MSc. Thesis

Integrated Line Planning and Decision Making for the Electrification of Bus Transport Systems

MSc. Transport, Infrastructure & Logistics

Delft University of Technology

Paul Simon Schön

July 4, 2024

Abstract

To aid the efforts of electrifying public bus systems and to reduce the impact public transit has on emissions in cities, it is necessary for planners to be well informed when making the decision to electrify the transport system. The methodological approach shows that there is value in considering the development of a new transport system over simply electrifying an existing transport network, as there is a trade-off between investment characteristics and passenger satisfaction. Significant cost savings could be realized by jointly considering line planning, infrastructure location, vehicle, battery and fleet size. This has been investigated by developing a decision making model that selects a lineplan seeking to maximize passenger satisfaction, whilst taking into account, vehicle scheduling, battery capacity and charging station location. The model demonstrated that there is a clear trade-off between these aspects and that a multitude of pareto optimal solutions that can be obtained, depending on weighting that decision-makers assign to the objectives. The simplicity of turning the objective of this methodology into a revenue-maximization function, should make this model a suitable starting point for more serious modelling of investment decision in the future.

Contents

Chapter 1

Introduction

1.1 Background

Considering the impact that public transportation has on carbon dioxide emissions, with 16% of carbon dioxide emissions originating from the road sector (Electric Vehicles, IEA, 2022)[17], electrification of road vehicles will be an increasingly important necessity for meeting emission goals, like the EU emission targets for 2030 and 2050. This involves the realm of public transport. The IEA sees potential in the electrification of bus transit systems not only in reducing total carbon dioxide emissions but also in the mitigation of air pollution and the creation of jobs, estimated to be 30% higher than would be achieved by simply building more roads. In order for this transition to take place, the availability of corresponding charging infrastructure is critical. Finally battery electric bus systems (BEB) offer a destinct advantage in comparison to rail as they require less investment as well as construction time and therefore have the potential to make a fast positive impact, even in large urban areas (Sustainable Recovery, Transport, IEA (2020))[16]. If the charging system considered is similar to the TOSA Bus system, in development by Carrosserie Hess and ABB Sécheron, short charging is possible at stops without the need for long charging breaks or battery replacement. This can make fully electric bus systems a viable consideration for municipalities seeking to run fully electric rapid transit systems as soon as possible.

Developing a suitable approach for infrastructure investment is going to be difficult. This is mainly due the many aspects that contribute to the cost of changing from an already existing bus transport network to an electric powered one. To no suprise, changing the main drive influence factors such as power and range, potentially impacting capacity and scheduling, but also other aspects that may be relevant to decision makers like safety, lifetime or disposal. Challenging is also the interplay between these aspects and how they affect the bus transport system as a whole. Considering that the bus fleet will constitute a significant part of the total investment cost, it is relevant to know how long the new fleet will last, how quickly the batteries degrade, how routing affects these characteristics. How route length, vehicle size and passenger capacity affect depth of discharge and battery

Figure 1.1: TOSA Energy Transfer System;
Oliver.auge,
https://creativecommons.org/licenses/by-sa/3.0/,
retrieved from: `https://en.wikipedia.org/wiki/TOSA_(bus)`

degradation over time. A similar situation presents itself when considering possible fast charging stations, as these would another aspect to the investment costs. When considering the interplay of fast charging stations at stops paired with battery electric vehicles these interactions between system components become increasingly more difficult to consider simultaneously.

It is therefore not suprising that the research community has therefore already started investigating the role of battery electric vehicles and their suitability in bus electric transport systems. When considering the research in the quantitative modelling field, it can be said that the characterisitics of batteries have recieved a lot of attention. Similarily whilst the field of lineplanning is fairly old in itself, the majority of novel research focussing on new methodology aims to produce methodologies with additional aspects into consideration, for example by including passenger choice. The field has therefore remaines disassembled and porous, and crevices between various fields or study remain.

Electrification of bus lines will require careful consideration of infrastructure decision making to avoid a balooning of costs. Operations research has proven suitable in guiding decision makers in the planning process in many domains including the realm of public transportation. Whilst electrification of bus transport infrastructure has been considered in several papers, including location planning of charging stops, fleet and battery sizing, few papers exist that investigate whether transport networks need to be redesigned or merely adapted to accommodate battery electric busses. Therefore interesting is still the effect that line planning for a whole bus network can have on these costs.

This thesis aims to address one of these gaps. Namely to determine how the change to battery electric vehicles impacts considerations related to the line plan of the bus transport system. When considering this gap several questions emerge. Is there a destinct advantage of running busses with larger batteries compared to smaller batteries capable of quick charging? Is there an advantage when considering vehicle scheduling and charging station location jointly and how will this impact the passenger satisfaction of the resulting line plan? Can investment costs be minimized jointly through a methodological approach, what would be limitations of such an approach and is it feasible to extend the methodology in order to apply it to real-life situations? What would be the real-life usefulness for actual decision making? This thesis aims to address a few of them.

This thesis will be structured in the following way: the problem description and research design can be found in this chapter, chapter 2 will review the literature. Followed by the mathematical formulation in chapter 3, the model being summarized in section 3.2. Chapter 4 presents the results, in which the model is applied to a hypothetical node network. Conclusion and Recommendations can be found in chapter 5.

1.2 Problem Description

In order to capitalize on the potential of battery electric bus systems it is crucial for operators to have a comprehensive understanding of the costs involved. This research aims to develop a novel line planning methodology capable to investigate trade-offs between vehicle types and charging station locations as well as their impact on the line plans produced and the attractivity of these plans to potential users. Since the development of an attractive line plan is a difficult task in itself it should come to no suprise that taking additional aspects into account will constitute an additional challenge. An aspect that is difficult in line planning for example is determining line plan utilization, that is what the service frequency should be, or alternatively how many vehicles should serve a line. This depends on the attractivity of said line, more attractive lines need to transport more passengers and require a higher service frequency for a fixed vehicle size. Since a higher service frequency generally means more vehicles are required and more passengers means more energy is needed to transport the bus, battery electric busses clearly impact this aspect. Questions of vehicle size and battery capacity clearly become relevant when considering this aspect. The same is true when considering battery capacity, line length and charging station location. Assuming that vehicles start fully charged, maximum line length will depend on battery capacity as well as the number passengers, which depends on attractivity of the line, when no charging stations are placed at stops. The placement of charging stations can mitigate this issue, but will produce additional investment costs. Through transfers it is also possible to produce a line plan consisting only of short lines, however this will come with the downside of requiring more vehicles to operate the larger number of lines. A final additional challenge to this problem is the realization that the number of vehicles does not exactly correspond to the sum frequency of the lines. Since vehicles can move between lines and service multiple lines as long as they have enough time to move between lines in between, the number of vehicles can be greatly reduced through vehicle scheduling. Since the vehicle schedule depends on the number of available vehicles and since lines on which non operate cant be serviced, the vehicle schedule and line plan are tightly linked. This relationship is usually mediated through timetabling of vehicles, in which case the available vehicles as well as the line plan is already know. In this specific case neither is available to determine a timetable and the vehicle schedule and lineplan need to be determined simultaneously. The frequency of a line determines the number of times it is operated by a vehicle, whilst the the vehicle schedule determines if this is feasible. The fleet size will depend on the number of vehicles required to produce a schedule that can service all lines at the frequencies required.

To summarize; the costs associated with the electrification of a bus transport system are the following costs:

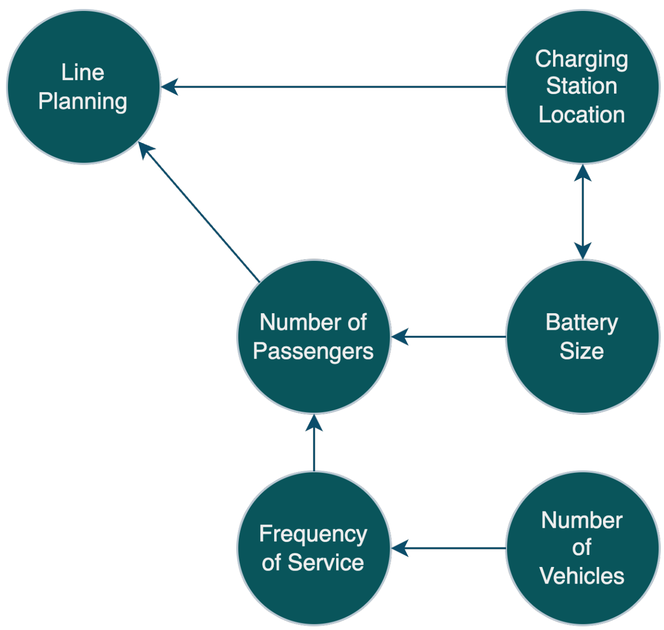


Figure 1.2: Problem Overview

- Charging Station location: depends on the chosen line plan and the battery size, as cost savings can be achieved by sharing charging stations between lines and also depends on the battery size as the charge between charging stations needs to be sufficient for passenger needs. Furthermore charging stations impact battery size and line length.

- Battery Size: This depends on the length of a line and the number of passengers on a given line and therefore the line plan. The more passengers are travelling a long a line, the more charge is used up. Serviceable line length may be constrained by battery size if no charging stations are placed. Finally varying battery sizes can have varying costs. For this research it is assumed that vehicle types are fixed beforehand and have a characteristic battery size and passenger capacity.
- Fleet size: This also depends on the number of passengers travelling along a line, as well as the vehicle types considered, as a given vehicle capacity affects the frequency of service. It also depends on the vehicle schedule and its compatibility with the line plan. Lines for which no vehicles can be scheduled, cant be serviced, no matter to demand. In order to best meet passenger demand a line plan and vehicle schedule need to be determined simultaneously.

These are the costs predominantly considered in this research. As can be seen line-planning has a significant impact on the costs associated with electrification, which is why developing a model to determine a line-plan is the main focus of this study.
An in depth investigation of possible user or service benefits associated with electrification are therefore not considered, although line-planning has the potential to improve service quality, for example by considering the number of transfers or waiting time. In order for a fair comparison of the model outcomes it will still be necessary to determine the service quality to some extent. Therefore the number of passengers using the network will be used as a main indicator for the attractiveness of the resulting line plan. These can be compared to existing network performance or constrained in the model to estimate costs given the network quality. The total number of users will be estimated based on the travel times in the network.
The following KPI's are therefore:

- Number of passengers
- Fleet Size
- Charging station investment cost
- Battery investment cost

1.3 Research Design

This research aims to investigate current line planning methodology and identify how it can be used to help decision making in planning electric bus transport networks. Furthermore it aims to develop a line planning model, where the battery size is determined jointly with the location of charging stations as part of the line plan, to investigate whether line planning models can be adapted to assess these additional aspects together.

With the introduction of electrification will we need to redefine the way new lines are constructed or is electrifying an existing line plan sufficient?

This research question can be explored by answering the following subquestions:

- How are line plans determined for public bus transport systems?
- What is the current state of the art in modelling of bus network electrification?
- In what way are these different modelling areas linked?
- What current models are suitable in developing a new methodology?
- What might such a new methodology look like?
- How could such a model be used to help decision makers plan battery electric bus transport systems?

To answer the subquestions of the research proposal the reviewed literature is going to be critically examined. This will form the starting point for developing a new model. The answers to these questions will form a guide towards a suitable selection of relevant models that can aid in the development process of a new methodological approach.

Research Sub-question	**Method**
How are line plans determined for public bus transport systems?	Literature Study
What is the current state of the art in modelling of bus network electrification?	Literature Study
In what way are these different modelling areas linked?	Literature Study
What current models are suitable in developing a new methodology?	Literature Study
What might such a new methodology look like?	Methodological Approach
How could such a model be used to help decision makers plan battery electric bus transport systems?	Review of Methodological Approach

The methodological approach will try to define a model structure capable of estimating a line plan that takes into consideration route choice, service frequency, fleet size and charging station location.

The model will have to make the following decisions, presented here in a brief overview:

- What lines l are selected as part of the line plan. A line is a simple path, without cycles, between stops in the network. A line corresponds to the path vehicles take in the network.

- Which route r are chosen from a set of routes C_{ij} that could be taken between origin i and destination j. If a route is chosen, this means that a line or set of lines needs to be available that facilitates travelling along this route. Routes need to be chosen such that they are attractive to potential passengers in order to maximize the usage of a new line plan.

- What frequencies are used on the lines chosen. This is required to estimate the number of vehicles required to satisfy the demand on the lines and is therefore a critical component of the electrification and operational costs, due vehicle and battery size considerations. This decision aspect makes vehicle scheduling and the timetabling pattern for the lines considered necessary, to get an accurate estimate of required fleet size.

- How vehicles are scheduled. The scheduled will depend on the frequency setting selected per line. The frequency setting will limit the number of total passengers travelling along a line, this limits the demand that can be satisfied. Therefore frequencies need to be set in such a way that they allow the maximum number of passengers on the most attractive routes. They themselves needs to be compatible with a possible vehicle schedule.

- What vehicle types will be used. This determines the battery size of the vehicles themselves. The model determines a single battery size used for all vehicles but should be capable to determine an optimal battery size from a set of varying sizes. Various vehicle sizes are not considered and the passenger capacity of vehicles is taken as fixed. This can be justified by most bus operators already operating the same kind of vehicle with a fixed passenger capacity.

- At what stops charging stations will be located. This depends on the depth of discharge of the battery at any given stop. This depth of discharge depends on battery characteristics of the vehicle type selected, the travel time between stops, and the passengers transported between stops.

An overview of a possible solution is presented in figure 1.3. This model will then be tested to identify possible short comings of this approach.

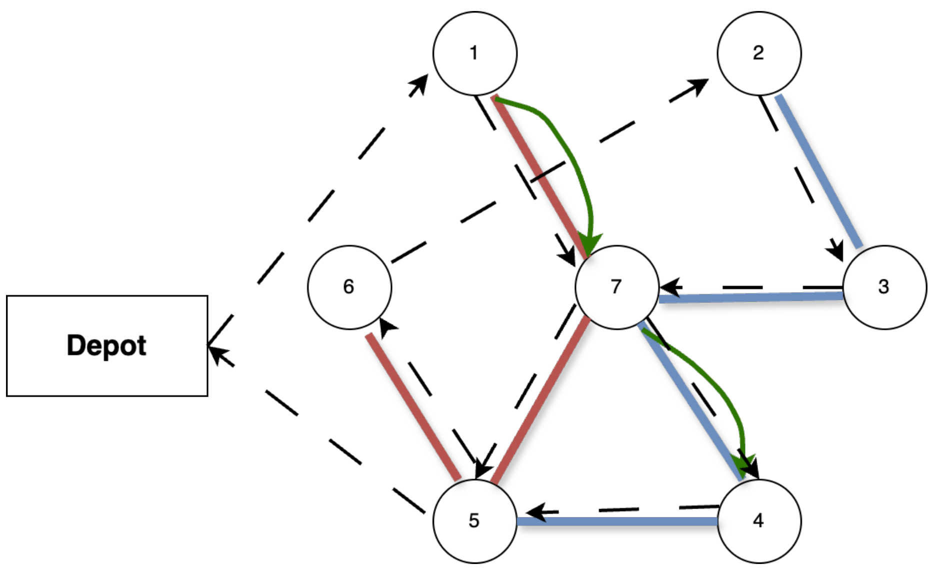


Figure 1.3: Network overview showing arcs for two lines (thick arcs in blue and red), a passenger route (green arrow) and a bus schedule (dotted arrow)

Chapter 2

Literature Review

2.1 Literature Review

The following paragraphs will give an overview of the literature reviewed. The body of literature can be divided into the following categories: previous research on electric bus transport systems, as well as models to aid decision making in relation to electric busses and charging stations, as seen in 2.1.1 and line planning as a wider domain; viewed from a strategic planning perspective, whilst considering various modelling approaches and their suitability in developing new methodology in 2.1.2.

2.1.1 Electrification and Battery Electric Busses

The application of operations research in the electrification domain can be segmented in several ways. Some aim to improve charging regimes, either via battery sizing based on fleet usage and vehicle routing regimes for optimal battery use, others try and optimally locate charging stations for battery electric busses (BEB) within a transport network. Since this area has only recently become a lot more relevant, the body of research is still relatively young.

Early studies have frequently focussed on feasibility and taken a data driven approach to estimate the energy demand, for example Sinhubera, Rohlfsa and Sauer (2012)[37] or Rogge, Wollny and Sauer (2015)[28] who assume that charging will take place at terminal stops only. Similarily Xylia et. al (2017)[41] assume in their study of the Stockholm bus network that electric charging of vehicles is most suitable at major transportation hubs and bus terminals. Combining geographic data analysis and an optimization model to determine energy balances they find a high concentration of electrification around major hubs and the potential to electrify a third of bus routes in the city.

In comparison Kunith et. al. (2016)[23] take the batteries charging behaviour into consideration and consider jointly charging station and transformer locations. They find, studying a subnetwork of the Berlin transport system, that the majority of charging locations are at intersections with other lines, due to the synergies with existing transformer infrastructure at metro stations. This model was extended the subsequent year in Kunith et. al. (2017)[22] to include battery size. Detailed simulation of battery usage has also been conducted by Sebastiani, Lüders, and Fonseca (2016)[32] locating charging stops for the city of Curitiba. An (2020)[1] considers the case of a bus network operator that aims to find the necessary fleet size of homogeneous electric vehicles and wants to locate the optimal loaction of charging stations in the network to maintain operations. In order to determine energy consumption all of these studies require data on driving regimes and timetables to estimate battery usage. A significant contribution to this area has been made by Sharif Azadeh, Vester, Maknoon (2022)[34]. In their research they propose a model for the electrification of a bus line, considering battery purchase cost, charging station installation cost and battery degredation costs jointly in a bi-objective formulation.

Alternatively authors have focussed on more operational aspects like charging regimes and vehicle scheduling. For example Wang, Huang, Xu and Barclay (2017)[40] schedule vehicle trips to charging stations based on an existing bus timetable, taking into account waiting and deadheading. He, Liu and Song (2020)[15] develop a model to schedule the fast charging of a battery electric bus system to minimize overall charging costs, for example due varying electricity prices due to varying demand. Yıldırım, and Yıldız (2021)[42] develop the Electric bus fleet composition and scheduling problem, in which various vehicle types and schedules are determined for a set of public transit trips to minimize operational costs. They show that electrification leads to considerable cost savings, even despite increasing the necessary fleet size.

The majority of papers therefore either focus on the strategic aspects, specifically infrastructure location or on the operational aspects like vehicle scheduling and driving regimes. What has recieved little attention is the role that network design and line planning has on the performance of an electric bus network. Despite the impact that demand has on energy expenditure or the cost savings possible by locating charging stops at line intersections, considering the wider transport network has recieved little attention in the literature so far.

2.1.2 Line Planning

Line planning is a common topic of interest in the transportation domain. In the "Handbook in Operations Research and Managment Science" Vol. 14 (C. Barnhart and G. Laporte, 2007), specifically Chapter 2 "Public Transit" Desaulniers and Hickman (2007)[9] give a general overview of the use of operations research in public transit planning. The areas of use are strategic planning, tactical planning, operational planning as well as real time control. All of these areas could be considered part of the line plan, however here line planning is considered part of the strategic planning horizon. Usually these problem perspectives are also considered in that order as they build on decisions made prior. In order to effectively design a battery electric bus system all of these aspects need to be evaluated. As was seen in the previous section, progress has already been made in the realm of infrastructure decision making as well as operational planning, whilst the role of line planning in electrification has recieved little attention, despite it being the first step in any bus network design. Due to its importance, line planning in itself has recieved significant attention in the past. Desaulniers and Hickman (2007)[9] give an overview over the progress made in the line planning and the history of the modelling approaches.

Finding a solution that freely generates lines as a series of archs between nodes is very challenging to solve. Magnanti and Wong (1984)[26], present a variety of mixed integer linear network models (like the commodity flow problem) and state that network sizes above 50 nodes are "extremely difficult to solve" suggesting these problems are NP-hard in general. This has also been demonstrated for several other problem formulations by other authors (Schöbel and Scholl, 2006[31]; Borndörfer, Grötschel, Pfetsch 2007[4]). Beyond complexity, difficulty also arises due to the magnitude of variables required to consider all possible arcs that can be used to generate lines, the service quality (travel time and frequency) and the resulting demand on the network. Developing a model which takes all factors into account simultaneously is very challenging.

Silman, Barzily and Passy (1974)[36] evaluate the generation of routes and service frequencies seperately. First the route network is developed from a skeleton network and subsequently the service frequency is determined based on route length, capacity and total fleet size. First routes offering the best travel times between a given OD pair are determined, then from these, the final routes are selected based on the frequency selection and fleet allocation constraints. This modelling structure is very practical and has found its way into other recent models. Furthermore the authors highlighted that its quick run-time makes it a suitable tool for planners.

Ceder and Wilson (1986)[6] developed another approach that constructs routes sequentially by seeking to minimize excess travel time, defined as the travel time between OD pairs plus the transfer time and minus the shortest path travel distance. The procedure is subject to route length bounds, and maximum number of routes and minimum service frequency constraints. Feasible routes are found by taking a set of terminal nodes and eliminating infeasible routes via a breadth-first search. This results in the construction of the network link by link and gives the decision maker the opportunity to adjust constraints during the process, however may be difficult to use effectively in large scale networks. The work was further extended by Israeli in 1995[18]. This approach was able to solve a network of 8 nodes and 14 links, however has not been applied to networks of larger size.

Another alternative method was produced by van Nes, Hamerslag, Immers (1988)[39]. This model was able to solve a Dutch network with 182 nodes over 115 zones with 8 routes. Here routes and frequencies are determined simultaneously by maximizing the number of direct trips subject to the maximum fleet size. The number of trips made is determined with use of a gravity function. The model has a non-linear objective function and many variables and therefore depends on the use of heuristics to solve. The heuristic used to solve this problem, determines how much increasing the frequency on one route improves the objective function. The frequency is then increased on the routes that have the most impact until budget and vehicle constraints are met. The model is reported to be capable of solving networks of up to 250 nodes, 150 zones, and 750 possible routes, according to van Nes (1988)[39].

Another segmentation of the line planning field is presented by Schöbel (2012)[30]. Here the problems are categorized by their objective functions and feasibility of the lines. Regarding the objective function, models either seek to minimize the cost of the operation or maximize the number of travellers or direct travellers. These models frequently include constraints on the minimum level of service as well as capacity. Schöbel also highlights that considering the number of "real passengers" along an edge frequently lead to a "chicken-egg" problem as passenger behaviour depends on the line plan and the line plan depends on the number of passengers travelling along a line. Furthermore the field can be segmented based on the restrictions made for the lines chosen. Either lines are selected from a set of potential lines, termed a line pool, or lines are constructed as part of the optimization process. An simple model determining the most convenient lines out of a line pool, subject to budget constraints was developed by Schöbel and Scholl (2006)[31]. Both simultaneous and prior generation of lines rely on heuristics to determine the lines, the first for constructing the line pool, the second for determining the lines during the optimization procedure. The latter requires certain rules or assumptions to prevent the formations of lines that are too long or too short, to ensure network connectivity and prevent sub-tours. An overview of this can be found in Gattermann, Harbering, Schöbel (2017)[12].

Heuristics have recieved significant attention in solving the line planning problem. Several heuristic solution approaches and meta-heuristic approaches have been developed and have proven successful. An overview of many of these can be found in Desaulniers and Hickman (2007)[9], as well as Schöbel (2012)[30]. Specific notice should be taken of column generation approaches that have proven very successful at solving multi-level problems. These problem formulations are capable of considering the line plan at strategic level whilst also passenger flows, acknowledges the "chicken-egg" nature of line planning problems in general. They are capable of constructing lines dynamically taking passenger flows into consideration. Bussieck, Winter, Zimmermann (1997)[5]; Borndörfer, Grötschel, Pfetsch (2007)[4]; Nachtigall, Jerosch (2008)[27] and very recently Bertsimas, Ng, and Yan (2021)[3] are notable papers on column generation. Bertsimas, Ng, and Yan (2021)[3] also explain how their generation of lines depends on preprocessing of the edge set, to speed up the solution process. In order to avoid sub-tour elimination constraints in the subproblem, an adaption of the shortest path problem, the edges need to construct the lines need to be aligned with a geographic direction to prevent cycles. This reduces the number of considered edges and speeds up the subproblem.

More recently there has been an aim to include discrete choice modelling into the demand and route choice estimation of the modelling process. Klier and Haase (2015)[20] proposed a line planning model that takes a binary logit model to estimate demand for public transit and develop a line plan that optimizes the number of users. The model assumes that the routes established have sufficient capacity and makes no attempt at measuring utilization. Very recently Hartleb, Schmidt, Huisman, Friedrich (2023)[14] propose a line planning model incorporating mode choice of travellers based on the routes made available by the lines selected in the model. The goal is maximizing service provider revenue (number of travellers). This allows for a line plan capable of determining the passenger distribution along lines and achieved by considering the journey time between OD pairs, and selecting a set of routes with minimal journey time. This is very similar to the model of Silman, Barzily and Passy (1974)[36] and likewise even includes a penalty for transfers. This property is very useful for the cost of electrification as estimating the number of passengers is relevant for determining the size of a bus battery. In the model the linking between routes and infrastructure has a quite simple structure, similar to the model proposed by Schöbel and Scholl (2006)[31] and similarily depends on a line pool from which suitable lines are selected. A downside is that the model neither determines the service frequency along the lines, nor does any form of vehicle routing. Either of these is required to estimate the fleet size, and therefore the investment cost of electrification. Instead the model assumes that travellers distribute evenly across equally good routes and determines the capacity required, rather than incorporating a strict limit. Still this paper provides a useful starting point.

Frequency Setting and Passenger Assignment

Approaches like the ones of Silman, Barzily and Passy (1974)[36] or Ceder and Wilson (1986)[6] determine the line plan first and then the set the frequencies for the given lines latter. This is sequential approach makes the bus transport system design problem tractable. Generally speaking developing a line plan and finding frequencies jointly is quite challenging due to the combinatorial nature of the resulting problem. Ceder (1984)[7] suggests that the number of busses required for a given period be equal to the average maximum number of passengers travelling over the period divided by the capacity of the vehicle times the load factor. This is akin to a service frequency if the busses are spread out equally and the time period is representative of a periodic service. In comparison Han and Wilson (1982)[13] relate passenger flows to the bus capacity and service frequency, realising that the flows over a given time period cannot exceed the bus passenger capacity times the service frequency (the number of busses in a given period). In this way uncongested assignment, through passenger flow and frequency can be linked.

Timetabling and Vehicle Scheduling

Vehicle scheduling is placed in the realm of operational planning and often after timetabling according to Desaulniers and Hickman [9], whilst timetabling is considered part of the tactical planning process. The difference between the two can described as the difference between creating a fixed schedule for running times and frequencies that will persist for a given time horizon, compared to asset scheduling that uses this schedule and seek to minimize costs. The reason they need to be treated separately is because to determine fleetsize, simply considering the service frequency and number of routes is not sufficient. A given set of routes, taking a certain amount of time, requiring a given frequency, suggests the need for a fixed amount of busses, as suggested by the fleet size constraint in Han and Wilson (1982)[13], however it would be more accurate to say that it requires a fixed amount of bus arrivals. This is where timetabling comes into play, translating frequencies into bus arrivals. Subsequently these arrivals can be scheduled for an actual bus fleet, as scheduling may result in time saving and superior asset allocation, the required fleet size is actually less than simply assigning a bus to each route. Crucially this links frequency with fleet size.

Research on timetabling has mainly focussed on determining the timetable either in such a way that passenger waiting times are minimized or that interchanges between lines are feasible. For example Sheffi and Sugiyama (1982) [35] develop a schedule for a single line, such that waiting times are minimized based on arrival times. Salzborn (1980) [29], Knoppers and Muller (1995)[21] focus on developing timetables such that either transfer waiting time is minimized for passengers and Ceder et al. (2001)[8] develops an approach for a timetable that synchronizes routes. This approach is useful when buslines have already been generated and their intersection points are known. Ceder and Wilson (1986) [6], suggesting a model of the whole bus transport system design problem, are amongst the first to consider the relationships between lineplanning, timetabling and vehicle scheduling. Here timetabling depends on already given lines with frequencies, from which point arrival and departure times can be generated, which are then used in the vehicle scheduling. Importantly fleet size is considered, but only as a constraint and the methodology therefore does not consider determining fleet size in itself. Since then both timetabling and vehicle scheduling have recieved separate attention as independent elements of the bus transport design process.

Vehicle scheduling has focussed on scheduling a fleet of vehicles from a single depot to the timetabled trips, with the aim of reducing transportation costs incured during the operation of the vehicles. An example of the how the problem can be formulated is presented by Freling and Paixão (1995)[11]. Löbel (1998) [25] offers methodology for the multiple depot vehicle scheduling problem. Generally an extensive overview of scheduling problems can be found in Desrosiers et al. (1995)[10].

2.2 Results from the literature

How are line plans determined for public bus transport systems?

As can be seen in the literature review, there are several different methods of route generation. The first methodological approach, presented by the likes of Silman, Barzily and Passy (1974)[36] or Ceder and Wilson (1986)[6] seeks to generate lines sequentially based on various performance measures, with the service frequency being determined seperately from the line selection procedure. This can be contrasted with approaches aimed at optimizing some overarching performance measure that develops a line plan by selecting lines from a predetermined line pool, as described by Schöbel and Scholl (2006)[31].

Recent advancements in line planning for public transit systems have incorporated discrete choice modeling into the demand and route choice estimation process. Klier and Haase (2015)[20] introduced a line planning model that utilizes a binary logit model to estimate public transit demand and optimize the number of users. However, the model assumes routes have sufficient capacity and does not measure utilization.

More recently, Hartleb, Schmidt, Huisman, and Friedrich (2023)[14] proposed a line planning model that integrates mode choice of travelers based on available routes. The objective is to maximize service provider revenue by considering journey times between origin-destination pairs and selecting routes with minimal travel time. This approach is akin to the model proposed by Silman, Barzily, and Passy (1974)[36] and includes penalties for transfers.

What is the current state of the art in modelling of bus network electrification?

The current state of the art in the application of operations research in the electrification domain involves various approaches aimed at optimizing charging regimes and infrastructure placement for battery electric buses (BEBs) within transportation networks. Early studies, such as those by Sinhubera, Rohlfsa, and Sauer (2012)[37], Rogge, Wollny, and Sauer (2015)[28], and Xylia et al. (2017)[41], have focused on feasibility assessments and data-driven approaches to estimate energy demand and identify suitable charging locations, often concentrating on major transportation hubs and bus terminals.

In contrast, recent research has delved into more nuanced considerations, such as battery charging behavior and transformer infrastructure synergies, as demonstrated by Kunith et al. (2016, 2017)[23][22]. Detailed simulations of battery usage and charging stop locations, exemplified by Sebastiani, Lüders, and Fonseca (2016)[32] and An (2020)[1], have also contributed to understanding the complexities of BEB operations and infrastructure requirements.

Operational aspects, including vehicle scheduling and charging strategies, have garnered attention from researchers like Wang, Huang, Xu, and Barclay (2017)[40], He, Liu, and Song (2020[15]), and Yıldırım and Yıldız (2021)[42]. These studies aim to optimize charging schedules, minimize overall charging costs, and determine the optimal fleet composition and scheduling strategies for electric bus fleets.

A significant contribution to this area has been made by Sharif Azadeh, Vester, and Maknoon (2022)[34]. Their research proposes a model for the electrification of a bus line, considering battery purchase cost, charging station installation cost, and battery degradation costs jointly in a bi-objective formulation. This holistic approach integrates strategic and operational considerations, highlighting the importance of comprehensive planning for the successful electrification of bus networks.
However, despite advancements in strategic infrastructure placement and operational scheduling, the role of network design and line planning in optimizing electric bus network performance has received limited attention. Considerations such as demand impact on energy expenditure and cost savings from strategic charging stop placements at line intersections have been overlooked in the literature. Addressing these aspects could provide valuable insights into enhancing the efficiency and sustainability of electric bus networks.

In what way are these different modelling areas linked?

The level of discharge depends on the number of passengers, as well as the size of the battery. Therefore the selection procedure for electric busses and batteries depends on the usecase of these busses. In order to have an optimal use case, these assets need to be used in the most efficient manner, for example by carrying as many passengers as possible on each leg of the journey. Furthermore the location of charging stations also effects the size of the battery, as a better network of charging stations can allow for more frequent recharging and therefore smaller battery sizes.

These considerations therefore strongly depends on the bus load of a bus along a route, this passenger load depends on the attractiveness of the line plan on which these busses operate. Determining an attractive lineplan that simulatenously takes these electrification aspects into consideration, allows for an optimal and therefore cost effective transition of a bus network, away from fossil fuel usage.

What current models are suitable in developing a new methodology?

The approaches currently most promising appear to be the approaches of Schöbel and Scholl (2006)[31] combined with the passenger mode choice of models such as the one of Hartleb, Schmidt, Huisman, and Friedrich (2023)[14]. The reason for this is that the passenger flows as determined by the mode choice model can be used to determine battery charge levels, which can make use of the research under taken by other researchers.

Summary

Battery electric vehicles are a relatively new area of research and have recieved a lot of attention recently. The majority of this transport planning related research focusses on the location of charging stations or on the operational aspects of fleet operation. Important aspects considered are fleet size, number of charging stations placed and battery sizes selected, as these strongly influence the cost of infrastructure or fleet purchases. The focus of

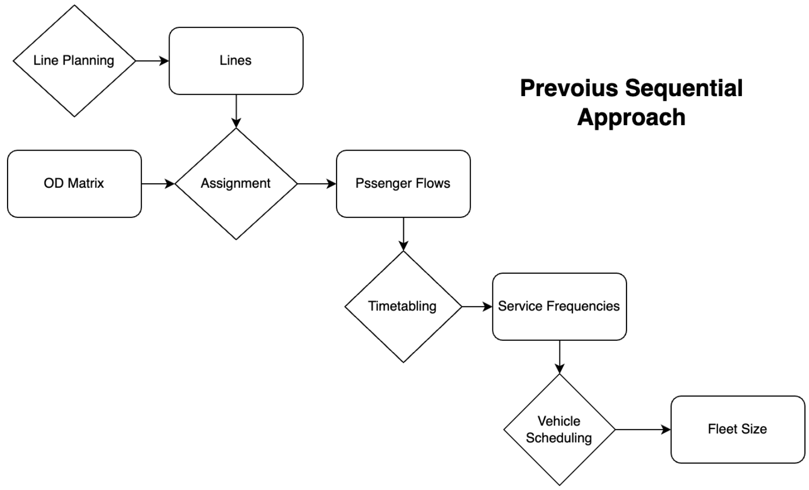


Figure 2.1: Previous approach in line planning

this literature can therefore be summarized as focussing on fleet and infrastructure location decision making. These studies require an existing timetable or line plan.

Line planning on the otherhand has been a well established field of research. This research aims at establish a line plan to either maximize the number of travellers willing to travel or reduce the cost given service requirements. A difficulty is determining lines and service frequencies jointly. The latter being required to establish operational costs. Recent research in the domain has focussed on producing more realistic models, for example by including passenger decision making through the use of discrete choice modelling.

Producing models at the intersection of strategic decision making and operational decision making has been a challenge to researchers. When considering BEB transport systems however this intersection is becoming especially relevant and has recieved no attention thus far. Developing a methodology for producing a line plan that can take the current state of the art in electrification into consideration, can prove to very useful to planners wanting to electrify their transport systems and should therefore recieve more attention in this field.

The previous approach as seen in the literature attempts to design a public transit system by sequentially solve various submodels and use the outputs in subsequent subroutines. This opens an opportunity for a new approach in which some of these components are evaluated simultaneously, potentially allowing for a better solution. The previous approach tries to create a line plan, followed by establishing the service frequency and time table, taking the OD matrix through passenger assignment. This service frequency can then be used to establish a timetable, which can be used for in vehicle scheduling to determine the fleetsize. Through this way, the design of a public transit system design process goes through various strategic and tactical models until a final design is found. After this the placement of electric charging infrastructure can be determined subsequently.

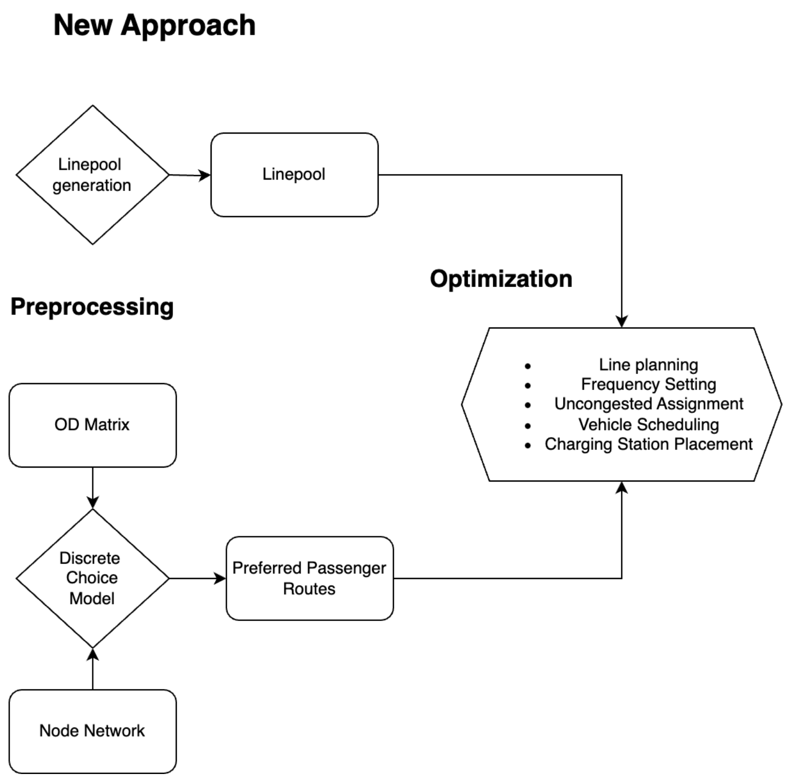


Figure 2.2: New approach

Alternatively a new strategic approach can attempt to establish a lineplane with already determined service frequency and vehicle scheduling simulataneously whilst also offering to assess the placement of charging stations in paralell. This new approach makes use of preprocessing a linepool and passenger route choice as is shown in the figure 2.2

Chapter 3

Methodology

3.1 Mathematical Formulation

In order to answer the following research question, a new methodology is proposed that describes the decision making model, capable of estimating a line plan that takes into consideration route choice, service frequency, fleet size and charging station location.

What might such a new methodology look like?

3.1.1 General Definitions

The aim of this model is to determine a line plan for an electric innercity bus transportation. This plan takes into account the passenger satisfaction, the number of vehicles required as well as the location of electric charging stations. The model seeks to optimize the system against the worst case scenario, namely during peak hours, for which it aims to establish a cyclical timetable as part of the line plan.

1. **General Settings** Passengers choose routes, available routes depend on the lines that exist. A line exists if it can be serviced by a bus schedule and the schedule needs to be cost efficient.

    - Let H denote the planning horizon, subdivided into discrete intervals. For each discrete interval, we represent it by t such that $t \in \{0, ..., T\} = H$.
    - While utilizing the existing locations of stops, our primary objective is to ascertain the bus line connecting them.
    - For every stop, we postulate the presence of an origin-destination demand matrix corresponding to each time interval.
    - Given a time horizon, typically chosen as a peak hour, we aim to establish a cyclical timetable for bus operations. By representing the bus timetable as a repeating pattern within this horizon, we can approximate daily operational activities and effectively capture the movement of the buses.

- Stops N. The network comprises a set of stops denoted by N, with each stop represented as $i \in N$. The depot is specifically noted as 0 within N.

- Let R be the set of potential routes, where each route r is defined by a sequence of stops: $r = \{s, \ldots, e\}$. Here, s denotes the starting stop and e the ending stop. If the length of a route exceeds a certain factor of the shortest distance between s and e, it's presumed unattractive to potential passengers and thus excluded from the set R. This max route length constraint acts as a cut-off to limit the time in preprocessing. This parameter can be changed depending on general distances and network sizes consider.

3.1.2 Route Choice and Passenger Demand

Routes are determined based on passenger demand, enabling the availability of attractive routes is key to developing a good line plan. This part of the model aims to enable attractive routes, by selecting feasible lines from the line pool.

The basic components of this part of the model are characterized by the modelling the following decisions:

Passenger Demand

By modelling the electrification of a metropolitan transport system use can me made of already existing data and network topography. It is assumed that the location of the stops will not change and only new lines joining these stops will be considered. Therefore all the stop locations are known and it is assumed that the demand of travellers willing to embark or disembark at a given stop is known as well. This can be formatted into an Origin-Destination demand matrix D_{ij}. This matrix encompasses all travellers travelling from origin i to destination j. It is also assumed that this demand will not change with a change in the line plan, which is not strictly true due to the induced demand effect, however is required to keep the model development tractable. Induced demand can mean that travellers choose to travel to alternative locations based on changed levels of accessibility regarding exchangeable destinations, for example by changes in the disutility of travel regarding alternative shopping or entertainment locations. A description of induced demand can be found in Lee Jr et. al. (1999) [24]. To include this level of analysis however, an indepth study of landuse would be required which is outside of the scope of this study. Furthermore the majority of travel happens between work and home, and whilst some residents might consider changing their work or home locations to better suit the new line plan, it is assumed that the majority of people will remain in there respective locations and that therefore the change in travel demand between stop locations will not change significantly with changes in the line plan.

Given that demand is fixed, it will still be necessary to estimate the satisfaction of the line plan produced, by estimating the number of passengers willing to travel with the new line plan. This is done with the use of a simple discrete choice model that estimates the probability of using a route between an origin and a destination. The use of discrete choice models for use in transportation planning has been well established, for example by the likes of Ben-Akiva et. al. (1985) [2]. By multiplying the probability of using a route with the travel demand between two locations can offer an estimate of the number of passengers willing to travel along a given route. If the model can produce a line plan that attracts similar numbers of passengers along a route as an already existing line plan then it is also reasonable to assume that the effect of induced demand will remain small. Overall the inclusion of this measure is used to estimate passengers satisfaction of the new line plan.

The utility of a route can be expressed in the following way:

$$U_r = \beta_t \cdot T_r + \epsilon \tag{3.1}$$

Where U_r is the utility of route $r \in C_{ij}$, β_t is a coefficient for travel time, T_r is the travel time along path r and ϵ the error term. Whether a route is attractive to passengers depends on the travel time for a given route. A route here is defined as a simple path, a path without cycles, between stops along the network. This results in the set C_{ij}, the set of all possible routes between origin i and destination j. The share of passengers for each origin and destination pair $ij \in OD$ is determined by use of a simple logit model:

$$P_{ij}^r = \frac{e^{U_r}}{\sum_{k \in C_{ij}} e^{U_k}} \tag{3.2}$$

The fraction of passengers P_{ij}^r for each route r $\in C_{ij}$ can therefore be determined before hand by finding all simple paths between all nodes in the network and computing the travel time and utility. The benefit of using this model structure to measure passenger satisfaction is also useful as it it can easily be extended later on to include other aspects that might have an impact on route choice or mode choice. As was done in the work of Johann Hartleb, Marie Schmidt, Dennis Huisman, Markus Friedrich (2022), this model does not explicitly consider the utility of alternative modes and simply combines their utility into one, namely the utility of not choosing the new lineplan. This means that this model could in the future also be extended to include pricing, by adding travel price to the utility funciton for example.

Passenger Demand with Mode and Route Choice

To further improve the measure of passenger satisfaction the demand estimation can be extended to include mode choice into the preliminary analysis. This can be achieved by replacing the simple logit model with a nested logit model instead. A different model becomes necessary when considering different types of modes as the independence of irrelevant alternatives no longer holds when considering routes chosen by varying modes. A reason for this is that correlations may exist between routes of the same mode, meaning that substitution of routes across varying modes are no longer independent.[38] As part of the generalized extreme value model family the nested logit model is a suitable choice to achieve this aim.

The nested logit model is suitable for this level of analysis as it allows clustering alternatives, in this case routes, in nests corresponding to the modes considered. The independence of irrelevant alternatives property will hold within a nest, therefore holds for the route choice considered for a single mode, but not for options between nests. Furthermore the nested logit model has the destinct advantage of offering a closed form solution for the choice probilities of routes considered, which simplifies assessing choice probabilities in preprocessing.
The new model structure proposed in this model is to include the impact of a single alternative, namely the car into this level of analysis. The car has the distinct advantage of offering direct travel or a direct route, which always correspond to the shortest path. Considering this as an option in the analysis will improve realism in determining the number of passengers willing to use the bus rapid transport system. The use of the nested logit also ensures that the choice of cut off for the route length doesnt actually impact route choice as significantly as may be assumed, as the number of passengers willing to choose long routes will be low in comparison to shorter and more direct routes. A schematic diagram of the nested logit structure can be seen in figure 3.1.

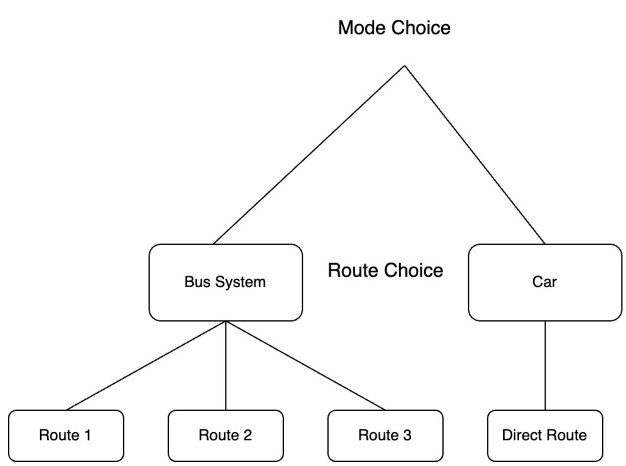


Figure 3.1: Diagram of the Nested Logit Model Structure

The choice probabilities can then be calculated by the following equation:

$$P_{ij}^{car} = \frac{e^{U_{car}/\lambda_{car}}(e^{U_{car}/\lambda_{car}})^{(\lambda_{car}-1)}}{e^{U_{car}/\lambda_{car}} + (\sum_{k\in C_{ij}} e^{U_{bus,k}/\lambda_{bus}})^{\lambda_{bus}}} \tag{3.3}$$

$$P_{ij}^{bus,r} = \frac{e^{U_{bus,r}/\lambda_{bus}}(\sum_{k\in C_{ij}} e^{U_{bus,k}/\lambda_{bus}})^{(\lambda_{bus}-1)}}{e^{U_{car}/\lambda_{car}} + (\sum_{k\in C_{ij}} e^{U_{bus,k}/\lambda_{bus}})^{\lambda_{bus}}} \tag{3.4}$$

Estimating Passenger Interest

The estimated interest in the line plan depends on the routes that can be served by the line plan. This depends on the routes $r \in C_{ij}$ selected for each origin destination pair i, j. Interest can be defined as the sum of the product of total demand between and the probability of each route for each origin-destination pair i, j.

- $max \sum_{ij\in OD} \sum_{r\in C_{ij}} D_{ij} P_{ij}^{r} y_r$

Where y_r is a binary variable, whether route $r \in C_{ij}$ is selected, D_{ij} a parameter indicating the total demand over the time horizon and P_{ij}^{r} the probability of selecting route $r \in C_{ij}$ between each origin-destination pair i, j.

3.1.3 Line planning problem

Solving the line planning problem is the core goal of this methodological approach. Here line planning is considered as the tasks revolving around the strategic aspects of a line plan. Specifically this part of the model aims to determine the paths that an established service would take, to serve stops along the way. Or in other words, what stops are part of a given line and what lines will be included in the model.

Another aspect that is considered in the line planning problem, is the selection of a service frequency. As has been seen in the literature review, jointly determining a line-plan and service frequency is quite a challenge. Therefore the line planning problem has taken an approach guided by both the works of Silman, Barzily and Passy (1974)[36] as well as Schöbel (2012)[30]. Here, the approach of Schöbel is taken, and used to generate a line pool. From this pool lines are selected, based on how attractive they are in enabling routes that are attractive to passengers. Reminiscent of Silman, Barzily and Passy's approach the service frequency is then used to select for lines out of this pool. However, instead of starting with a skeleton network, the entire network is considered and line selection is done jointly with frequency selection by "pruning" the set of lines used in the model. This can be done with a set partitioning approach, in which we select a frequency for every single line in the line pool. Lines that are not selected will recieve a frequency of 0. This selection process is done with the help of a binary variable that is equal 1 for a given combination of a line, service frequency and service pattern. Only one such combination is viable for each line.

Moreover, the scalability of the problem can be regulated by imposing constraints on the line pool itself. For instance, constraints such as minimum or maximum running times can be introduced to influence the size and composition of the line pool, thereby introducing a level of control in the computational complexity of the optimization process. Overall by managing to integrating these elements, this approach offers a systematic framework for addressing the difficulty of simulataneously selecting lines and frequency.

Line selection

- We assume that we enumerate all possible bus lines. However, in the computational results we can apply the following heuristic to limit the number of lines.
- Algorithm GenerateBusLines(start, end, graph):
  1. Initialize an empty list: feasibleRoutes.
  2. Use a search algorithm (e.g., Depth First Search) to find all paths between start and end.
  3. For each path:
  a. If it satisfies the directness criterion, continue; else, discard.
  b. If the number of stops is within the defined range, continue; else, discard.
  c. If it serves key areas or has a good distribution of stops, continue; else, discard.
  d. If it doesn't overlap significantly with existing routes, continue; else, discard.
  e. If its duration is within acceptable limits, add to feasibleRoutes.
  4. Return feasibleRoutes.

- We model the bus line as an ordered set of stops. Each one-way line, l, is characterized as an ordered set of stops: $l \in L = \{s, \ldots, e\}$.

Frequency Setting

- $\sigma_l^{fp} \in \{0, 1\}$: A binary variable that is set to one if line $l \in L$ is chosen for service with frequency f and service pattern p.

1. We assume that each selected line has an evenly distributed service.
2. Frequency shows the starting time of the first service after the beginning of the horizon period.
3. Note: Frequency shows the total number of services for the entire planning horizon.
4. For any given frequency, the starting time can either coincide with time zero or align with the commencement of the second run. If the latter option is chosen, the final service might extend beyond the planning horizon, which is acceptable.

5. For clarity, we refer to the combination of frequency and starting time as the "service pattern," denoted by p.

Related Constraint: This constraint ensures that for each line l only one service pattern p with corresponding frequency f is selected. If a line isn't chosen, its service frequency is set to zero, (for which there is only one service pattern).

- $\sum_{f \in F} \sum_{p \in P_f} \sigma_l^{fp} = 1 \quad \forall l \in L$

3.1.4 Linking Route Choice and Line Selection

This section details how the set partitioning regime for the line planning model can be connected with the selection, affected by the nested logit model. The result will be a line plan that selects lines that facilitate routes that maximize passenger satisfaction.

The line selection depends on the two previously definied σ_l^{fp} and y_r as well as the following variable:

- Let $w_{ij}^l \quad \forall \quad i \in N, j \in N, l \in L$: A continuous variable taking the value total passengers over the time horizon between stops i and $j \in N$ on line l.

Related Constraints: The first constraints ensure that if stops i, j are not part of a line l, w_{ij}^l is set to zero. The second that if i, j are not subsequent stops on a line l, w_{ij}^l is set to zero. Finally w_{ii}^l is also set to zero.

- $w_{ij}^l = 0 \quad \forall l \in L, \forall i, j \in N \nsubseteq l$
- $w_{ij}^l = 0 \quad \forall l \in L, \forall i, j \in N \subseteq l | i = l_n \wedge j = l_{n+1}$
- $w_{ii}^l = 0 \quad \forall l \in L, \forall i \in N$

Further more the follwoing parameter:

- α_i^r: A binary parameter set to one if station i is included in route r. This parameter will be used to ensure that the sum of all stops is covered in the model. It is assumed that this is a requirement for the line plan.

Related Constraints: Every station must be part of at least one route served.

- $\sum_{l \in L} \alpha_i^r y_r \geq 1 \quad \forall i \in N$

Considering the following constraints, route choice can be linked with line selection:

- $\sum_{l \in L_r} w_{ij}^l = \sum_{r \in R_{ij}} D_{ij} P_{n=1,ij}^r y_r \quad \forall i, j \in N$

This constraint sums all w_{ij}^l for each line $l \in L_r$, a subset of the line pool, all lines that cover all routes r, which pass stops $i, j \in N$ and equates this value to the total interest in routes $r \in R_{ij}$ passing by stops $i, j \in N$. This results in the model deciding how many passengers are carried on each line, between any two stops, based on the routes chosen.

- $\sum_{l \in L_r} \sum_{f \in F} \sum_{p \in P_f} f \sigma_l^{fp} \geq y_r \quad \forall i, j \in N, r \in Cij$

This constraint ensures that if the service frequency is set to zero for all lines $l \in L_r$ that cover a given route $r \in C_{ij}$, then this route can not be selected.

If a route is selected, then this requires a non zero service frequency for at least one line out of the set of lines covering said route. Selecting a route also results in a positive passenger loading on lines covering a pair of stops. How these passengers flows are connected to the line selection is explained in subsquent parts of the methodology.

3.1.5 Vehicle Scheduling

To model the fleet size more accurately it is necessary to include vehicle scheduling in the model. Therefore it is now assumed that the demand matrix D_{ij} covers a certain time period, say two morning peak hours. Then this time period can be segmented into several time steps $t \in T$, Where T is the segmentation of the planning horizon H. Taking all possible available lines it is now possible to construct a time space graph of activities (lines) and time steps. Every activity $a \in A_t$ has at least one possible subsequent activity, namely returning to the depot. All other possible activities require that the sum of start time, the time to do the activity and the time required to travel to the next activity is less than the timestep in which the next activity is available. Modelling this flow of vehicles between activities can be done by considering arcs from activity a at time t to activity a' at time t'. To determine the total number of vehicles required, it is necessary to count the number of vehicles going to the depot.

On this graph we now try to schedule subsequent activities. By allowing vehicles to serve multiple lines, the total number of vehicles required reduces significantly. This requires that activities occur at a specific service frequency and that these frequencies correspond to the service frequency of the lines selected as well as what the sequence of activities, that an individual vehicle can perform, is.

In order to select a frequency and a corresponding frequency pattern (the times at which the task performed on the time horizon) a binary parameter b_a^{fp} will be created, that is equal to 1, if activity a is compatible with various frequency patterns. For example if the frequency of a service is once an hour, then $b_a^{fp} = 1$ only for the activities starting at the

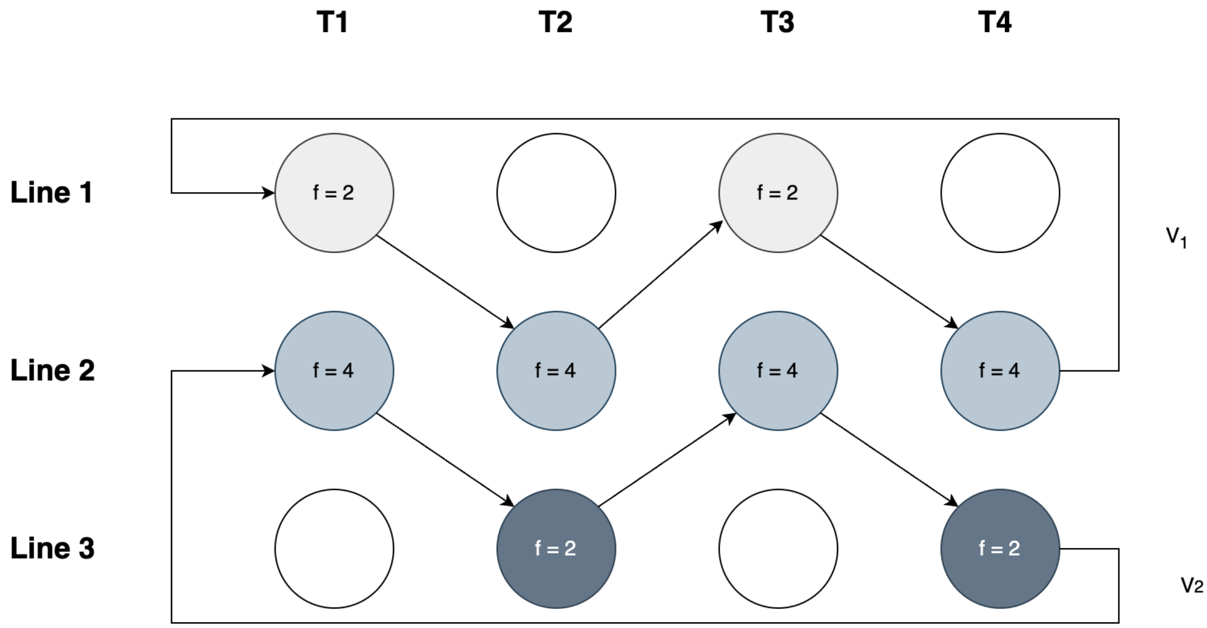


Figure 3.2: Schematic of scheduling vehicles between lines Arcs represent vehicles respectively.

time (step) at which this pattern services a given line.

Correspondingly each line needs to select a frequency and a pattern. For this, the binary variable σ_l^{fp} can be used as well, which is 1 if pattern p and frequency f are selected for line l and 0 otherwise. Lines that are not to be included in the line plan will have their frequency set to 0.

Activity Scheduling

We define an activity as providing a service on one bus line at a given time step. Here our goal is to determine the number of buses. To do that we connect activities that are compatible. Two activities are compatible if the start time of the subsequent activity is after the start time, duration and travel time of the initial activity.

A subsequent activity corresponds to the servicing of the same line or another line at a different time. Taking the intervals $t \in T$ and defining multiple starting times for all lines, a network with activities as nodes, can be generated. Here an arc from a to a' represents whether a' can be done after activity a. This is possible if the start time of activity a' is more than the start time plus the duration of activity a as well as the travelling time required between the last node of the line. This graph will be used to determine a schedule of activities. In order to service an activity a bus is used and this schedule is therefore used to determine the number of vehicles required.

Activity Scheduling Methodology

1. **Parameters and Variables**

   - Parameter p^f distribution pattern for given frequencies.
   - b_a^{fp} Binary parameter whether activity $a \in A_t$
     is compatible with frequency f and corresponding pattern p.
   - $\sigma_l^{fp} \in \{0,1\}$ variable whether line $l \in L$ has frequency f with pattern p.
   - $x_{tarc} \in \{0,1\}$ variable indicating whether an arc from activity $a \in A_t$ at time $t \in T$ to activity $c \in A_a$ at time $r \in T$ is selected

2. **Set of subsequent activities**
   The set of subsequent activities $a_{a'}$, possible after activity a will include activity a' if start time of activity a' is greater than the start time of activity a plus the duration of activity a.
   Related Constraints: If an activity is scheduled, a corresponding frequency distribution, compatible with the start time of the activity, needs to be selected. Activities correspond to lines. Node $x_{ra'ta}$ represents an arc from activity a' at time $r \in T$ to activity a at time $t \in T$. The depot node D is excluded, as it does not correspond to a line.

   - $\sum_{r=0}^{t} \sum_{a' \in A_a} x_{ra'ta} = \sum_{f \in F} \sum_{p \in P_f} b_a^{fp} \sigma_a^{fp} \quad \forall a \in A_t/\{D\}, t \in T$

3. **Scheduling**

   - $\sum_{r \in T} \sum_{a' \in A_r} x_{ra'ta} = \sum_{r \in T} \sum_{a' \in A_r} x_{tara'} \quad \forall a \in A_t, t \in T$
     This constraint ensures flow conservation. The number of ingoing arcs (vehicles) needs to equal the number of outgoing arcs. This ensures vehicle conservation.

   - Ensure vehicle conservation:
     For each time step t and activity a,
     if $a = D$,
     $\sum_{r=0}^{T} x_{r,D,t,a} = \sum_{r=0}^{T} x_{t,a,r,D}$
     else,
     $\sum_{r=0}^{T} x_{r,D,t,a} = \sum_{r=0}^{T} x_{t,a,r,D}$, where $a = A[r]$ or $D = A[r]$.
     The constraint states that the number of outgoing arcs from the depot D needs to equal the ingoing arcs to the depot D. Specifically it ensures that if an arc is generated from the depot node, to the depot node, no additional activities are scheduled subsequently, because only depot activities can be done from then on. It further ensures vehicles conservation between nodes that are not part of

the depot. These "depot-depot" arcs, do not contribute to the total number of vehicles and are therefore artifacts of the model.

- Leave depot at the beginning:
  For each time step t and activity a,
  $\sum_{t=1}^{T} \sum_{a \in A_t} x_{0,D,t,a} =$ Total number of vehicles
  This constraint counts the number vehicles leaving the depot, going to an activity, that does not correspond to the depot.

- Return to depot at the end:
  $\sum_{r=0}^{T-1} \sum_{a \in A_r} x_{r,a,T,D} =$ Total number of vehicles
  This constraint counts the number vehicles returning to the depot, coming from an activity, that does not correspond to the depot. These two constraints work because the previous constraint related to vehicle conservation requires all activities that end at the depot, continue doing the "depot activity" until the final timestep, at which point they are counted. This constraint also ensures that the number of vehicles counted the depot at time zero is equal to the number of vehicles that are at the depot at the final time step.

- Ensure that vehicles do not travel backwards in time:
  For each time step r, activity a, time step t, and activity a',
  if $r \geq t$, then $x_{r,a,t,a'} = 0$
  This constraint should already be covered by the activity scheduling, however allows for faster reduction of variables by the model.

3.1.6 Electrification

The final component of the model is the inclusion of battery sizing and charging station location. Through the introduction of an "energy level" decision variable the state of charge of the battery between successive stops can be tracked. At any given stop this energy level is influenced by, the distance covered and the number of people transported, the amount charged at a stop and the energy level at the previous stop. The nature of this energy balance is affected by the location of charging stations (whether the battery can be charged at a stop, otherwise the amount charged is equal to zero), as well as the battery size (which determines the maximum and minimum values the energy level can attain).It is assumed that vehicles start at the first stop with full charge.

Batteries are selected jointly as part of the vehicle type arrangement, therefore only combinations of passenger seating capacity and battery size can be considered, as it simplifies the analysis and especially the formulation of a corresponding objective function significantly. Charging stations can be located at any stop, but cost a significant amount to install, which can also be reflected in the objective function.

Electrification Methodology

A crucial component of the electrification methodology is the introduction of the w_{ij}^{lf} variable. This variable will link the passenger flow w_{ij}^{l} as determined by the route choice, with the energy expended by transporting passengers. It breaks the flow of passengers over the total time horizon into fractions that correspond to the service frequency selected, as described previously. Thereby it also finalizes the selection of lines and ultimately concludes the linking of route choice and line selection.

It should be noted that this variable is not directly connected with vehicle scheduling. Vehicle scheduling corresponds with the frequency set, which does correspond with w_{ij}^{lf} however. It is therefore assumed that the number of passengers per vehicle is distributed uniformly for each vehicle scheduled on a given line. That means that if a line has a frequency of 4 and a total passenger flow of 100 over the time horizon, 25 passengers are assumed to be on each scheduled vehicle and 25 passengers will be used for the energy balance of a vehicle.

1. **Parameters and Variables** *Assumption: Before the start of each activity the battery is fully charged.*

   - β_{il} Charging level at stop $i \in N$ on line l
   - δ_{il} Charging amount at stop $i \in N$ on line l
   - h_v Binary variable vehicle type selection
   - z_i Whether a charging station is located at stop $i \in N$
   - w_{ij}^{lf} Variable indicating the number of travellers from $i \in N$ to $j \in N$ on line l for frequency setting f.
     Related Constraints: The first constraints ensure that if stops i, j are not part of a line l, w_{ij}^{lf} is set to zero. The second that if i, j are not subsequent stops on a line l, w_{ij}^{lf} is set to zero. Furthermore w_{ii}^{lf} is also set to zero. Finally if the frequency f is zero, w_{ij}^{lf} is also zero.
   - $\sum_{f \in F} w_{ij}^{lf} = 0 \quad \forall l \in L, \forall i, j \in N \not\subseteq l$
   - $\sum_{f \in F} w_{ij}^{lf} = 0 \quad \forall l \in L, \forall i, j \in N \subseteq l | i = l_n \wedge j = l_{n+1}$
   - $\sum_{f \in F} w_{ij}^{lf} = 0 \quad \forall l \in L, \forall i \in N$
   - $w_{ij}^{lf} = 0 \quad \forall f = 0, l \in L, \forall i \in N$

2. **Energy balance**

   - $\beta_{il} = \beta_{jl} - \zeta C_{ij} \sum_{f \in F} w_{ij}^{lf} + \delta_{jl} \quad \forall i, j \in S_l, l \in L$

- $\sum_{v \in V} C^v_{min} h_v \leq \beta_{il} \leq \sum_{v \in V} C^v_{max} h_v \quad \forall i \in S_l, l \in L$

  These constraints ensure an energy balance between stop nodes on a given line. The energy level at a stop needs to equal the energy level at the previous stop, minus the energy deducted based on travel time and passengers transported, plus a possible charged amount. The energy level is bound by maximum and minimum charge depths.

3. **Charging Station Location**

   - $\delta_{il} \leq d_c z_i \quad \forall i \in S_l, l \in L$
     The amount charged at a station can at most be the max charge amount, if a charging station is placed at the stop.

4. **Battery Selection**

   - $\sum_{v \in V} h_v = 1$
     Only one vehicle type and therefore battery type can be selected. The battery type determines battery size.

5. **Model Linking**

   - $w^l_{ij} \leq \sum_{f \in F} f wij^{lf} \quad \forall i, j \in N, l \in L$

     This constraint ensures that the total number of passengers travelling on a line over the time horizon is split equally on each vehicle.

   - $w^{lf}_{ij} \leq \sum_{p \in P} Q \sigma^{fp}_a \quad \forall i, j \in N, l \in L$

     This constraint ensures that the total number of passengers travelling on a vehicle does not exceed the vehicle capacity. Furthermore it links the frequency selected for a line, with the corresponding w^{lf}_{ij} variable.

3.2 Full Model

The following section summarizes the resulting decision making model used to generate a line plan for the electrification of a passenger bus transport system, taking into consideration various battery sizes and charging station location. The model consists of 10 variable types (5 binary, 5 continuous), 20 constraints in the basic version, that can be extended by additional constraints related to minimum demand, maximum fleet size or number of charging stations placed.

3.2.1 Sets, Parameters and Variables

Sets

Sets and Indices	Description
$i, j \in N$	Set of Station Nodes
C_{ij}	Travel cost from node i to node j
D_{ij}	Travel demand from node i to node j
$l \in L$	Set of all lines in the linepool
$r \in R_{ij}$	Set of all routes from node i to node j
S_r	Set of all stops on route r
L_r	Set of all lines covering route r
S_l	Set of all stops on line l
P_{nij}^{r}	Probability of route r from i to j in nest n
$f \in F$	Set of service frequencies considered ([1,2,3,4,6,12,0])
$p \in P_f$	Set of service patterns for frequency f
$t \in T$	Set of time steps covering the time horizon H
$a \in A_t$	Set of activities that can be scheduled at time t
$a' \in A_a$	Set of activities that can be scheduled after activity $a \in A_t$
$v \in V$	Set of vehicle types

Table 3.1: Overview of all sets in the model

Parameters

Parameters	Description
α_i^r	Binary parameter set to one if station i is included in route r
α_i^l	Binary parameter set to one if station i is included in line l
b_a^{fp}	Binary parameter whether activity $a \in A_t$ is compatible with frequency f and corresponding pattern p
p^f	Distribution pattern for frequency f
ζ	Charge used per passenger per travel time
d_c	Maximum charge amount at charging stations
C_{min}^v	Minimum permissible charge level for vehicle type v
C_{max}^v	Maximum permissible charge level for vehicle type v
Q	Passenger capacity

Table 3.2: Overview of all parameters in the model

Variables

Variables	Description
y_r	Binary variable indicating whether route r is active in the model
$\sigma_l^{fp} \in \{0,1\}$	Binary variable indicating whether line l is chosen for service with frequency f and service pattern p
w_{ij}^l	Positive continuous variable taking the value total passengers over the time horizon between stops $i, j \in N$ on line l
w_{ij}^{lf}	Positive continuous variable taking the value of passengers between stops $i, j \in N$ on line l per service, depending on frequency f
x_{tarc}	Binary variable indicating whether an arc from activity $a \in A_t$ at time $t \in H$ to activity $c \in A_r$ at time $r \in H$ is selected
β_{il}	Continuous positive variable for charging level at stop $i \in N$ on line l
δ_{il}	Continuous positive variable for charging amount at stop $i \in N$ on line l
h_v	Binary variable indicating which vehicle type (battery type) v is selected
z_i	Binary variable indicating whether a charging station is located at stop $i \in N$
Λ_{fleet}	Continuous positive variable Total fleet size

Table 3.3: Overview of all variables in the model

3.2.2 Objective

$$max \quad \tau \sum_{ij \in OD} \sum_{r \in R_{ij}} D_{ij} P_{n=1,ij}^r y_r - \upsilon \Lambda_{fleet} - \phi \sum_{i \in N} z_i \tag{3.5}$$

The objective constists of the three objectives: passenger satisfaction, fleetsize and number of charging stations. The can be optimized jointly by defining the objective as a revenue function where the three objectives are optimized jointly in monetary terms. Here the τ, υ and ϕ are weighting coefficients. For satisfaction τ for example could constitute the ticket price, whilst υ and ϕ could represent the respective costs associated with fleet size and infrastructure investment. These coefficients need to be evaluated based on the problem context.

3.2.3 Constraints

$$\sum_{l\in L}\alpha_i^r y_r \geq 1 \quad \forall \quad i\in N \tag{3.6}$$

$$\sum_{l\in L_r} w_{ij}^l = \sum_{r\in R_{ij}} D_{ij}P_{n=1,ij}^r y_r \quad \forall \quad i,j\in N \quad n=1 \tag{3.7}$$

$$\sum_{l\in L_r}\sum_{f\in F}\sum_{p\in P_f} f\sigma_l^{fp} \geq y_r \quad \forall \quad i,j\in N, r\in C_{ij} \tag{3.8}$$

$$\sum_{r=0}^{t}\sum_{a'\in A_a} x_{ra'ta} = \sum_{f\in F}\sum_{p\in P_f} b_a^{fp}\sigma_a^{fp} \quad \forall \quad a\in A_t, t\in T \tag{3.9}$$

$$\sum_{r=0}^{T} x_{r,D,t,a} = \sum_{r=0}^{T} x_{t,a,r,D} \quad \forall \quad t\in T, a\in A_t \tag{3.10}$$

$$\sum_{r=0}^{T} x_{r,D,t,a} = \sum_{r=0}^{T} x_{t,a,r,D} \quad \forall \quad t\in T, a\in A_t = D \tag{3.11}$$

$$\sum_{r=0}^{T} x_{t,a,r,D} = \sum_{r=0}^{T} x_{t,a,r,D} \quad \forall \quad t\in T, a\in A_t, a\in A_r, D\in A_r \tag{3.12}$$

$$\sum_{t=1}^{T}\sum_{a\in A_t} x_{0,D,t,a} = \Lambda_{fleet} \quad \forall \quad t\in T, a\in A_t \tag{3.13}$$

$$\sum_{r=0}^{T/\{T_n\}}\sum_{c\in Ar} x_{r,c,T,D} = \Lambda_{fleet} \tag{3.14}$$

$$x_{r,a,t,c} = 0 \quad \forall \quad r\in T, a\in A_r, t\in T, c\in A_c, r\geq t \tag{3.15}$$

$$\sum_{f\in F} w_{ij}^{lf} = 0 \quad \forall l\in L, \forall i,j\in N \nsubseteq l \tag{3.16}$$

$$\sum_{f\in F} w_{ij}^{lf} = 0 \quad \forall l\in L, \forall i,j\in N \subseteq l | i = l_n \wedge j = l_{n+1} \tag{3.17}$$

$$\sum_{f\in F} w_{ij}^{lf} = 0 \quad \forall l\in L, \forall i\in N \tag{3.18}$$

$$w_{ij}^{lf} = 0 \quad \forall f=0, l\in L, \forall i\in N \tag{3.19}$$

$$w_{ij}^{l} \leq \sum_{f \in F} f w_{ij}^{lf} \quad \forall l \in L, i, j \in N \tag{3.20}$$

$$w_{ij}^{lf} \leq \sum_{p \in P_f} Q \sigma_l^{fp} \quad \forall l \in L, i, j \in N, f \in F \tag{3.21}$$

$$\beta_{il} = \beta_{jl} - \zeta C_{ij} \sum_{f \in F} w_{ij}^{lf} + \delta_{jl} \quad \forall i, j \in S_l, l \in L \tag{3.22}$$

$$\delta_{il} \leq d_c z_i \quad \forall i \in S_l, l \in L \tag{3.23}$$

$$\sum_{v \in V} C_{min}^{v} h_v \leq \beta_{il} \leq \sum_{v \in V} C_{max}^{v} h_v \quad \forall i \in S_l, l \in L \tag{3.24}$$

$$\sum_{v \in V} h_v = 1 \tag{3.25}$$

Chapter 4

Numerical Results

In order to estimate the feasibility of this methodological approach the model was implemented in python and solved using the commercial gurobi solver. The model was run by generating completely random input matrices which only depend on normalization procedures to define the time horizons and time scales.

When running the model a Macbook Air M1 with 8GB of RAM, up to 9 stops can be considered in the model.

The computational performance of this approach strongly depends on the linepool generated for selection. Line length or running time constraints in the lines considered have defining impact on the overall linepool and therefore the lines considered. The larger the line pool the longer the optimization procedure is going to take, and the more data needs to be loaded simultaneously.

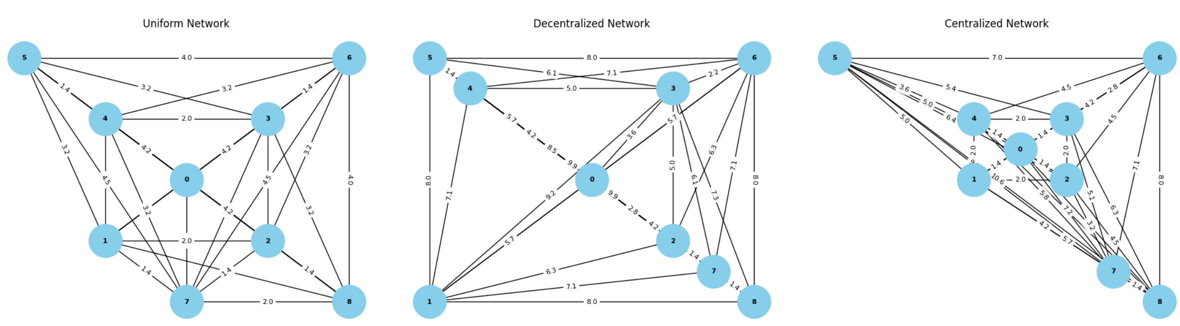
In order to run the discreet choice procedure a couple of parameters need to be assumed or measured prior to the optimization procedure. These include the β sensitivities used in the utility functions, but also the λ fitting parameters. Extension of the utility functions to include other aspects is ofcourse straight forward if the data is available and does not have an impact on the model. Unfortunately the discrete choice model used for joint route and mode choice, specifically regarding the sensitivity parameters, could not be validated, as that would have required a thorough estimation procedure based on real world data. The values chosen; namely -1 for the travel time β parameter and 0.5 for λ_{PT} and 1.5 for λcar where loosely based off Shakeel et. al (2016) [33], as a rough guidance, given that the models are not exactly the same. Their model was estimated for the city of Sydney, as well as taking additional parameters, like access time, travel cost and number of transfer into account.

Section 4.1 will explain what networks have been chosen to investigate aspects of the model and how they impact passenger behaviour, these will be used as a basis for all further subsequent analysis. Section 4.3 will investigate how the λ and β in the utility estimation influence the model, Section 4.4 explores the impact of line-pool size on performance, Section 4.5 will analyse the trade-off between infrastructure investment, fleet size and passenger service satisfaction and finally Section ?? will offer an extension of the model for concrete battery sizing.

4.1 Instances

In order to test the model under varying network responses, 3 similar node networks were generated to which 4 different demand matrices can be applied. Distances are expressed in minutes of travel time. Node proximity is therefore not determined by physical characteristics but on a network basis, where neighbouring routes are faster to reach that nodes further away. This can be justified through the assumption that the underlying road network allows for a connection between all origins and destinations and that therefore only the total duration of travel time, from one node to the next, is relevant for the network representation. Accordingly the network is fully connected. Travel demand, represented by an OD-matrix is considered on a per person basis.

The first node graph consists of roughly equally spread out nodes, organized around a 5 node cluster in the center. This graph is going to referred to as the base network. The remaining two networks are roughly derived from this graph. The second network considers an exploded case, in which nodes around the center are moved further out. This results in generally longer travel times, especially via routes that travel through multiple nodes. By keeping the distance between distant nodes small, this also represents a more "decentralized" network as nodes are clustered around the periphery. The final network considers the case where only the most outward nodes are moved further out, resulting in a close center proximity with larger distances to the outter edges. The result of which are long travel times to and away from the center, as well as between nodes not in the center. The three graphs can therefore be summarized as uniform network, decentralized network and centralized network. The three networks shown in figure 4.1.



(a) (b) (c)

Figure 4.1: (a) Uniform Network (b) Decentralized Network (c) Centralized Network

The demand matrices were constructed similarily. The first considered uniform random demand, between any origin and destination pair. Demand was distributed uniformly between 0 and 110. This was achieved simply by generating a random matrix. When looking at the uniform network, the inner nodes here are the 5 central nodes, labled 0-4. The remaining nodes are termed the outter nodes. The second demand matrix consisted of high demand to and from inner nodes and lower demand to and from outter nodes. This was done by splitting the demand matrix into corresponding segments that each recieved differently generated random numbers, distributed uniformly across from 40 to 85, for high demand and from 0 to 40 for low demand. The third demand matrix considered the opposite case with high outter demand across nodes 5-8, also distributed uniformly from 40 to 85, for high demand and from 0 to 40 for low demand. Finally the last demand matrix was constructed as a sparse matrix, with random numbers generated from 0 to 180. The resuling demand characterisitcs are summarized in Table 4.1 below:

Matrix	**Size**	**High Demand**	**Low Demand**	**Density**	**Total Demand**
Uniform	9x9	-	-	0.88	3939
High Inner Demand	9x9	{0,1,2,3,4}	{5,6,7,8}	0.88	3917
High Outer Demand	9x9	{5,6,7,8}	{0,1,2,3,4}	0.88	3929
Sparse	9x9	-	-	0.42	3919

Table 4.1: Demand Matrices

From this 12 instances were created to test the model. One for each network structure, paired with demand. They are listed in Table 4.2 below:

Instance	**Network**	**Demand**
1	Uniform	Uniform
2	Uniform	Inner
3	Uniform	Outer
4	Uniform	Sparse
5	Imploded	Uniform
6	Imploded	Inner
7	Imploded	Outer
8	Imploded	Sparse
9	Exploded	Uniform
10	Exploded	Inner
11	Exploded	Outer
12	Exploded	Sparse

Table 4.2: Instances

These instances will be used to investigate the trade-off between the varying objectives, as described in Section 4.5, a subset of these may be used to investigate other aspects of the model. The following subsection will offer insight into the way the performance of the scenarios differ from each other.

4.2 Instance Performance

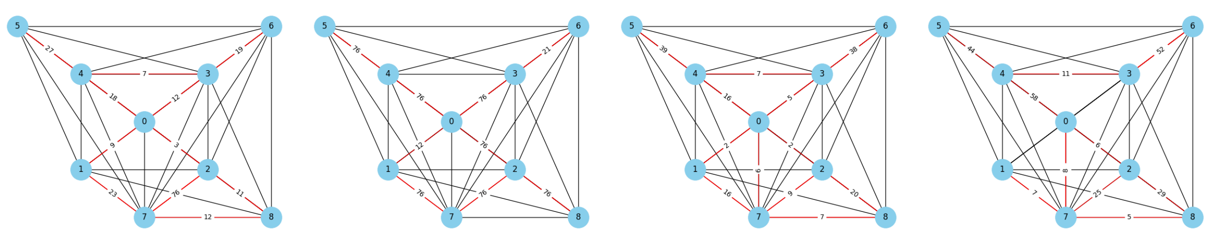
To compare the general performance of the scenarios fairly across the objectives (line-plan satisfaction, fleet size and infrastructure required), the values in the objective function need to be normalized. Otherwise it is possible that the optimization prefers the objective with the largest magnitude, as the sum total of across all objectives can be influenced most in this way. To ensure a fair comparison the objectives will be normalized in the following way:

$$\theta_1 = \frac{X_1 - X_1^{min}}{X_1^{max} - X_1^{min}} \tag{4.1}$$

This way objective X is normalized such that $0 \leq \theta_1 \leq 1$. Since all objectives have the same magnitude range, the algorithm will seek to treat them equally when finding the maximum. The objective used to compare the scenarios is the following:

$$max \quad \theta_{sat} - \theta_{fleet} - \theta_{infra} \tag{4.2}$$

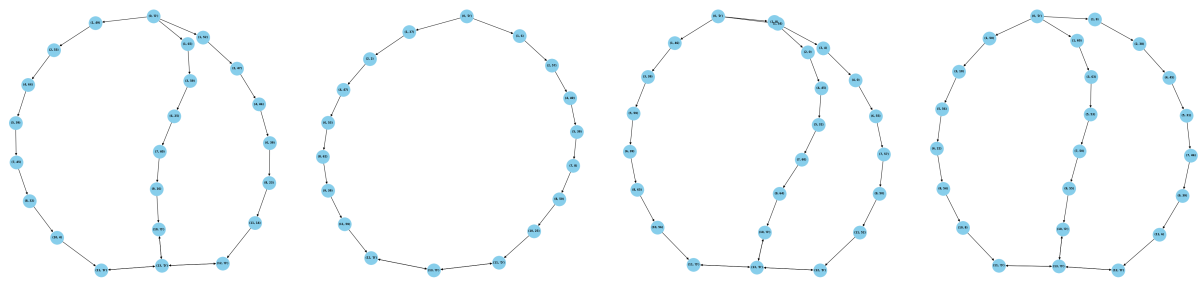
The following Figure 4.2 shows the network structure and passenger flows for instances 1-4, across the normal network for the 4 types of demand. The red lines represent active lines across the network, the numbers indicating the total number of passengers travelling along that arc over the total time horizon.



(a) (b) (c) (d)

Figure 4.2: Passenger Flow: (a) Normal Demand (b) Inner Demand (c) Outer Demand (d) Sparse Demand

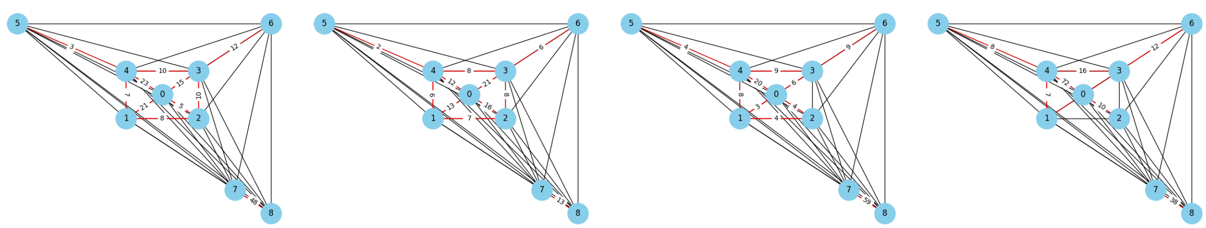
Figure 4.3 shows the vehicle flows for instances 1-4. The diagramme shows the depot as both source and sink, the number of out and ingoing arcs to the depot represents the number of vehicles. The depot is marked as activity "D". Once returned to the depot, the only subsequent valid activity is remaining at the depot. Subfigure (a), (c) and (d) require three vehicles, whilst for the inner demand case (b), 2 vehicles are sufficient. It should be noted that 76, is the vehicle capacity.



(a) (b) (c) (d)

Figure 4.3: Vehicle Flow: (a) Normal Demand (b) Inner Demand (c) Outer Demand (d) Sparse Demand

The following was repeated for 2 other node networks, with Figures 4.4, 4.5 and Figures 4.6, 4.7 representing the imploded and exploded case respectively.

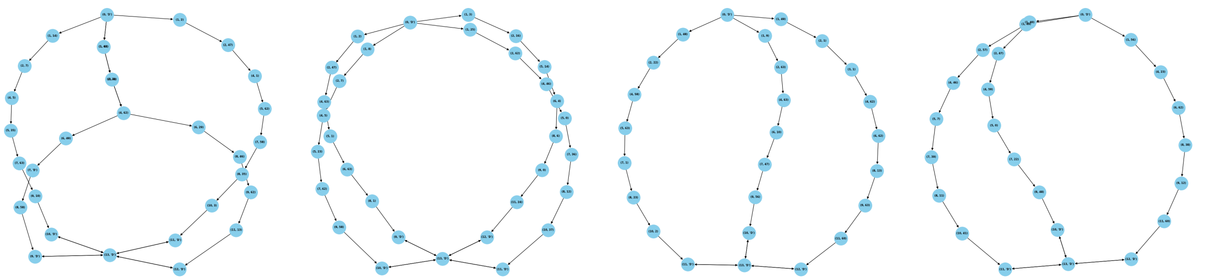


(a) (b) (c) (d)

Figure 4.4: Passenger Flow: (a) Normal Demand (b) Inner Demand (c) Outer Demand (d) Sparse Demand

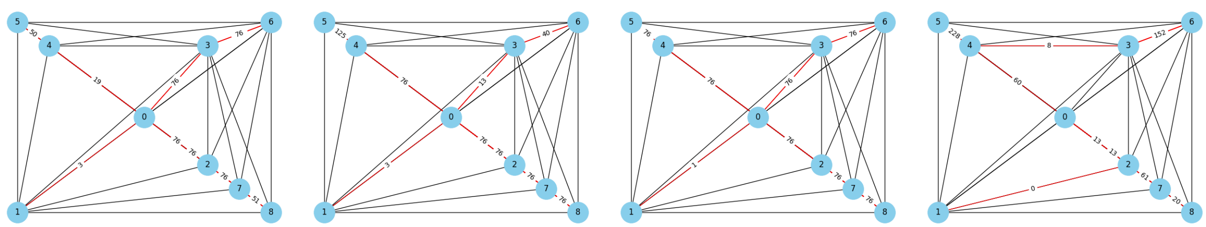
4.3 Sensitivity Analysis of Choice Parameters

The nested logit model used for mode and route choice is strongly dependent on the estimated parameters. To implement and test the model, that this research proposes, parameters were chosen based on Waller (2016) [33], who estimated them for the city of Sydney. These parameters have a strong influence on the number of people travelling, their sensitivity for travel time as well as the distinction between the car and public transit nest. As the estimation of the nested logit is part of the preprocessing, this is technically not part of the model proposed, however the following section will offer some insight into the sensitivity of the model to changes in choice model parameters. The following table 4.3



(a) (b) (c) (d)

Figure 4.5: Vehicle Flow: (a) Normal Demand (b) Inner Demand (c) Outer Demand (d) Sparse Demand



(a) (b) (c) (d)

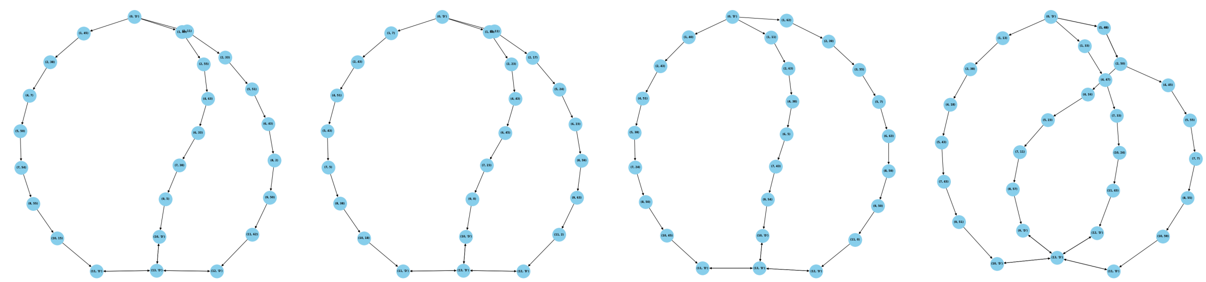
Figure 4.6: Passenger Flow: (a) Normal Demand (b) Inner Demand (c) Outer Demand (d) Sparse Demand

shows how changing various model parameters impact the total number of routes selected by the model.

ASC (PT)	β_{Time}	λ_{PT}	λ_{Car}	**Service Satisfaction**	**Service Disatisfaction**
2.5	-0.025	2	6	11.10%	0.23%
5	-0.025	2	6	11.19%	0.01%
2.5	-0.05	2	6	11.36%	0.24%
2.5	-0.025	4	6	11.05%	0.04%
2.5	-0.025	2	12	11.10%	0.23%

Table 4.3: Effect of discrete choice parameters on Service Satisfaction

The Service Satisfaction expressed here as a percentage is the number of passengers willing to choose to travel by bus on this network. The service dissatisfaction on the other hand is the number of people that choose to travel, but prefer to take the car. In other words; those who prefer direct travel via the fastest possible route or shortest path. The remaining demand could not be satisfied by the network and would have preferred alternative routes through lines, not offered by the network.



(a) (b) (c) (d)

Figure 4.7: Vehicle Flow: (a) Normal Demand (b) Inner Demand (c) Outer Demand (d) Sparse Demand

An effect that is clearly visible is that the discrete choice parameters have a profound impact on the service disatisfaction. Larger values of λ_{car} relative to λ_{PT}, as well as more negative values and larger absolute values for β_{Time} relative to the ASC, especially impact service disatisfaction. On the other hand service satisfaction is also impacted by the size of the linepool. More lines to choose from can enable more satisfying routes and therefore impact service satisfaction. This can be seen in table 4.6.

4.4 Computational Performance and Limitations

The model performance is strongly dependent on the size of the linepool. As this model framework is highly combinatorial, the time required to find an optimal solution increases significantly with the number of variables. Multiple decision variables depend on the total number of lines, since many variables are instantiated and indexed by a single line. Furthermore activities, are also definied as lines that can be started at a certain time, so the vehicle scheduling is also strongly impacted by the number of lines. Therefore the problem becomes increasingly more complex the more lines are considered in the model.

In order to influence performance and total number of variables, linepool constraints have been introduced. This is done by adding only those simple paths between any two nodes to the linepool, if the sum of the arcs (total trave time) is between a minimum and maximum value.

To ensure that the scenarios remain comparible, different linepool constraints were required.

The following table 4.4 exemplifies how linepool constraints impact the computational performance of the model. It shows the effect of changing the linepool constraints on both the total linepool size, as well as the number of variables in the model (columns) and its impact on the resulting run time. Increasing the span between linepool constraints

Network	Min Line Length (min)	Max Line Length (min)	Total Linepool Size
Uniform	10	30	66
Imploded	10	65	64
Exploded	10	40	64

Table 4.4: Linepool constraints for the Networks

greately increases the linepool, resulting in a significant increase in lines considered by the model. This greately impacts the number of variables and therefore computational runtime.

Uniform Network (9 Nodes)							
Min Length	Max Length	Linepool Size	Total Number of columns	Total Runtime	MIP Gap	Service Satisfaction	Service Disatisfaction
20	30	48	402,171	3.97	5%	10.81%	0.23%
10	30	66	715,983	6.82 s	5%	11.10%	0.23%
10	40	150	3,414,359	41.92 s	5%	17.75%	0.23%

Table 4.5: Effect of Linepool constraints on the uniform network with 9 nodes

The size of the linepool has large influence on the computational performance of the model. In order to control runtimes and match the size the model to the computational resources available, controlling the size of the linepool becomes a critical task in preprocessing for real world application. An aspect that should be considered is that the modeller requires prior insight into design characteristics of the resulting lines. What are the shortest possible routes that should be considered? What are the longest lines that should be scheduled? Considering that lines shouldnt exceed a unit of the time horizon for the vehicle scheduling, it is therefore necessary to make these two decisions in tandem. Furthermore, depending the on the distances of nodes and the number of nodes themselves, the longest line length is constrained naturally by the longest paths taken through the network anyway, which should be taken into consideration when constraining the linepool.

The following table 4.6 shows how varying the number of nodes impacts computational performance. Node distance and demand matrices were generated using entirely random matrices for the sake of this comparison, a caveat being that this impacts linepool size as distances are not effected by symmetries or triangle inequalities.

Random Network/Random Demand				
Number of Nodes	Linepool Size	Total Number of columns	Total Runtime	MIP Gap
5	222	7,224,913	265.93 s	5%
6	240	8,454,941	322.43 s	5%
7	214	6,761,453	187.75 s	5%
8	247	6,418,286	268.17 s	5%

Table 4.6: Effect of network size at (roughly constant linepool size)

As can be seen the number of nodes has less impact on the computational performance than the total linepool size. In theory this means that networks of varying sizes can be simulated by the model, in practice however, the linepool size changes based on the number of nodes, specifically by increasing the total amount of simple paths that can be taken from any given origin and destination. The linepool size thus increases naturally simply by virtue of having more options of getting from any origin to any destination. This means that constraining the linepool size becomes more challenging with increasing network size.

4.5 Trade-off Behaviour

This section aims to investigate the trade-off between service satisfaction and costs. The objectives of maximizing service satisfaction, minimizing fleet size and minimizing charging station placement and their sensitivities are all interesting to decision makers.

Line selection produces passenger flows over lines, the number of passengers is constrained by vehicle size, whilst a single vehicle has a fixed battery size and therefore a fixed amount of energy available for transporting passengers, there is a relationship between fleet size and charging station placement. The model can compensate placing less charging stations by scheduling more vehicles and vice versa. To investigate this relationship the epsilon-constraint method was applied to the model.

In order to facilitate this trade-off the various parameters have been set to fairly extreme values. This is necessary as the networks condsidered in the study are significantly larger, as is the running time of the buses. This means that the values need to be adjusted to the scale of the problem. The following table 4.7 shows the values used, as well as reference values, more inline with real world application.

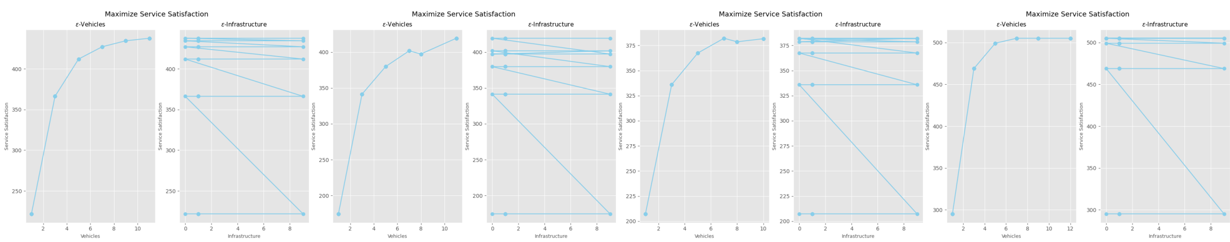
	Model Values	**Values from Literature**
Vehicle Passenger Capacity	76	76[19]
Vehicle Battery Capacity [kJ]	11250	720000
Vehicle discharge rate [kJ per passenger per minute]	11.66	11.66
Station Charging amount [kJ]	9000	600kW over 20s

Table 4.7: Values used in the model, as based on the literature. See footnote.[1]

For the epsilon-constraint method, service satisfaction was discretized into values from 900-50, fleet-size from 1-12 and charging stations from 1-9. The values were chosen based on

[1] In order to facilitate a stronger drive of the model to manage "energy consumption" the vehicle battery capcity was reduced by a factor 2^6. Bus route 106 was taken as reference for passenger capacity, battery capacity and vehicle discharge rate per passenger per minute (trip energy consumption) Jinhua Ji et. al. (2022). The station charging value was retrieved from the ABB flash charging station project description, retrieved on May 2024. `https://library.e.abb.com/public/4e4c1be356f24e8aa3a74e2042d73860/ABB_factsheetA4_TOSA_reference_EN.pdf?x-\sign=aEnksAKEXUVQeVE/VYHh+M8joow/SpO4nFEu0HJjkvZjkZ/SmCRfiKsCXi8KZ1T`

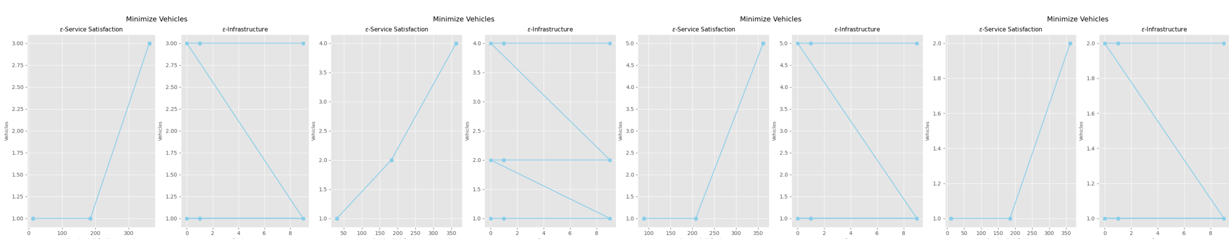
the following rationale; the maximum values of 900, 12 and 9 for for service satisfaction, fleetsize and charging stations were taken from the maxizing service satisfaction only, the minimum value charging stations was set to 0 as vehicles start fully charged when leaving the depot, in theory making it permissible not to require any additional electrification infrastructure, a minimum of 1 vehicle is necessary to establish a service and 50 for service satisfaction, which was discretized in units of 50, leading to minimum value. For each scenario the optimization was run, progressively tightening a constraint for two objectives, whilst optimizing the other. This resulted in the following plots 4.8, 4.9 and **??**. Since the optimization procedure would not always converge, only values were plotted that resulted in an optimial solution.



(a) (b) (c) (d)

Figure 4.8: Passenger Flow: (a) Normal Demand (b) Inner Demand (c) Outer Demand (d) Sparse Demand

As can be seen in 4.8 the methodology will alternate between placing charging stations and vehicles in order to increase service satisfaction. More vehicles means more passengers can be transported, stations allow for reusing vehicles, whilst few charing stations require more vehicles to compensate for the charge consumed by passengers.



(a) (b) (c) (d)

Figure 4.9: Passenger Flow: (a) Normal Demand (b) Inner Demand (c) Outer Demand (d) Sparse Demand

When minimizing charging stations the model will resort to configurations in which no stations need to be placed. It will compensate this by using as many vehicles as allowed to facilitate as much service satisfaction as possible. If these service satisfaction constraint is at odds with the fleet size constraint the opimization procedure will not be able to produce a result. All of this is strongly indicative that fleet size is a function of charging infrastructure. The more charging infrastructure is placed, the lower the fleet size can be. More vehicles means, less charging infrastructure is required.

A take away of this is that determining accurate weighting for the objectives to bring them into proper relation with each other is definitely necessary to use this methodolgy as effectively as possible.

Chapter 5

Conclusion & Recommendations

5.1 Conclusion

In conclusion, this study has provided a comprehensive examination of line planning optimization for public transit systems, culminating in the development of a sophisticated model designed to address critical aspects of route selection, service frequency determination, vehicle choice, and scheduling simultaneously. Drawing upon insights gleaned from an exhaustive literature review, the model integrates elements of discrete choice modeling and mode choice considerations to optimize passenger distribution and journey times along transit lines.

This novel approach showed that lineplanning, frequency setting and vehicle scheduling can all be assessed jointly via an optimization approach. It also showed that passenger flows depend on the node locations of the transport networks and that despite this similar fleet sizes can be used to service a network, whilst still serving all nodes. Another conclusion that can be drawn is that battery and fleet sizing, as well as charging infrastructure are linked. A larger fleet can compensate for smaller battery sizes, as can more infrastructure and vice versa.

While the model represents a significant extension to the reportoir of line planning methodologies, it is crucial to acknowledge its computational constraints, especially when confronted with large-scale networks featuring numerous nodes. As can be seen, due to the combinatorial nature of this problem, the number of lines in the linepool significantly impacts run times. The size of the linepool depends on the number of lines, can be controlled however in the preprocessing step. Despite its innovative approach, the model's ability to handle extensive datasets is restricted, as evidenced by its limitation to consider only up to 9 stops during testing on the MacBook Air M1 with 8GB of RAM. This discrepancy underscores the challenge of managing vast amounts of data simultaneously within system memory, highlighting the need for more efficient data processing and loading mechanisms to enhance the model's usability and scalability. Another aspect is the number of assumptions required to justify the simplifications this approach makes, these need to be carefully evaluated to ensure that realworld implementation remains feasible.

However, despite these limitations, the model serves as a valuable framework for exploring the intricate interplay between strategic and operational considerations in public transit planning. By concurrently evaluating multiple factors, it offers valuable insights into optimizing transit networks for efficiency, sustainability, and passenger satisfaction.

Moving forward, addressing the computational constraints while retaining the model's analytical depth will be paramount. Additionally, continuous validation and refinement through empirical testing and real-world applications will be essential to ensure the model's practical utility and effectiveness in addressing the complex challenges inherent in urban transportation planning and management. Despite the hurdles, the model stands as a promising tool for guiding decision-making processes and advancing the discourse on transit network optimization in an ever-evolving urban landscape.

5.2 Recommendations

The following chapter will offer a reflection on opportunities this methodology provides for future planners and model developers as well as offer perspective on further research that can be undertaken to improve this methodology.

This research has shown methodology is feasible for evaluating effective use of resources for small transport networks. Instead of going through several stages in the transport system design process, this methodology can be used instead. This methodology can also be used to evaluate sub-components of transport network especially if multiple vehicle depots are used. If it is possible and feasible to divide a given study area into several subareas with individual depots, its possible to apply this approach to find a lineplan and fleetsize for these subregions. This also means that this approach is suitable in updating a transport system parts at a time, rather than all at once. The routes already present in the current transport system can be transfered into the linepool together with some alternatives to see if improvements can be made to the line plan. At the very least this model offers a new way of linking frequency to lines and vehicles, which may prove useful in other model methodologies. There is a clear trade-off between computational complexity and higher realism. A methodology that takes passengers flows, frequency and vehicle scheduling into account when selecting lines can produce better solutions regarding all of these aspects than the previous sequential approach, however there is a trade-off between network size that can be evaluated as well as the practicality of evaluating all of these aspects at the same time. It may be useful to fix intermediate aspects in the design process, for the sake of planning and communication, or to stick to the current precedent of clearly dividing strategic from tactical planning aspects. Additionally this approach does generate entirly new lines as some other methodologies do, but instead relies on picking preselected lines from the linepool. Should there be a greater emphasis on finding and developing new lines entirely, more focus will be on the linepool generation and techniques for that, or other methodologies may be more suitable.

Opportunities for further research can be grouped in four categories mainly centered around the methodological approach: the formulation of the model and extension and inclusion of more detail, computational aspects to improve performance, consideration of geographic aspects and application to real-world study areas and practicality assessments.

Further research with use of this model into the characteristics of energy use per passenger per minute, battery size and charging amount as well as there relationship to line length, number of vehicles and charging station placement may be of further interest. This also depends on the relationship to node distances, travel times, line pool constraints employed. As these determine the time spent travelling between stops and there for the energy amound discharged per minute. Further investigation into these aspects can surely reveal interesting relationships between these variables that could prove valuable in the future.

There certainly is opportunity to further improvement of this model in a multitude of ways. The general framework for approaching electrification and vehicle scheduling is certainly promising and more nuance, context, specific aspects and additional realism can certainly be included to improve the methodological approach. Examples of this include but are not limited to: the extension of the utility function for higher realism, for example by including stops; considering battery characteristics, like degredation as seen in Azadeh, Vester, Maknoon (2022) [34]; bidirectional line servies or dead-heading of vehicles, a more realistic modelling of passenger assignment and vehicle energy consumption. Furthermore there is, without doubt, also an opportunity to improve the model in terms of formulation, this is evident alone by the depot-depot artifact nodes that are created, indicating that the current formulation could still be improved. Most importantly however additional research has the chance to focus on the utility function exclusively, to turn it into a fully functional revenue function that is fully capable of making trade-offs between vehicle types, income generated by the passenger flows and charging station investment. Since the structure naturally lends itself to revenue maximization, and since the investigation of the pareto frontier has shown that optimization does produce points in the dominant set, only an investigation into appropriate coefficients may be required to investigate realistic cost trade-offs.

During testing on the MacBook Air M1 with 8GB of RAM, it was observed that the model could effectively handle up to 9 stops. While this limitation falls short of representing a realistic network size, the primary issue lies in the model's inability to accommodate large volumes of data simultaneously due to memory constraints. Enhancing data processing and loading mechanisms could significantly enhance the model's usability, making it more adaptable to larger and more complex transit networks. Exploring alternative coding paradigms or procedures may also offer avenues for improving computational efficiency and scalability.

An intriguing avenue for further exploration involves incorporating geographic features into the model's analysis. Presently, the model primarily relies on travel times measured in minutes and utility functions to represent demographic characteristics. However, integrating additional geographic elements unique to specific cities or regions could offer valuable insights into line planning considerations beyond travel times. Factors such as terrain characteristics, population density patterns, and landmark locations could influence route selection and service optimization, thereby enriching the model's predictive capabilities.

Furthermore, ongoing efforts in validating and refining the model through empirical testing and real-world applications are crucial. By subjecting the model to diverse transit planning scenarios and evaluating its performance against real-world data, researchers can ensure its practical utility and effectiveness in addressing the multifaceted challenges of urban transportation planning and management. Continuous refinement and validation will be essential to ensure that the model remains relevant and robust in the dynamic landscape of urban mobility.

Bibliography

[1] Kun An. Battery electric bus infrastructure planning under demand uncertainty. *Transportation Research Part C: Emerging Technologies*, 111:572–587, 2020.

[2] Moshe E Ben-Akiva and Steven R Lerman. *Discrete choice analysis: theory and application to travel demand*, volume 9. MIT press, 1985.

[3] D. Bertsimas, Y. S. Ng, and J. Yan. Data-driven transit network design at scale. *Operations Research*, 69(4):1118–1133, 2021.

[4] R. Borndörfer, M. Grötschel, and M.E. Pfetsch. A column-generation approach to line planning in public transport. *Transportation Science*, 41(1):123–132, 2007.

[5] Michael R Bussieck, Thomas Winter, and Uwe T Zimmermann. Discrete optimization in public rail transport. *Mathematical programming*, 79:415–444, 1997.

[6] A. Ceder and N.H.M. Wilson. Bus network design. *Transportation Research Part B: Methodological*, 20(4):331–344, 1986.

[7] Avishai Ceder. Bus frequency determination using passenger count data. *Transportation Research Part A: General*, 18(5-6):439–453, 1984.

[8] Avishai Ceder, Boaz Golany, and Ofer Tal. Creating bus timetables with maximal synchronization. *Transportation Research Part A: Policy and Practice*, 35(10):913–928, 2001.

[9] Desaulniers and Hickman. Chapter 2: Public Transit. In C. Barnhart and G. Laporte, editors, *"Handbook in Operations Research and Managment Science"*, volume 14th. North Holland, 2007.

[10] Jacques Desrosiers, Yvan Dumas, Marius M Solomon, and François Soumis. Time constrained routing and scheduling. *Handbooks in operations research and management science*, 8:35–139, 1995.

[11] Richard Freling and José M Pinto Paixao. Vehicle scheduling with time constraint. In *Computer-Aided Transit Scheduling: Proceedings of the Sixth International Workshop on Computer-Aided Scheduling of Public Transport*, pages 130–144. Springer, 1995.

[12] P. Gattermann, J. Harbering, and A. Schöbel. Line pool generation. *Public Transport*, 9:7–32, 2017.

[13] Anthony F Han and Nigel HM Wilson. The allocation of buses in heavily utilized networks with overlapping routes. *Transportation Research Part B: Methodological*, 16(3):221–232, 1982.

[14] J. Hartleb, M. Schmidt, D. Huisman, and M. Friedrich. Modeling and solving line planning with mode choice. *Transportation Science*, 57(2):336–350, 2023.

[15] Yi He, Zhaocai Liu, and Ziqi Song. Optimal charging scheduling and management for a fast-charging battery electric bus system. *Transportation Research Part E: Logistics and Transportation Review*, 142:102056, 2020.

[16] IEA. Sustainable recovery, transport, 2020.

[17] IEA. Electric vehicles, 2022.

[18] Yechezkel Israeli and Avishai Ceder. Transit route design using scheduling and multi-objective programming techniques. In *Computer-Aided Transit Scheduling: Proceedings of the Sixth International Workshop on Computer-Aided Scheduling of Public Transport*, pages 56–75. Springer, 1995.

[19] Jinhua Ji, Yiming Bie, Ziling Zeng, and Linhong Wang. Trip energy consumption estimation for electric buses. *Communications in Transportation Research*, 2:100069, 2022.

[20] M.J. Klier and K. Haase. Urban public transit network optimization with flexible demand. *Or Spectrum*, 37:195–215, 2015.

[21] Peter Knoppers and Theo Muller. Optimized transfer opportunities in public transport. *Transportation Science*, 29(1):101–105, 1995.

[22] Alexander Kunith, Roman Mendelevitch, and Dietmar Goehlich. Electrification of a city bus network—an optimization model for cost-effective placing of charging infrastructure and battery sizing of fast-charging electric bus systems. *International Journal of Sustainable Transportation*, 11(10):707–720, 2017.

[23] Alexander Kunith, Roman Mendelevitch, Anne Kuschmierz, and Dietmar Goehlich. Optimization of fast charging infrastructure for electric bus transportation–electrification of a city bus network. In *Proceedings of the IEEE 16th International Conference on Intelligent System Applications to Power Systems, Montréal, QC, Canada*, pages 19–22, 2016.

[24] Douglass B Lee Jr, Lisa A Klein, and Gregorio Camus. Induced traffic and induced demand. *Transportation Research Record*, 1659(1):68–75, 1999.

[25] Andreas Löbel. Vehicle scheduling in public transit and lagrangean pricing. *Management Science*, 44(12-part-1):1637–1649, 1998.

[26] T.L. Magnanti and R.T. Wong. Network design and transportation planning: Models and algorithms. *Transportation science*, 18(1):1–55, 1984.

[27] K. Nachtigall and K. Jerosch. Simultaneous network line planning and traffic assignment. In *8th Workshop on Algorithmic Approaches for Transportation Modeling, Optimization, and Systems (ATMOS'08)*. Schloss Dagstuhl-Leibniz-Zentrum für Informatik, 2008.

[28] Matthias Rogge, Sebastian Wollny, and Dirk Uwe Sauer. Fast charging battery buses for the electrification of urban public transport—a feasibility study focusing on charging infrastructure and energy storage requirements. *Energies*, 8(5):4587–4606, 2015.

[29] FJM Salzborn. Scheduling bus systems with interchanges. *Transportation Science*, 14(3):211–231, 1980.

[30] A. Schöbel. Line planning in public transportation: models and methods. *OR spectrum*, 34(3):491–510, 2012.

[31] A. Schöbel and S. Scholl. Line Planning with Minimal Traveling Time. In L.G. Kroon and R.H. Möhring, editors, *5th Workshop on Algorithmic Methods and Models for Optimization of Railways (ATMOS'05)*, volume 2 of *OpenAccess Series in Informatics (OASIcs)*, Dagstuhl, Germany, 2006. Schloss Dagstuhl–Leibniz-Zentrum für Informatik.

[32] Mariana Teixeira Sebastiani, Ricardo Lüders, and Keiko Veronica Ono Fonseca. Evaluating electric bus operation for a real-world brt public transportation using simulation optimization. *IEEE Transactions on Intelligent Transportation Systems*, 17(10):2777–2786, 2016.

[33] Kiran Shakeel, Taha Rashidi, and Steven Waller. Choice set formation behavior: Joint mode and route choice selection model. *Transportation Research Record: Journal of the Transportation Research Board*, 2563:96–104, 01 2016.

[34] Sh. Sharif Azadeh, J. Vester, and M.Y. Maknoon. Electrification of a bus system with fast charging stations: Impact of battery degradation on design decisions. *Transportation Research Part C: Emerging Technologies*, 142:103807, 2022.

[35] YOSEF Sheffi and MORIHISA Sugiyama. Optimal bus scheduling on a single route. *Transport*, 60:68, 1982.

[36] L.A. Silman, Z. Barzily, and U. Passy. Planning the route system for urban buses. *Computers & operations research*, 1(2):201–211, 1974.

[37] Philipp Sinhuber, Werner Rohlfs, and Dirk Uwe Sauer. Study on power and energy demand for sizing the energy storage systems for electrified local public transport buses. In *2012 IEEE vehicle power and propulsion conference*, pages 315–320. IEEE, 2012.

[38] Kenneth Train. *Discrete Choice Methods With Simulation*, volume 2009. 01 2009.

[39] R. van Nes, R. Hamerslag, and L.H. Immers. *The design of public transport networks*, volume 1202. National Research Council, Transportation Research Board, 1988.

[40] Yusheng Wang, Yongxi Huang, Jiuping Xu, and Nicole Barclay. Optimal recharging scheduling for urban electric buses: A case study in davis. *Transportation Research Part E: Logistics and Transportation Review*, 100:115–132, 2017.

[41] Maria Xylia, Sylvain Leduc, Piera Patrizio, Florian Kraxner, and Semida Silveira. Locating charging infrastructure for electric buses in stockholm. *Transportation Research Part C: Emerging Technologies*, 78:183–200, 2017.

[42] Şule Yıldırım and Barış Yıldız. Electric bus fleet composition and scheduling. *Transportation Research Part C: Emerging Technologies*, 129:103197, 2021.