environment: The role of public transportation. Technical report, American Public Transportation Association, July 2002.

Kenneth A. Small and Clifford Winston. The demand for transportation: models and applications. In *Essays in transportation economics and policy.* Brookings Institute, 1999.

Leon Stenneth, Ouri Wolfson, Philip S Yu, and Bo Xu. Transportation mode detection using mobile phones and gis information. In *Proceedings of the 19th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, pages 54–63. ACM, 2011.

Ran Tao and Jean-Claude Thill. Spatial cluster detection in spatial flow data. *Geographical Analysis*, 48(4):355–372, 2016.

Transport Committee of the London Assemply. Bus services in london. Technical report, Greater London Authority, 2013.

Transport for London. London Underground. `https://tfl.gov.uk/corporate/about-tfl/what-we-do/london-underground`. Accessed: 2017-04-04.

Transport for London. Transport for london budget 2017/18. Technical report, Transport for London, 2017.

Martin Trépanier, Catherine Morency, and Bruno Agard. Calculation of transit performance measures using smartcard data. *Journal of Public Transportation*, 12(1):5, 2009.

David Uniman, John Attanucci, Rabi Mishalani, and Nigel Wilson. Service reliability measurement using automated fare card data: Application to the london underground. *Transportation Research Record: Journal of the Transportation Research Board*, (2143):92–99, 2010.

Mariko Utsunomiya, John Attanucci, and Nigel Wilson. Potential uses of transit smart card registration and transaction data to improve transit planning. *Transportation Research Record: Journal of the Transportation Research Board*, 1971(1):119–126, 2006.

Catherine Elizabeth Vanderwaart. Planning transit networks with origin, destination, and interchange inference. Master's thesis, Massachusetts Institute of Technology, 2016.

Anthony J Venables. Evaluating urban transport improvements: cost–benefit analysis in the presence of agglomeration and income taxation. *Journal of Transport Economics and Policy (JTEP)*, 41(2):173–188, 2007.

Cecilia Viggiano, Haris N Koutsopoulos, Nigel HM Wilson, and John Attanucci. Journey-based characterization of multi-modal public transportation networks. *Public Transport*, pages 1–25, 2015.

Cecilia Viggiano, Haris N Koutsopoulos, John Attanucci, and Nigel HM Wilson. Inferring public transport access distance from smart card registration and transaction data. *Transportation Research Record: Journal of the Transportation Research Board*, (2544):55–62, 2016.

VR Vuchic and A Musso. Theory and practice of metro network design. *Public transport international*, 40(3/91), 1991.

Mark Wardman. Review and meta-analysis of uk time elasticities of travel demand. *Transportation*, 39(3):465–490, 2012.

Peter R White. *Public transport: its planning, management and operation.* UCL Press Limited, 1995.

Lúis G Willumsen and Juan de Dios Ortuzar. *Modelling transport.* John Wiley & Sons, 2011.

Nigel HM Wilson, Jinhua Zhao, and Adam Rahbee. The potential impact of automated data collection systems on urban public transport planning. In *Schedule-based modeling of transportation networks*, pages 1–25. Springer, 2009.

Yadan Yan, Zhiyuan Liu, Qiang Meng, and Yu Jiang. Robust optimization model of bus transit network design with stochastic travel time. *Journal of Transportation Engineering*, 139(6):625–634, 2013.

Chaowei Yang, David W Wong, Ruixin Yang, Menas Kafatos, and Qi Li. Performance-improving techniques in web-based gis. *International Journal of Geographical Information Science*, 19(3):319–342, 2005.

IO York. Factors affecting bus-stop times. trrl project report 2. Technical report, Transport and Road Research Laboratory, Crowthorne, 1993.

Linda Zemotel and David Montebello. Interregional corridors: Prioritizing and managing critical connections between minnesota's economic centers. *Transportation Research Record: Journal of the Transportation Research Board*, (1817):79–87, 2002.

Yu Zheng, Quannan Li, Yukun Chen, Xing Xie, and Wei-Ying Ma. Understanding mobility based on gps data. In *Proceedings of the 10th international conference on Ubiquitous computing*, pages 312–321. ACM, 2008.

Xi Zhu and Diansheng Guo. Mapping large spatial flow data with hierarchical clustering. *Transactions in GIS*, 18(3):421–435, 2014.

Yiwen Zhu. Passenger-to-train assignment model based on automated data. Master's thesis, Massachusetts Institute of Technology, 2014.