FLEXIBLE TRANSIT NETWORK DESIGN WITH AND WITHOUT BRANCHING UNDER SPATIALLY HETEROGENEOUS DEMAND

BY

ANTOINE PETIT

THESIS

Submitted in partial fulfillment of the requirements for the degree of Master of Science in Civil Engineering in the Graduate College of the University of Illinois at Urbana-Champaign, 2014

Urbana, Illinois

Adviser:

Associate Professor Yanfeng Ouyang
ABSTRACT

While public transportation systems are usually designed with fixed routes, this work presents an alternative flexible-route transit system. Flexible transit vehicles do not operate on fixed paths but travel within predetermined areas in response to trip demand in order to provide door-to-door service. The main advantage of this system is that passenger access time to and from transit stops is removed. To design the optimal route layout and service operation, continuum approximation is used to reduce the computation burden and formulate the problem in terms of a few decision variables. Unlike many continuous models, passenger distribution is not assumed to be homogeneous over space. Since travel patterns are typically not uniform in urban and suburban areas, this thesis will consider a heterogeneous passenger distribution. In order to adapt to both global and local demand variations, several transit system designs will be investigated. Thus, the purpose of this thesis is (i) to investigate the benefits of Daganzo (2010a) hybrid structure over grid structure under heterogeneous demand with flexible routes; and (ii) to investigate the benefits of allowing branching local tubes within the transit system. For (i), we derive the agency and user cost metrics of the proposed models and seek optimal network layout, service area of each bus and bus headway, to minimize the total generalized cost. For (ii), we use the framework provided by Ouyang et al. (2014) and the power-of-two concept from Roundy (1985) to design a grid flexible transit network with local tubes. The same cost metrics as in (i) are derived on a local scale. Considering a low-to-moderate demand level and several spatially heterogeneous demand distributions, it is found out that hybrid structure is beneficial over grid structure, and that transit network with local tubes allows a reduction of the system cost, with respect to a homogeneous transit network. A sensitivity analysis is performed on the branching structure design. It is found that branching does not depend on the total number of passengers. Finally, several future research leads are presented to enhance the transit network design.
I would like to give special thanks to my adviser, Professor Yanfeng Ouyang, for his support and for giving me the opportunity to continue our beneficial collaboration. His guidance was extremely valuable in the development of this thesis.
# TABLE OF CONTENTS

CHAPTER 1 INTRODUCTION ........................................................................................................... 1

CHAPTER 2 LITERATURE REVIEW ............................................................................................. 4
  2.1 Flexible transit description and case experiences ............................................................. 4
  2.2 Transit network design ...................................................................................................... 8
  2.3 Continuum approximation ............................................................................................... 11
  2.4 Heterogeneous demand literature ..................................................................................... 12
  2.5 Solution methods ............................................................................................................ 13

CHAPTER 3 FLEXIBLE TRANSIT SYSTEM DESIGN COMPARISON UNDER SPATIALLY HETEROGENEOUS DEMAND: GRID NETWORK VERSUS HYBRID NETWORK ......................................................................................................................... 15
  3.1 Grid network ..................................................................................................................... 16
  3.2 Hybrid network .................................................................................................................. 24
  3.3 Conclusion ......................................................................................................................... 32

CHAPTER 4 FLEXIBLE TRANSIT SYSTEM DESIGN UNDER SPATIALLY HETEROGENEOUS DEMAND WITH LOCAL TUBES ................................................................................................................................................................................. 33
  4.1 Methodology ...................................................................................................................... 34
  4.2 Continuum approximation ............................................................................................... 35
  4.3 Agency investment ............................................................................................................. 36
  4.4 User costs ........................................................................................................................ 36
  4.5 Mathematical model ......................................................................................................... 37
  4.6 Numerical results ............................................................................................................. 39
  4.7 Conclusion ......................................................................................................................... 46

CHAPTER 5 FUTURE RESEARCH ............................................................................................ 47
  5.1 Flexible hybrid system with local tubes .......................................................................... 47
  5.2 Semi-flexible transit system ............................................................................................ 49
  5.3 Flexible modes ................................................................................................................. 50
  5.4 Other ............................................................................................................................... 51

CHAPTER 6 CONCLUSIONS ..................................................................................................... 53

REFERENCES ............................................................................................................................ 55

APPENDIX A PROOFS ............................................................................................................... 63
CHAPTER 1
INTRODUCTION

According to the U.N., world population will reach 9.6 billion persons by 2050. Besides, 3.7 billion were living in urban area in 2012 and they will be approximately 6.4 billion in 2050 (World Health Organization, 2014), which represents an increase of 73%. As a result, urban sprawl happened as an extent of the process of urbanization; in which population migrates to suburban areas around the city. Along with the urban sprawl, traffic increased in urban areas and transportation infrastructure could not support all demand (Sarzynski et al., 2006). Thus, congestion became a burning issue in urban regions, and increased delay to all users, fuel consumption, and CO2 discharge. For instance, American drivers have experienced a 150% increase on the annual delay since 1982 (U.S. Department of Transportation Statistics, 2014). In addition, oil price has tripled since the early 2000’s (U.S. Energy Information Administration, 2014), putting even more pressure on customers’ expenses. Schrank et al (2012) estimated that on average, an auto commuter experiences a yearly delay of 3.8 hours, discharges 380lbs of CO2, and wastes a total of $818 in congestion. While population continues to increase, congestion will continue to spoil urban areas and harm commuters, unless viable solutions are found.

Modal shifting is one of the possible solutions to mitigate traffic. It refers to commuters switching from private vehicles to public transportation. Several modes are available to the users (e.g. bus, rail, Metro, ferry) and fit different situations. Bus is by far the most common mode used in the U.S. and serves approximately 50% of the passenger trips. Between 2004 and 2012, in the U.S., transit use has grown 15% whereas highway ridership has decreased 1% (American Public Transportation Association, 2013). This proves that users have already started shifting from driving to using public transit. If a transit system ensures service quality in terms of comfort, safety and travel time reliability, commuters will be more likely to use it. Hence, by developing reliable and sustainable transit networks, ridership will certainly
continue to increase.

To design a viable, efficient and green transit system, essential decisions have to be made regarding service mode, network structure, coverage and operating mode. Different network structures have been investigated: grid networks (Holroyd, 1965); radial systems (Byrne, 1975); corridors (Wirasinghe et al., 1977); hub-and-spoke (Newell, 1979); hybrid system (Daganzo, 2010a). These works managed to design a transit system in terms of stop spacing, service frequency, and, if need be, line spacing, which minimize the total system cost corresponding to the user travel time and the agency investment.

Fixed routes are usually designed for bus transit systems and are known to perform well in dense population areas. Yet, in low demand areas, optimal spacing tends to be larger and then exposes users to risks. Quadrifoglio et al. (2006) assessed the performance of a flexible shuttle service, and proved possible benefits of a system in which vehicles pick up and drop off passengers. While focusing on the hybrid network structure (Daganzo, 2010a), Nourbakhsb and Ouyang (2012) also established advantages to flexible transit system under low-to-moderate homogeneous demand levels. Such a system eliminates access time, since passengers do not need to walk to and from bus stations, and reduces possible infrastructure investments. In addition, it can be designed for operating modes such as “paratransit” or “safe ride”. These modes are mostly aimed at elderlies and disable people who are not able to commute to the bus stops. Yet, most of the demand responsive systems have not been optimized.

Typically, urban areas dictate transit trips distribution and result into heterogeneous travel demand. While traditional bus networks were designed under heterogeneous demand using continuum approximation (Ouyang et al., 2014) or flexible frequency (Daganzo, 2010b), few are known about flexible transit design under heterogeneous demand. One might think that the system could be improved by making allowance for the local conditions and thus by designing a heterogeneous network structure.

The objective of this thesis is to (i) investigate how the design and operations of a flexible transit system are affected by heterogeneous demand, in the case of a grid and a hybrid structure, and to (ii) use a continuum approach to improve the flexible transit system under heterogeneous demand and reduce generalized costs by inserting local tubes within the main ones. For that purpose, we will also consider several spatial demand distributions within
range provided by Nourbakhsh and Ouyang (2012) and analyze the transit system structure for those scenarios. First we will design a grid network with flexible transit and then expand this work to the hybrid structure provided by Daganzo (2010a). Next, we will explore another feature, branching, allowing local tubes to better serve regions with higher demand density.

This thesis is organized as follows. Chapter 2 reviews previous work related to traditional and flexible transit design. Chapter 3 presents analytical formulae derived for user costs and agency investment considering spatially heterogeneous demand on both a grid and a hybrid transit network. Chapter 4 will introduce the local tube system, considering the grid transit system. Finally, ideas for future work are provided in Chapter 5.
CHAPTER 2
LITERATURE REVIEW

This chapter provides an overview of existing flexible transit systems around the world, and major contributions in transit network design and operations. Then, continuum approximation methods are introduced along with solution methods for transit design problems. Finally, passenger demand literature and solution methods are presented.

2.1 Flexible transit description and case experiences

2.1.1 Adaptive transit

Adaptive transit is a response to population growth patterns. It adapts mass transit services and technologies to better serve spread-out areas. Cervero and Beutler (1999) divide adaptive transit into three classes: technological innovations, bus-based service reforms, and small vehicle service reform. Table 2.1 provides an overview of these classes along with examples and real-world implementations.

Technological innovations allow vehicles to achieve higher speeds and add flexibility to traditional transit services. They can be divided into two types: Bus Rapid Transit and Flexible technologies (e.g. track-sharing). Bus Rapid Transit, or BRT, features exclusive rights of way, signal priority, off-board fare collection and advanced technologies that increase the bus speed and the level of service. France is one of the most advanced countries in terms of BRT network; with in particular a 170-km network of exclusive busways and bus lines (Cervero and Beutler, 1999). Another example of technological innovation involves track-guided buses. Buses are steered along dedicated tracks to achieve high speed but can also exit the guideway to operate as regular surface street buses and to function as distributors. Germany and UK have implemented guided buses systems.
The second class of Adaptive transit involves the rearrangement of bus routes to provide more direct service to the users and reduce travel time. Bus-based service reforms include seamless transferring, tangential routing and flexible transit. Time-transfer systems were first implemented in Edmonton and Calgary, Canada, where suburban lines are synchronized with base routes between city centers to reduce transfer time. Tangential routing is another form of bus route reorganization, performed in numerous cities in the United States. Rearranged routes provide suburb-to-suburb service and buses operate along freeways or beltways, as in Montgomery County, Maryland; San Diego, Riverside, California; or Washington D.C. The last type of bus-based service reform, flexible transit, will be developed in Section 2.1.2.

<table>
<thead>
<tr>
<th>Classes and types</th>
<th>Examples</th>
<th>Case Experiences</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1. Technological innovations</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bus Rapid Transit</td>
<td>Busways, priority schemes</td>
<td>Quebec (Canada), Curitiba (Brazil)</td>
</tr>
<tr>
<td>Flexible technologies</td>
<td>Track-guided buses, rail track-sharing</td>
<td>Adelaide (Australia), Karlsruhe (Germany)</td>
</tr>
<tr>
<td><strong>2. Bus-based service reforms</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Seamless transferring</td>
<td>Timed transfers, coordinated scheduling</td>
<td>Edmonton (Canada), Portland, OR</td>
</tr>
<tr>
<td>Tangential routing</td>
<td>Crosstown surface and express routes</td>
<td>Houston, TX, San Jose, CA</td>
</tr>
<tr>
<td>Flexible routing</td>
<td>Route deviation, dial a ride</td>
<td>Ft Worth, TX, Broward County, FL</td>
</tr>
<tr>
<td><strong>3. Small-vehicle service reforms</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shuttles</td>
<td>Home-end shuttles, work-end shuttles</td>
<td>Albany, NY, Contra Costa County, CA</td>
</tr>
<tr>
<td>Circulators</td>
<td>Activity-center circulators</td>
<td>Montgomery County MD</td>
</tr>
<tr>
<td>Zonal networks</td>
<td>Cellular services</td>
<td>Tidewater, WA, Hamilton, OH</td>
</tr>
<tr>
<td>Subscription services</td>
<td>Vanpools, club buses</td>
<td>Seattle, WA, Houston, TX</td>
</tr>
<tr>
<td>Paratransit</td>
<td>Jitney, shared-ride taxis</td>
<td>Mexico City, San Juan PR, Miami, FL</td>
</tr>
</tbody>
</table>

Table 2.1. Adaptive transit and case experiences

The third class of Adaptive transit takes advantage of small-capacity vehicles (e.g. van, taxi, minibus). Such vehicles provide more flexibility and passengers are less likely to experience delay due to boarding and alighting. The most popular form is paratransit, which features flexible-route service. It consists of privately owned minibuses or jitneys. Although some of
those services are open to general public, they are usually designed for assisting elderlies and
disables. Besides, some US cities also allow private minibus and jitneys operators (New York
City, Miami) to ply their trade under certain conditions. A promising advance in paratransit
operations was experimented in Germany, where centralized computers take real-time
decision to best serve waiting customers. For a comprehensive review on the “smart
paratransit”, see Behnke (1993). Other examples of small-vehicle systems include circulators
that serve activity centers such as college campuses, commercial centers, and medical centers
(Contra Costa County, CA), and subscription vans (Seattle, OR; Houston, TX).

2.1.2 Flexible route transit

Typical flexible transit, also known as “demand responsive transportation” (DRT), allows
drivers to make detours according to passenger requests, either along a fixed route or around
designated checkpoints or within a corridor. Figure 2.1 illustrates some modes of flexible
transit service. There are a total of 6 modes: route deviation, point deviation, demand-
responsive connector, request stops, flexible-route segments, and zone route (Transportation
Research Board, 2004). In either case, flexible transit takes the following features: flexibility
in routing and scheduling, demand-responsiveness, door-to-door service, and safe access.
Another form of flexible transit, called “dial-a-ride”, operates under phone request to the
dispatcher. It consists of cars, vans, or small buses that may be equipped with wheelchair-lift.

Demand responsive systems already exist in North America, in Europe, in Australia and in
Japan. Some of those are displayed in Table 2.2. In Europe (Switzerland, Luxembourg) and
in Australia, they usually serve specific rural areas, where population density is low and no
other public transit system is available. In the UK, DRT links rural areas with district centers
or train stations (Enoch et al., 2006). In North America, similar systems exist in Washington
State, in California, and in Missouri. In Washington D.C. however, flexible transit is part of
the global transit system. Some lines operate in fixed routes in urban areas while they provide
door-to-door service in the suburbs. In Winnipeg, Canada, a low-density city, a Bus Rapid
Transit has been implemented and provides flexible transit to the users at nonpeak times, in
addition to the existing fixed-routes transportation system (Transportation Research Board,
2004).
2.1.3 Benefits

One of the major benefits of DRT is the elimination of the access time since users are picked up directly at their origin and dropped off at their destination. Because of some particular conditions (disability, night time, adverse weather), significant discomfort can be perceived as commuters are walking to the bus stop. It also solves safety issues that may arise at night.

In addition, literature provides a wide variety of other benefits for flexible transit systems, including increasing ridership (Durvasula et al., 1998), and reducing the burden associated with transferring (Wachs, 1976; Cevero and Beutler, 1999). In addition, Pratelli claimed that flexible transit are still beneficial for serving areas with demand densities too high for door-to-door services but not high enough for fixed-route service, and Farwell stressed that flexible transit allows to combine the regularity of fixed-route service with the flexibility of demand-responsive services (as cited in Transportation Research Board, 2004). Quadrifoglio and Li (2009) and Nourbakhsh and Ouyang (2012) also gave credit to flexible transit systems in low-demand areas. Comparative performances with other systems (e.g. fixed-route system, taxis) were investigated in terms of user costs and agency investments.
2.2 Transit network design

2.2.1 Transit network design problems

The transit network design problem (TNDP) has been studied by researchers since the early 1970s. Usually, such problems are formulated as linear or non-linear programming with an objective function and a set of constraints. Objective function generally consists in total welfare maximization; profit maximization; travel cost minimization; or capacity maximization. See Van Nes and Bovy (2000) and Kepaptsoglou and Karlaftis (2009) for a review of different objective functions in the field of transit network design. The former investigates the impact of different objectives on the optimal value of design variables and the latter reviews research on transit network design based on objective function in particular. Typical constraints include variables ranges, minimum and maximum occupancy (Estrada et al., 2011), maximum fleet size (Ceder and Wilson, 1986; Fan and Machemehl, 2004), and maximum budget (Dubois et al., 1979; Chien et al., 2001; Bagloee and Ceder, 2011). Variables are usually related to the network structure (route and stop spacing, zone size) and the vehicle dispatch (frequency, fleet size).

As for fixed-route transit, numerous system structures have been investigated. Only the following will be addressed in this section: (i) feeder; (ii) grid; (iii) radial; (iv) hub-and-spoke; and (v) hybrid. Daganzo (2010b) provided a comprehensive review of most of those structures.

Hurdle (1973) found optimal line spacing and headway between buses for a parallel feeder lines system by minimizing agency and users costs. Kuah and Perl (1988) expanded on Hurdle’s work by determining the three major design variables: route spacing, stop spacing and headway. Flexibility was also studied since stop spacing was first allowed to vary from a route to another and then on a single route. Tsao and Schonfeld (1984) developed a branched zonal transit service. Design variables are zone boundaries, headway and local route lengths. The system is optimized by minimizing operator costs and users costs.

Holroyd (1965) and Byrne (1976) first studied rectangular areas. The former investigated a grid system and the former calculated optimal design variables for a set of parallel bus routes in a rectangular area. Imam (1998) designed a grid network and determined optimal route spacing, frequency, and fare, with and without vehicle capacity constraints. More recently, irregular grid layouts have been developed (Chien and Schonfeld, 1997; Chien et al., 2003;
Ouyang et al., 2014). The route spacing may vary over space because of street location, heterogeneous demand, and other variable zone characteristics. Thus, Chien and Schonfeld (1997) divided the service area into multiple zones with different demand levels, geometry conditions, and other characteristics and optimized the network variables for each zone. Chien et al. (2003) designed a grid network based on preexisting street layout that also includes diagonal routes. Finally, Ouyang et al. (2014) implemented a grid network that allows routes to form local lines in areas with higher demand density. The high density of routes mainly reduces access costs in higher demand neighborhoods. The ideas presented in this article will be largely used in this thesis.

Researchers also focused on radial structures (Byrne, 1975; Wirasinghe et al., 1977; Vaughan, 1986). Byrne (1975) designed a radial transit lines system and determined optimal headway and radial lines positions. Wirasinghe et al. (1977) extended this work by adding direct bus service. Under mode competition, stop spacing, feeder bus zone boundaries, and headway were optimized. Vaughan (1986) investigated a radial system with radial and ring routes and found optimal stop spacing and headway between buses.

Other networks include hub-and-spoke (Newell, 1979) and hybrid structure (Daganzo, 2010a; Smith, 2014). Newell (1979) discussed the benefits of a hub-and-spoke system over a basic grid system in terms of travel time and operation costs. Yet, design variables were not determined. Daganzo (2010a) designed a hybrid structure that combines grid system in the central area and hub-and-spoke system in the peripheral area. Headway, stop spacing and zone size are optimized while minimizing both agency and user metrics (infrastructure length, operation costs, travel time, number of transfers).

2.2.2 Flexible-route systems

According to TRB (2004), four main elements are to be determined in order to design a flexible transit system: (i) where vehicles operate; (ii) boarding and alighting locations; (iii) schedule; and (iv) advance notice requirements. This thesis will mainly focus on (i), (ii) and to a lesser extent (iii). (i) pertains to the operation modes mentioned in Section 2.1.2. (ii) refers to the status (checkpoint, required but non timed stop, optional stops) and the location of the stops along the route or within the zone of operation. (iii) is a combination between scheduled times and times determined by demand. (iv) is related to the request that
passengers make for pick-up or drop-off at their desired location. Additional parameters involve fare determination or coordination with other services.

In order to determine these previously mentioned elements, there has been significant research for the different flexible modes. Farwell and Marx (1996) designed a request stops system (“flex-route”) in Northern Virginia. Planning, implementation and evaluation of the system are presented. Fu (2002) also proposed an analytical model for flex-route service design and gave an insight of the relationship between the system performance and design parameters. User and operator costs are derived and a simulation analysis is provided to address the limitations of the model. Additionally, Malucelli et al. (1999) introduced a Demand Adaptive System with checkpoint and optional stops (request stop system) and developed models for various operation strategies characterized by the way to treat the service requests. Durvasula et al. (1998) developed a feasibility study to implement a route deviation system in Southern Virginia. The study includes route selection, software design for scheduling and dispatch, and issue discussion. Li et al. (2007) designed a hybrid DRT system in California. The system presented involves a set of checkpoint the vehicles may or may not visit when operating while it operates as a fixed-route transit system in some portions of the route. In addition, Daganzo (1984) presented a comparative study in terms of waiting and riding times for a similar system (dial-a-ride with checkpoints) with fixed-route transit and typical dial-a-ride system. It also provides guidelines to implement transit systems according to the demand level. Nourbakhsh and Ouyang (2012) developed a model to optimize a zone route system with transfer points implemented in Daganzo hybrid network (2010a). User costs and agency investments are derived and performance of the system is compared with those of other transportation system structures. This work will serve as a base for this thesis. Aldaihani et al. (2004) designed a hybrid transit network that involves a grid fixed-route system and a flexible zone route service. The grid system partitions the service zone into squares where on-demand vehicles operate. By minimizing the total cost function, the optimal number of zones is determined.

Researchers also focused on scheduling problems in the case of a dial-a-ride system. Savelsbergh and Sol (1995) provided a detailed review of the pickup and delivery problem that is intended for a DRT system. Wilson and Hendrickson (1980) also presented a review of various service models for DRT systems. Quadrifoglio et al. (2007) developed a solution
method for scheduling in zone route system with checkpoints. The insertion heuristic is proved to provide near-optimal solution for low-demand scenarios.

<table>
<thead>
<tr>
<th>Service name</th>
<th>DART</th>
<th>Omnilink</th>
<th>PubliCar</th>
<th>Flex route</th>
<th>MetroAccess</th>
<th>Cango</th>
</tr>
</thead>
<tbody>
<tr>
<td>Location</td>
<td>Winnipeg, Canada</td>
<td>Stafford and Prince William County, Virginia</td>
<td>Switzerland</td>
<td>Minnesota Valley, Minnesota</td>
<td>Washington DC</td>
<td>Hampshire County, UK</td>
</tr>
<tr>
<td>Flexible mode</td>
<td>Demand responsive connector</td>
<td>Route deviation</td>
<td>Zone route</td>
<td>Route deviation</td>
<td>Zone route</td>
<td>Point deviation</td>
</tr>
<tr>
<td>Area type</td>
<td>Suburban area</td>
<td>Urban area with very low population density</td>
<td>Rural area</td>
<td>Suburban area</td>
<td>Urban area</td>
<td>Rural area</td>
</tr>
<tr>
<td>Days of operation</td>
<td>Weekday evenings and weekends</td>
<td>Weekday</td>
<td>Weekday</td>
<td>Weekday</td>
<td>All week</td>
<td>Monday-Saturday</td>
</tr>
</tbody>
</table>

Table 2.2. Implementations of flexible transit

2.3 Continuum approximation

The design of public transit systems involves cumbersome inputs because of specific on-site details that create large discrete data sets. Furthermore, costly algorithms are required to treat complex problems. Therefore, researchers developed alternative strategies to simplify the problem in terms of keys decision variables, and to find near-optimal solutions. Continuous models allow to reduce the number of variables, for instance from a set of departure times to a single headway, or from a set of stop locations to a single spacing, while giving useful insight over long-horizon decisions.

Thus, continuum approximation models have been used extensively for transportation systems optimization. They have been used to optimize vehicle dispatch (Newell, 1971; Clarens and Hurdle, 1975), to solve facility location problems (Newell, 1973) and vehicle
routing problems (Francis and Smilowitz, 2006), to design public transit networks (Hurdle, 1973; Byrne, 1975; Wirasinghe et al., 1977; Wirasinghe and Ghoneim, 1981; Chang and Schonfeld, 1991; Daganzo, 2010a; Ouyang et al., 2014), and for various logistics problems (Daganzo, 2005). A detailed review of CA models for logistics systems design as of 1996 can be found in Langevin et al. (1996).

Daganzo et al. (2012) discuss the major benefits of models with only few variables: fewer data requirements reduced computational complexity, improved system representation, transparency, and insightfulness. Therefore, this thesis will follow a continuum approximation model to address the impact of spatially heterogeneous passenger demand on the design of a flexible transit system.

2.4 Heterogeneous demand literature

For simplicity, continuous models tend to treat trip demand as homogeneous over both time and space. However, such settings rarely occur in reality. For instance, travel demand tends to be concentrated in rush hours and near centers of high interest. As for temporal variations, Newell (1971) considered a travel demand that is a smooth function of time to optimize a bus dispatch. His work was then extended in Salzborn (1972), using a two-stage methodology that minimizes the fleet size and then passenger waiting time. Hurdle (1973) also used a time-dependent demand to optimize a network of feeder bus lines. Clarens and Hurdle (1975) analyzed a commuter bus system with a temporally heterogeneous passenger demand. They proposed a two-zoning operation, one during peak and one during off-peak. Additionally, Chang and Schonfeld (1991) designed a multiple period model to design a bus transit system, while the demand varies over a daily cycle. Multiple period models allow some characteristics of the system to change within each period (e.g. headway), whereas others remain constant over the cycle. Yet, Daganzo (2010b) established that if the demand does not change quickly with time, the system design would remain a near-optimal solution. Adapting the vehicle dispatch frequency is a simple strategy that will barely affect the optimal cost found with homogeneous settings.

Spatial demand disparity can be more problematic and have a greater impact on the network design. Byrne (1975) and Vaughan (1986) considered a demand density that varies radially to determine optimal headway and spacing for a radial transit network. Hurdle (1973) also
allowed travel demand to vary along the feeder line, although he did not provide any numerical example. In addition, Wirasinghe and Ghoneim (1981) optimized the stop spacing along a bus route with a non-uniform demand distribution. Demand was modeled with cumulative functions of origin and destinations along the route. Daganzo (2010b) treated heterogeneous demand along corridors using an OD matrix of trip selection rates. Byrne (1976) addressed cost formulation for grid systems under spatially heterogeneous demand. Yet, a numerical example was only provided for homogeneous demand. Another strategy regarding spatially heterogeneous demand is to split the service region into homogeneous zones that will have specific demand levels. Thus zonal demand can be computed by integrating the demand density function over the zone, as for homogeneous demand. Clarens and Hurdle (1975) provided a practical example of a bus system design in the region of San Francisco with a spatially heterogeneous demand density. The area was divided in zones with various demand densities to easily compute both passenger costs and agency costs. Similarly, Chien and Schonfeld (1997), Chien et al. (2003) designed a grid network with different zonal distributions. Chien and Schonfeld (1998) also developed a model for jointly optimizing a rail transit route and its feeder bus route in a rectangular region that considers spatial demand variations. Ouyang et al. (2014) considered a continuous spatial demand density to find the optimal spacing of routes in a grid system and to optimize the total cost. User costs and agency investments were expressed for local conditions and then integrated over the whole network.

2.5 Solution methods

According to researchers, the transit route network design problem is a difficult to solve optimization problem (Newell, 1979; Baaj and Mahmassani, 1991; Chakroborty, 2003). Baaj and Mahmassani (1991) observed nonlinearities and nonconvexities in the TNDP, while the discrete nature of the TNDP and its multi-objective nature increases the complexity of the problem, making it NP-hard. Chakroborty (2003) also investigated problems arising with the transit network design problem, such as discrete variables or transfer representation. In general, solution methods to transit network design problems are classified into two main categories: conventional and heuristic approaches (Kepaptsoglou and Karlaftis, 2009). The former includes analytical models and mathematical programming and addresses idealized situations. The latter exploits traditional heuristic algorithms and metaheuristics for more
practical problems. Given the large number of solution methods, the current section will only give an insight of those methods.

Most of the analytical approaches in transportation network design were addressed in the previous sections (Holroyd, 1965; Hurdle, 1973; Byrne, 1975; Clarens and Hurdle, 1975; Tsao and Schonfeld, 1984; Chang and Schonfeld, 1991; Chien and Schonfeld, 1997). However, Ceder (2001) pointed out that analytical methods were adapted to idealized situations but not for complete design. It listed the disadvantages of conventional methods such as the lack of precision, and the limited range of network sizes handled by those methods. Additionally, researchers presented mathematical programming formulations for the transit route network design problem. For instance, Ceder and Israeli (1998) proposed a nonlinear mixed-integer programming model. Constantin and Florian (1995) presented a nonlinear nonconvex programming formulation for optimizing the frequencies in a transit network. However, in most cases, heuristics were used to solve the problem. Thus, Chakroborty (2003) discussed the reasons why traditional methods have difficulties to solve the problem: numerous discrete decision variables, nonlinearity, and logical conditions.

Therefore a second category of solution methods appears as a credible alternative to conventional approaches. It contains numerous heuristic algorithms such as genetic algorithm (Pattnaik et al., 1998; Chakroborty and Dwivedi, 2002), local search (Dubois et al., 1979), simulated annealing (Fan and Machemehl, 2006), or hybrid heuristics (Zhao et al., 2005). Typical procedures include candidate route generation and route configuration. A set of candidate routes is generated using a heuristic algorithm with respect to some criteria (Baaj and Mahmassani, 1995) and then a subset of optimal routes is selected using another heuristic (Mauttone and Urquhart, 2009).
CHAPTER 3
FLEXIBLE TRANSIT SYSTEM DESIGN COMPARISON
UNDER SPATIALLY HETEROGENEOUS DEMAND:
GRID NETWORK VERSUS HYBRID NETWORK

Travel patterns are typically not uniform in urban areas, or suburban areas. Centers of interest (e.g. CBD, university, shopping malls, large-lot residential tract housing) are likely to attract the largest part of transit users. As a result, agency needs to adapt service to demand disparity in order to serve passengers properly. Thus, designing transit system accordingly to demand spatial variations is crucial. Therefore we will design a flexible transit system on a grid structure with main tubes in response to a heterogeneous demand. Then we will investigate the advantages of a hybrid system with respect to the grid system.

This chapter seeks to derive analytical formulations of the generalized cost to the system’s passengers and its operating agency in terms of key parameters and decision variables based on both grid and hybrid networks, while using a continuous demand density function based on Ouyang et al. (2014), to capture the spatial heterogeneity of the demand. Since the grid network is part of Daganzo hybrid structure (2010a), the results found for the former structure will serve as a basis for the mathematical derivation in the case of the latter structure.

The chapter is organized as follows. First, the grid network problem is introduced in Section 3.1.1 along with important characteristics and parameters. Next, Section 3.1.2 displays the derivation of agency investment and user costs. Section 3.1.3 exhibits the mathematical model obtained from the previous costs computation and the solution method. Section 3.1.4 presents the optimal grid transit network and optimal frequency for different spatial trip distributions. Hybrid transit system is introduced in Section 3.2.1, and cost metrics are
derived in Section 3.2.2. Section 3.2.3 exhibits the mathematical model obtained from the previous costs computation and the solution method. Section 3.2.4 presents the optimal hybrid transit network and optimal frequency and allows for comparison with the grid system. Finally, Section 3.3 concludes this chapter.

3.1 Grid network

3.1.1 Problem settings

3.1.1.1 Grid network characteristics

The transit service region is a square of side $d$. Unlike fixed-route system where buses travel along a route and serve passengers at predetermined stops, flexible transit allows buses to operate within a narrow elongated area, called “bus tube”, to pick up and drop off passengers. The square is essentially divided into two hemispheres each containing $N$ bus tubes. Local streets in the region are assumed to form a grid network with constant spacing $s$. We further assume that buses only have longitudinal and lateral movements along the local streets network and sweep back and forth through the tube, as shown in Figure 3.1.

![Figure 3.1. Flexible transit structure (adapted from Nourbakhsh and Ouyang (2012))](image)
3.1.1.2 Passenger trip

Although buses travel without fixed routes or predetermined stops, we assume that transfer points are planned along each main tube. Hence, there are a total of $N^2$ transfer stops inside the central region (see Figure 3.1b).

A typical passenger trip can be described as follows. A passenger is picked up directly at his origin, travels as directly as possible and with the least transfers, and is dropped off either at his destination or at a transfer point in case a transfer is necessary. Whenever there is a tie among multiple alternatives, passengers randomly choose their bus route. Unlike traditional fixed-route systems, there is no walking time.

3.1.1.3 Demand distribution

Unlike Nourbakhsh and Ouyang (2012), this paper does not assume a uniform demand distribution. Instead, a heterogeneous demand is borrowed from Ouyang et al. (2014) to describe trip distribution from origin $(x_1, y_1)$ to destination $(x_2, y_2)$, with density $\delta(x_1, y_1, x_2, y_2)$ in units of passenger per time per unit area.

$$\delta(x_1, y_1, x_2, y_2) = \prod_{i=1}^{2} \left( a_1 + a_2 \sum_{j=1}^{2} \exp \left[ -(a_{3j}x_i - b_{ij})^2 - (a_{4j}y_i - b_{ij})^2 \right] \right)$$

(3.1)

where additional parameters $a_1, a_2, a_{31}, a_{32}, a_{41}, a_{42}, b_{11}, b_{12}, b_{21}, b_{22}, b_{11}, b_{12}, b_{21}, b_{22}$ will be addressed in Section 3.1.4. Several cases will be studied: mono-centric city, twin cities, asymmetric city, and commuter city.

From now on, the total number of passenger on the transit system during one hour will be defined by $T_{pax} = \iint_{d \times d} \iint_{d \times d} \delta$. As shown in Quadrifoglio and Li (2009) and in Nourbakhsh and Ouyang (2012), flexible-route transit system is relatively desirable under low demand (approximately 4 to 40 passengers/h/km²), while fixed-route transit system is desirable under high demand. Therefore, we will further consider low-to-moderate demand rates.

Additionally, we need to compute the expected rate per unit area per time that passengers start and end their trip at point $(x, y)$. 
\[ D_{\text{start}}(x, y) = \int_{\tilde{y}=0}^{d} \int_{\tilde{x}=0}^{d} \delta(x, y, \tilde{x}, \tilde{y}) \, d\tilde{x} d\tilde{y} \]  
\[ D_{\text{end}}(x, y) = \int_{\tilde{y}=0}^{d} \int_{\tilde{x}=0}^{d} \delta(\tilde{x}, \tilde{y}, x, y) \, d\tilde{x} d\tilde{y} \] (3.2)

3.1.2 Problem formulation

3.1.2.1 Agency investment

Basically, a bus travels in a single tube with headway \( H \), and cruising speed \( v \). It takes \( \tau_1 \) to make a stop (delay due to acceleration and deceleration) and \( \tau_2 \) to pick up or drop off passengers. Hence the time for a bus to stop is \( \tau = \tau_1 + \tau_2 \). Therefore, its average travel speed reduces to \( v_c < v \). Passengers can walk at speed \( v_w \) and \( \beta \) is the factor capturing discomfort associated with transfers.

Agency costs include the expected total vehicle distance traveled per hour of operation, \( Q \), and expected total fleet size in operation, \( M \). Also, since the buses do not operate on fixed routes, the infrastructure investment is minimal. We will assume that need for user interface (e.g. phone service systems, driver communication devices) has a negligible cost compared to traditional infrastructures. The formulae of the agency cost metrics are presented next and are derived in Appendix A (Result 1-3).

\[
Q \approx \frac{1}{H} \sum_{k=1}^{2N} \left[ 2 \left( d + \frac{1}{3} (H) \left( \frac{d}{H} \int_{T_k} (D_{\text{start}} + D_{\text{end}}) \right) \right) + \sum_{i=1}^{\infty} (i-1) \left( \frac{1}{2} \sum_{(x,y) \in T_k} P_{x,y}(i) \right) \right] 
\] (3.3)

where \( P_{x,y}(i) = \frac{(\|D_{\text{start}}(x,y) + D_{\text{end}}(x,y)\|H)^i}{i!} e^{-[\|D_{\text{start}}(x,y) + D_{\text{end}}(x,y)\|H]} \)

\( T_k \) denotes tube \( k \).

Then the bus fleet size is given by

\[ M = \frac{Q}{v_c} \] (3.4)
where the bus average travel speed is given by

\[
\frac{1}{v_c} = \frac{1}{v} + \tau T_{pax} \left( \frac{2 + e_T}{Q} \right)
\]  

(3.5)

and \( e_T \) denotes the expected number of transfers. It will be formulated in the next section.

3.1.2.2 User costs

User costs include the total time for passengers to travel from their origin to their destination, and other comfort-related parameters (e.g. transfers). To compute the user costs, we need to determine the average waiting time per passenger, \( W \), the expected travel distance per passenger, \( E \), the expected in-vehicle travel time per passenger, \( IVTT \), and the maximal vehicle occupancy, \( O \). All are given by the following formulae and derived in Appendix A (Result 4-6).

\[
W = \frac{H}{2} (1 + e_T)
\]  

(3.6)

where the expected number of transfers is given by \( e_T = \lambda^{(1)} / T_{pax} \) where \( \lambda^{(1)} \) is the number of passenger transferring exactly once. Basically, passengers either do not transfer or transfer once. Therefore, \( \lambda^{(1)} \) is given by \( \lambda^{(1)} = T_{pax} - \lambda^{(0)} = T_{pax} - \sum_{i=1}^{2N} \left[ \int \int_{T_i} \int \delta \right] \).

\[
E \approx \frac{2d}{3} \times \frac{QH/2N}{2d} = \frac{QH}{6N}
\]  

(3.7)

\[
IVTT = \frac{E}{v_c}
\]  

(3.8)

\[
O = \frac{HT_{pax}}{2N}
\]  

(3.9)

Although the maximum expected occupancy is not directly related to user costs, it could be helpful for the transit agency to assess the bus capacity needed when designing the transit system.
3.1.3 Mathematical model

The design problem comes down to find optimal combination of decision variables \( H \) (headway) and \( N \) (number of tubes in one hemisphere). For homogeneity, agency investments need to be converted into time equivalent. We let $Q$ denote the agency operation cost, $M$ denote the agency cost per vehicle-hour, and \( \mu \) the average value of one passenger-hour. Then $\pi_Q = \frac{Q}{T_{\text{pass}}\mu}$ and $\pi_M = \frac{M}{T_{\text{pass}}\mu}$ convert the corresponding agency costs into travel time equivalents. Also $\beta$ evaluates the travel discomfort associated with the expected number of transfers, and then $\frac{\beta}{V_T}$ converts the number of transfers into passenger travel time equivalent (Nourbakhsh and Ouyang, 2012). The objective function will be the sum of the agency investment, the user costs, and the transferring penalty, per passenger.

Hence, the optimization problem can be written as follows:

\[
\begin{align*}
\text{Min } z &= \pi_Q Q + \pi_M M + W + IVTT + \frac{\beta}{V_T} e_T \\
\text{s.t. } H &\geq 0, \quad N \in \{1, 2, \ldots, \left[ \frac{d}{s} \right] \}
\end{align*}
\]

We enforce $N \leq \left[ \frac{d}{s} \right]$ (where \( \left[ \cdot \right] \) is the floor function) to ensure that there is at least one local street within each tube, so that each bus can run within its assigned tube. It also ensures that there is at least one transfer point within the system.

3.1.4 Numerical results

In this section, the optimal grid transit network is designed for a square city with \( d = 10 \text{ km} \) and in the case of four distinct spatial demand distributions: (i) a mono-centric city, where origins and destinations of transit trips are clustered in the center; (ii) a twin city, where demand is clustered in two adjacent regions near the city center; (iii) an asymmetric mono-centric city, where origins and destinations are clustered on the edge of the city; and (iv) a commuter city, where two distinct clusters, one for trip origins and the other for trip destinations, occur in two distinct centers. Demand distribution parameters are borrowed from Ouyang et al. (2014) and presented in Table 3.1. In order to reach the low-demand scenario, Equation (3.1) was divided by 5, for a base demand of 2,000 passengers/hr. The parameters values used to compute the generalized cost are displayed in Table 3.2. They are
mostly taken from Daganzo (2010). It suggests that $\mu$ should be similar to the prevailing wage. A few are borrowed from Nourbakhsh and Ouyang (2012). Note that $v_w$ is usually 3-6 km/hr but here we also account for the discomfort associated with walking.

Figure 3.2 illustrates the marginal distribution of trip origins and destinations for the different demand distributions. The combination of boarding and alighting is obtained with the expected rate per unit area that passengers start and end their trip at point $(x,y)$ computed in (3.2). The ratio of the highest demand neighborhood over the lowest demand neighborhood is about 10, which is typical of metropolitan areas in the U.S.

<table>
<thead>
<tr>
<th>Case</th>
<th>Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$a_1$</td>
</tr>
<tr>
<td>(i)</td>
<td>-5.41</td>
</tr>
<tr>
<td>(ii)</td>
<td>0.17</td>
</tr>
<tr>
<td>(iii)</td>
<td>-6.09</td>
</tr>
<tr>
<td>(iv)</td>
<td>-8.35</td>
</tr>
</tbody>
</table>

Table 3.1. Demand distribution parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>$s$ (km)</th>
<th>$v$ (km/hr)</th>
<th>$v_w$ (km/hr)</th>
<th>$\tau$ (s)</th>
<th>$S_M$ ($/\text{veh-hr}$)</th>
<th>$S_Q$ ($/\text{veh-km}$)</th>
<th>$\mu$ ($$/\text{hr})</th>
<th>$\beta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Value</td>
<td>0.15</td>
<td>25</td>
<td>2</td>
<td>13</td>
<td>40</td>
<td>2</td>
<td>20</td>
<td>0.03</td>
</tr>
</tbody>
</table>

Table 3.2. Input parameters for the model
Given those parameters, the model described by (3.10) was programmed into Matlab and run on a 3.3 GHz CPU and 8 GB memory desktop computer. We compute the optimal headway and optimal number of tubes. These results are shown in Table 3.3 along with the cost and performance metrics.

We do not observe significant difference between the demand distributions in terms of cost metrics. In the four cases, the in-vehicle travel time has the highest contribution in the generalized cost. However, the expected number of transfers varies according to the trip demand distribution. Because the discomfort cost associated with transferring has a negligible impact on the generalized cost, there is finally no difference in the optimal design and total cost.
Table 3.3. Results summary for several heterogeneous distributions

<table>
<thead>
<tr>
<th>Demand distribution</th>
<th>Mono-centric city</th>
<th>Twin cities</th>
<th>Asymmetric city</th>
<th>Commuter city</th>
</tr>
</thead>
<tbody>
<tr>
<td>( T_{\text{pax}} )</td>
<td>2,000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Decision variables</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H (hr)</td>
<td>0.20</td>
<td>0.20</td>
<td>0.20</td>
<td>0.20</td>
</tr>
<tr>
<td>N</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>Agency metrics</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \pi_Q )</td>
<td>0.27</td>
<td>0.28</td>
<td>0.28</td>
<td>0.27</td>
</tr>
<tr>
<td>( \pi_M )</td>
<td>0.24</td>
<td>0.24</td>
<td>0.24</td>
<td>0.24</td>
</tr>
<tr>
<td>User metrics</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>W</td>
<td>0.19</td>
<td>0.19</td>
<td>0.19</td>
<td>0.20</td>
</tr>
<tr>
<td>IVTT</td>
<td>0.40</td>
<td>0.40</td>
<td>0.40</td>
<td>0.40</td>
</tr>
<tr>
<td>( \frac{\beta}{\nu_W} )</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
</tr>
<tr>
<td>Performance metrics</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \nu_c )</td>
<td>22.8</td>
<td>22.8</td>
<td>22.8</td>
<td>22.8</td>
</tr>
<tr>
<td>( e_r )</td>
<td>0.93</td>
<td>0.86</td>
<td>0.92</td>
<td>0.99</td>
</tr>
<tr>
<td>O</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>Generalized cost</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Z</td>
<td>1.12</td>
<td>1.12</td>
<td>1.13</td>
<td>1.13</td>
</tr>
</tbody>
</table>

The lack of flexibility in the network design does not allow for an appropriate design that accounts for the spatial demand disparities. Thus, there is no difference in the optimal transit system design in terms of decision variables, between the four heterogeneous distributions. Therefore, in the next section, we will present a strategy to design another transit network that allows for double coverage only in the central area, to better serve higher demand regions. Table 3.3 will serve as a benchmark for the results obtained in the next sections.
3.2 Hybrid network

3.2.1 Problem settings

3.2.1.1 Hybrid network characteristics

This section provides background for Daganzo hybrid transit structure (2010a), shown in Figure 3.3. Whereas Daganzo hybrid structure was primarily designed for fixed route networks, we will adapt the route layout for the flexible transit system. We will consider a square service region of side D that generates a heterogeneous trip pattern. Local streets in the region are assumed to form a grid network with constant spacing s. Figure 3.4 provides an overview of the hybrid flexible transit system including a central square region of side d < D and a peripheral region. Basically, the central region is similar to that described in Section 3.1. It is divided into two hemispheres each containing N bus tubes, in which buses operate with no fixed stops and travel only laterally and longitudinally along the street network. Hence, there is a total of \( N^2 \) transfer stops within the central square. \( \alpha \) denotes the ratio \( d/D \), and then \( \alpha^2 \) represents the percentage area of the total transit system with double coverage. The maximum width of the tubes is \( w = \frac{D}{N} \) at the boundary of the service region, and the width of the tubes within the central region is given by \( \alpha w \). Each bus travels in its assigned tube with headway H.

![Figure 3.3. Daganzo hybrid structure for fixed routes (adapted from Daganzo (2010a))](image-url)
3.2.1.2 Passenger trip

A typical passenger trip can be described as follows. A passenger is picked up directly at his origin, travels as directly as possible and with the least transfers, and is dropped off either at his destination or at a transfer point in case a transfer is necessary. Whenever there is a tie among multiple alternatives, passengers randomly choose their bus route. Unlike traditional fixed-route systems, there is no walking time. The major difference with the previous chapter regarding passenger trips is the expected number of transfer. Passengers travelling from the peripheral region to the peripheral region may have to experience two transfers, if those are in the same hemisphere, while the maximum number of transfer in the grid network was one.

Passenger distribution characteristics are the same as in Section 3.1. However, in this section, for the sake of simplicity, we will formulate the heterogeneous demand based on origin and destination regions. Thus, there will be four trip types: (i) Central-Central; (ii) Central-Peripheral; (iii) Peripheral-Central; and (iv) Peripheral-Peripheral. Each zone-to-zone demand rate is derived as follows.

\[ \lambda_{c-c} = \int_{y_2=LB}^{UB} \int_{x_2=LB}^{UB} \int_{y_1=LB}^{UB} \int_{x_1=LB}^{UB} \delta(x_1, y_1, x_2, y_2) \, dx_1 \, dy_1 \, dx_2 \, dy_2 \]  

\text{(3.11)}
\[ \lambda_{c-p} = \int_{y_2=0}^{UB} \int_{x_2=0}^{UB} \int_{y_1=LB}^{UB} \int_{x_1=LB}^{UB} \delta(x_1, y_1, x_2, y_2) \, dx_1 \, dy_1 \, dx_2 \, dy_2 - \lambda_{c-c} \]  
\[ (3.12) \]

\[ \lambda_{p-c} = \int_{y_2=LB}^{UB} \int_{x_2=LB}^{UB} \int_{y_1=0}^{UB} \int_{x_1=0}^{UB} \delta(x_1, y_1, x_2, y_2) \, dx_1 \, dy_1 \, dx_2 \, dy_2 - \lambda_{c-c} \]  
\[ (3.13) \]

\[ \lambda_{p-p} = \int_{y_2=0}^{D} \int_{x_2=0}^{D} \int_{y_1=0}^{D} \int_{x_1=0}^{D} \delta(x_1, y_1, x_2, y_2) \, dx_1 \, dy_1 \, dx_2 \, dy_2 - \lambda_{c-c} - \lambda_{p-c} - \lambda_{c-p} \]  
\[ (3.14) \]

\[ UB = \frac{D}{2} (1 + \alpha) \]  
\[ (3.15) \]

\[ LB = \frac{D}{2} (1 - \alpha) \]  
\[ (3.16) \]

Similar to Chien and Schonfeld (1997), Chien et al. (2003), and Smith (2014), we transform the continuous demand density into a set of discrete zone-to-zone demands. Since we are studying a low-to-moderate demand region, spatial variations of demand level will be moderate. Thus this approach efficiently approximates demand distribution and avoids cumbersome derivations for the generalized cost.

3.2.2 Problem formulation

3.2.2.1 Agency investment

All parameters related to agency investment are similar to those in Section 3.1. Agency cost is the sum of the expected total vehicle distance traveled per hour of operation Q and expected total fleet size M. Detailed derivation of these formulae can be found in Appendix A (Result 7).

\[ Q \approx \frac{1}{H} \left[ 2(2N \alpha D + 2N (1 - \alpha) D) + \frac{2}{3} \frac{aD}{N} \lambda_{c-c}H + \frac{2}{3} \frac{aD}{N} (\lambda_{c-c} + \lambda_{p-c})H + l_p (\lambda_{c-c} + \lambda_{p-c})H + 2l_p \lambda_{p-p}H + L_{extra} \right] \]  
\[ (3.17) \]

where \( l_p \) and \( L_{extra} \) are given in Appendix A (Result 7).
\[ M = \frac{Q}{v_c} \quad (3.18) \]

where \( P_{x,y} \{i\} \) is given by (3.3) and \( v_c \) is given by (3.5).

### 3.2.2.2 User costs

Similarly to Section 3.1.2, user costs include total time for passengers to travel from their origin to their destination, and other comfort-related parameters (e.g. transfers). W denotes the average waiting time per passenger, E, the expected travel distance per passenger, T, the expected in-vehicle travel time per passenger, and O the maximal vehicle occupancy.

\[
W = \frac{H}{2} (e_T + 1) \quad (3.19)
\]

where \( e_T = \frac{\lambda_{c-c}^{(1)} + \lambda_{p-c}^{(1)} + \lambda_{c-p}^{(1)} + \lambda_{p-p}^{(1)}}{T_{pax}} + 2 \frac{\lambda_{p-p}^{(2)}}{T_{pax}} \quad (3.20) \]

\[
E = \frac{q_pH}{2aN} \left[ \frac{\lambda_{c-c}^2}{T_{pax}^2} aD + \frac{\lambda_{p-p}^{11}}{T_{pax}^{12}} aD + \frac{\lambda_{p-c} + \lambda_{c-p}^5}{6T_{pax}} aD \right] + \frac{q_pH}{2D(1-a)} \left[ \frac{2 - 3a + a^3}{3} \right] D \quad (3.21)
\]

\[
T = \frac{E}{v_c} \quad (3.22)
\]

\[
O = \frac{H}{N} \max \left\{ \frac{\max \left\{ \lambda_{p-c} \lambda_{c-p} \right\}}{2}, \frac{\lambda_{p-p}}{32}, \frac{\lambda_{c-c}}{4}, \left\{ \frac{\lambda_{c-p} + \lambda_{p-c} + \lambda_{p-p}}{2} - \frac{\lambda_{p-p}}{8} \right\} \right\} \quad (3.23)
\]

The expression for \( W \) is the same as in (3.6). Other formulae are derived explicitly in Appendix A (Result 8-10).
3.2.3 Mathematical model

In Section 3.1, the decision variables were H and N. We now have to optimize the generalized cost with respect to a third variable, \( \alpha \), which denotes the ratio of the size of the central region over the size of the service region. The total cost is obtained by converting agency investment in unit of time. We will use the same coefficients \( \pi_Q \) and \( \pi_M \) as in the previous chapter. Thus, the objective function, \( z \), will be the sum of the agency cost, the user costs, and the transferring penalty, per passenger.

\[
\begin{align*}
\min z &= \pi_Q Q + \pi_M M + W + T + \frac{\beta}{v_w} e_T \\
\text{s.t.} & \quad H \geq 0, \quad N \in \{1, 2, \ldots, \frac{d}{s}\}, \quad \frac{s}{D} \leq \alpha \leq 1
\end{align*}
\]

The first two constraints are provided by (3.10). The third one ensures that there is at least one street in the central region and that the central region is not bigger than the service region.

3.2.4 Numerical results

In this section, the optimal size, structure and operation of a square city are studied under the four different spatial distributions mentioned in the previous sections: (i) mono-centric city; (ii) twin cities; (iii) asymmetric city; and (iv) commuter city. Density functions and parameters values can be found in Table 3.1, Table 3.2, and Section 3.1.4.

To solve this problem, we developed a steepest decent-based algorithm. Model (3.24) was programmed into Matlab and run on a 3.3 GHz CPU and 8 GB memory desktop computer. Table 3.4 displays optimal decision variables values and cost and performance metrics for the four spatially heterogeneous distributions previously mentioned. Table 3.5 allows for comparison with the grid structures presented in Section 3.1, and Table 3.6 allows for comparison with optimal network design under homogeneous demand (Nourbakhsh and Ouyang, 2012).

Comparing the optimal decision variables \( H, N \), and \( \alpha \), there are some noticeable differences between the four heterogeneous distributions. However, one can notice that all of the optimal generalized costs computed in this section are lower than those from the previous chapter.
The hybrid structure is more beneficial for the mono-centric city, with a 12% decrease of the total cost, since demand is concentrated within the central region. Double coverage is provided in a reduced area without penalizing the whole system. Hybrid network also allows a 10% decrease of the total cost for the twin cities. Total cost decrease for the asymmetric city and the commuter city are in the same order of magnitude (-9% and -8%, respectively), although slightly less than for the mono-centric city.

Hybrid structure decreases the number of bus tubes up to 35% (-35%, -35%, -25%, and -30%, respectively). These large reductions, due to the reduction of the double coverage, are balanced with the decrease of the headway (-20%, -20%, -15%, and -15%). Thus, each bus serves a larger tube but the frequency increases. In terms of region size ratio, $\alpha$, for cases (i) and (ii), the percentage of service area that has double coverage is reduced to 45% and 49%, while it remains higher than 50% for cases (iii) and (iv) (56% and 53% respectively). Regarding case (iii), the demand is clustered on the edge of the city but the service area that has double coverage is designed to be in the center of the service region. Regarding metrics, hybrid structure allows a reduction of all cost metrics. The main reduction occurs for waiting time that is reduced by 15 to 21%. For cases (i) and (ii), agency metrics are reduced by about 15% and 13%, while they are only reduced by 7 to 11% and 8% for cases (iii) and (iv). In-vehicle travel time is reduced by 3 to 5% in the four cases. However we can see a small increase in the expected number of transfers. But because the weight of the transfer penalty remains negligible, this does not affect significantly the generalized cost. Another interesting note is that critical occupancy increases in the four cases, especially in cases (iii) and (iv) (9%, 16%, 53%, 68%). In cases (iii) and (iv), because most of the travel demand is clustered in the peripheral region, passengers will have to transfer twice in the central region. In the grid system they were able to transfer at most once. Thus, a capacity constraint on the maximum occupancy could affect the optimal design.

Additionally, Table 3.6 provides insights on the impact of heterogeneous demand with respect to homogeneous demand (20 pax/hr/km²). First, heterogeneous happens to be beneficial to the system, since generalized costs are reduced by 8%, 6%, 5% and 4%. Though, all decision variables are increased in the heterogeneous scenarios: 7%, 7%, 13% and 13% for $H$, 30%, 30%, 50% and 40% for $N$, and 29%, 40%, 60% and 51% for $\alpha$. Regarding agency metrics, they are reduced from 4% up to 9% for cases (i) and (ii), while they are only reduced up to 4% in cases (iii) and (iv). Similarly to the previous paragraph, one can notice a
significant decrease in the transfer penalty cost. However, since the weight of this user metric remains very small in the generalized cost, it has no impact on the optimal design. As for performance metrics, the travel speed slightly increases in the heterogeneous scenarios while the expected number of transfer decreases. This relationship follows from Equation (3.5).

<table>
<thead>
<tr>
<th>Demand distribution</th>
<th>Mono-centric city</th>
<th>Twin cities</th>
<th>Asymmetric city</th>
<th>Commuter city</th>
</tr>
</thead>
<tbody>
<tr>
<td>T_{pax}</td>
<td>2,000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Decision variables</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H (hr)</td>
<td>0.16</td>
<td>0.16</td>
<td>0.17</td>
<td>0.17</td>
</tr>
<tr>
<td>N</td>
<td>13</td>
<td>13</td>
<td>15</td>
<td>14</td>
</tr>
<tr>
<td>α</td>
<td>0.45</td>
<td>0.49</td>
<td>0.56</td>
<td>0.53</td>
</tr>
<tr>
<td>π_{Q}</td>
<td>0.23</td>
<td>0.24</td>
<td>0.25</td>
<td>0.25</td>
</tr>
<tr>
<td>π_{M}</td>
<td>0.21</td>
<td>0.21</td>
<td>0.22</td>
<td>0.22</td>
</tr>
<tr>
<td>Agency metrics</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>User metrics</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>W</td>
<td>0.16</td>
<td>0.15</td>
<td>0.16</td>
<td>0.17</td>
</tr>
<tr>
<td>T</td>
<td>0.38</td>
<td>0.39</td>
<td>0.38</td>
<td>0.39</td>
</tr>
<tr>
<td>\frac{β}{v_w}e_T</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
</tr>
<tr>
<td>Performance metrics</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>v_c</td>
<td>22.4</td>
<td>22.5</td>
<td>22.6</td>
<td>22.5</td>
</tr>
<tr>
<td>e_T</td>
<td>0.96</td>
<td>0.93</td>
<td>0.93</td>
<td>0.99</td>
</tr>
<tr>
<td>O</td>
<td>10.91</td>
<td>11.63</td>
<td>15.25</td>
<td>16.77</td>
</tr>
<tr>
<td>Generalized cost</td>
<td>Z</td>
<td>0.99</td>
<td>1.01</td>
<td>1.03</td>
</tr>
</tbody>
</table>

Table 3.4. Results summary for several heterogeneous distributions
<table>
<thead>
<tr>
<th>Demand distribution</th>
<th>Mono-centric city</th>
<th>Twin cities</th>
<th>Asymmetric city</th>
<th>Commuter city</th>
</tr>
</thead>
<tbody>
<tr>
<td>Decision variables</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H (hr)</td>
<td>-20%</td>
<td>-20%</td>
<td>-15%</td>
<td>-15%</td>
</tr>
<tr>
<td>N</td>
<td>-35%</td>
<td>-35%</td>
<td>-25%</td>
<td>-30%</td>
</tr>
<tr>
<td>α</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Agency metrics</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>π₉Q</td>
<td>-15%</td>
<td>-14%</td>
<td>-11%</td>
<td>-7%</td>
</tr>
<tr>
<td>π₉M</td>
<td>-13%</td>
<td>-13%</td>
<td>-8%</td>
<td>-8%</td>
</tr>
<tr>
<td>User metrics</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>W</td>
<td>-16%</td>
<td>-21%</td>
<td>-16%</td>
<td>-15%</td>
</tr>
<tr>
<td>T</td>
<td>-5%</td>
<td>-3%</td>
<td>-5%</td>
<td>-3%</td>
</tr>
<tr>
<td>β/νₑبلاغ</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>Performance metrics</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>νₑبلاغ</td>
<td>-2%</td>
<td>-1%</td>
<td>-1%</td>
<td>-1%</td>
</tr>
<tr>
<td>eₑبلاغ</td>
<td>3%</td>
<td>8%</td>
<td>1%</td>
<td>0%</td>
</tr>
<tr>
<td>O</td>
<td>9%</td>
<td>16%</td>
<td>53%</td>
<td>68%</td>
</tr>
<tr>
<td>Generalized cost</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Z</td>
<td>-12%</td>
<td>-10%</td>
<td>-9%</td>
<td>-8%</td>
</tr>
</tbody>
</table>

Table 3.5. Percent change in metrics between grid and hybrid structure

<table>
<thead>
<tr>
<th>Demand distribution</th>
<th>Mono-centric city</th>
<th>Twin cities</th>
<th>Asymmetric city</th>
<th>Commuter city</th>
</tr>
</thead>
<tbody>
<tr>
<td>Decision variables</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H (hr)</td>
<td>7%</td>
<td>7%</td>
<td>13%</td>
<td>13%</td>
</tr>
<tr>
<td>N</td>
<td>30%</td>
<td>30%</td>
<td>50%</td>
<td>40%</td>
</tr>
<tr>
<td>α</td>
<td>29%</td>
<td>40%</td>
<td>60%</td>
<td>51%</td>
</tr>
<tr>
<td>Agency metrics</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>π₉Q</td>
<td>-8%</td>
<td>-4%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>π₉M</td>
<td>-9%</td>
<td>-9%</td>
<td>-4%</td>
<td>-4%</td>
</tr>
<tr>
<td>User metrics</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>W</td>
<td>-6%</td>
<td>-12%</td>
<td>-6%</td>
<td>0%</td>
</tr>
<tr>
<td>T</td>
<td>-5%</td>
<td>-3%</td>
<td>-5%</td>
<td>-3%</td>
</tr>
<tr>
<td>β/νₑبلاغ</td>
<td>-50%</td>
<td>-50%</td>
<td>-50%</td>
<td>-50%</td>
</tr>
<tr>
<td>Performance metrics</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>νₑبلاغ</td>
<td>4%</td>
<td>5%</td>
<td>5%</td>
<td>5%</td>
</tr>
<tr>
<td>eₑبلاغ</td>
<td>-4%</td>
<td>-7%</td>
<td>-7%</td>
<td>-1%</td>
</tr>
<tr>
<td>Generalized cost</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Z</td>
<td>-8%</td>
<td>-6%</td>
<td>-5%</td>
<td>-4%</td>
</tr>
</tbody>
</table>

Table 3.6. Percent change in metrics between homogeneous and heterogeneous for hybrid network
3.3 Conclusion

In this chapter, analytical formulations of the generalized cost for the system’s users and its operating agency were developed in terms of key parameters and decision variables for both a grid network and a hybrid network. Continuous demand distributions were used to capture the heterogeneity of the demand.

Numerical results for the design and operating frequency in mono-centric city, twin cities, asymmetric city and commuter city were presented. In a grid network, tube design and operating frequency do not depend on the shape of the distribution, since all cases presented lead to the same optimal system design. This result led to considering a system with more flexibility. Hybrid system optimal design showed some interesting results: (i) hybrid system allows a cost decrease up to 12% with respect to grid system; (ii) both agency and user metrics are reduced with a hybrid transit system; (iii) total cost decreases when the demand becomes more concentrated.

However there are a few limitations with the transit system presented in this chapter. The hybrid system enforce the double coverage to be in the center of the service region, and the width of the bus tubes is the homogeneous in the central region while demand is supposed to vary in this region. Double coverage only allows a rough flexibility by providing double service within region with higher demand density. Hence, we would like to allow more flexibility by design a system where bus tubes width depends on the local demand level. Therefore, an alternative strategy will be developed in the next chapter. It allows main tubes to split into local tubes, in regions with higher passenger demand, to facilitate the access to the transit system.
CHAPTER 4
FLEXIBLE TRANSPORT SYSTEM DESIGN UNDER SPATIALLY HETEROGENEOUS DEMAND WITH LOCAL TUBES

In the previous chapter, analytical formulations for the generalized cost to the agency and to the users were presented for a flexible transit system under heterogeneous demand in the case of grid and hybrid networks. Hybrid network demonstrated potential advantages in terms of costs and provides more flexibility to the transit network since it ensures double coverage within the central region and single coverage in the peripheral region, where the demand is assumed to be lower. In this chapter, we will focus on a third feature to build a flexible transit network. In order to allow more flexibility within the transit system, a strategy was developed to produce a heterogeneous tube layout and ultimately improve the overall system. It allows the insertion of local tubes in the zones with higher demand. Thus, this strategy ensures that local tubes are sufficiently dense in those neighborhoods. Intuitively, this strategy should be beneficial to the system compared to homogeneous grid layout. The goal of this chapter is to design a grid transit network with local tubes, under heterogeneous demand, in terms of headway, network size, and route width. The major difference is that better coverage can be provided wherever it is needed and not only within the central region as in Chapter 3.

The chapter is organized as follows. Section 4.1 describes the methodology used to design the local tubes. Section 4.2 displays the continuum approximations that are made in this chapter to compute demand rates. Next, costs metrics are derived in Section 4.3 and 4.4. Section 4.5 presents the constrained optimization model. Section 4.6 presents the optimal design for a grid network with local tubes. Finally, Section 4.7 concludes this chapter.
4.1 Methodology

In the following part we will try to improve the general grid transit system by allowing local tubes. Given the heterogeneity of the demand, one can assume that providing more service where the demand is higher will improve the overall system and reduce the optimal cost. Therefore, in the new system we will allow the main tubes to split into local tubes, similar to Ouyang et al. (2014). Main and local buses will travel back and forth along the length of their assigned routes. As main buses, local buses will also have transfer points along their road. Figure 4.1a illustrates the heterogeneous transit network with local tubes.

![Diagram](image)

**Figure 4.1. General scheme of a heterogeneous transit network**

Local buses will operate in local tubes and transfers between local and main buses will occur at the first transfer point that the local bus will meet in the main tube (also called spacing transfer). It will neither pick up nor drop off passengers in the main tube, except at the transfer point. For instance in Figure 4.1b, the local bus represented by the red dotted line does not make any stop within the main tube where the bus represented by the blue dotted line operates. Note that the bus in the main tube will operate within the other local tube. As an illustration to the worst-case passenger, consider a passenger, who starts his trip within a local tube; performs a spacing transfer to the main tube; performs a directional transfer; and finally reaches his destination within another local tube through a second spacing transfer.
Essentially, we will use a power-of-two scheme to design the grid layout (main and local tubes) provided by Roundy (1985) and displayed in Ouyang et al. (2014). Concretely, the spacing between transfer points will be defined by \( l(x, y) \):

\[
l(x, y) = \frac{l_0}{2^{k(x,y)}}
\]

where \( l_0 = \frac{d}{N} \) is the base spacing between N-S and E-W tubes and \( k(x,y) \in \mathbb{N} \).

Similar to Ouyang et al. (2014), we assume that the value of \( k(x,y) \) does not have sharp variations with respect to both \( x \) and \( y \). In other words, we further assume that every change in the routes layout involves the convergence (or bifurcation) of no more than two routes.

### 4.2 Continuum approximation

Prior to introducing the costs derivation, we need to compute the expected rate per unit area that passengers perform a directional transfer at point \((x,y)\), \( D_{d-transf}(x,y) \).

\[
D_{d-transf}(x,y) = \frac{1}{2} \int_{\bar{x}=0}^{d} \int_{\bar{y}=0}^{d} \left[ \delta(\bar{x},y,x,\bar{y}) + \delta(x,\bar{y},\bar{x},y) \right] d\bar{x} d\bar{y}
\]

(4.1)

Finally, we can approximate the expected rate per unit area that passengers perform a spacing transfer at point \((x,y)\), \( D_{s-transf}(x,y) \).

\[
D_{s-transf}(x,y) \approx \frac{1}{2} \left( f^2_E(x,y) + f^2_W(x,y) \right) \left[ \frac{\partial}{\partial x} k(x,y) \right]^{-1} + \frac{1}{2} \left( f^2_N(x,y) + f^2_S(x,y) \right) \left[ \frac{\partial}{\partial y} k(x,y) \right]^{-1} + \frac{1}{2} \left( f^1_E(x,y) + f^1_W(x,y) \right) \left[ \frac{\partial}{\partial x} k(x,y) \right] + \frac{1}{2} \left( f^1_N(x,y) + f^1_S(x,y) \right) \left[ \frac{\partial}{\partial y} k(x,y) \right]
\]

(4.2)

where

\[
f^1_E = \frac{1}{2} \int_{x_1=0}^{x} \int_{y_2=x}^{d} \delta(x_1,y_2) d\bar{y} d\bar{x}
\]

(4.3a)

and

\[
f^1_S = \frac{1}{2} \int_{x_2=x}^{d} \int_{x_1=0}^{y_1=0} \delta(x_1,y_1) d\bar{y} d\bar{x}
\]

(4.3b)

Besides, we can compute the expected fluxes of onboard passengers passing through \((x,y)\) in all four directions per time-distance. Let \( F_E, F_W, F_N, F_S \) denote the eastbound, westbound,
northbound, southbound, respectively. We will only display $F_E$, since the other fluxes can be derived using symmetry.

$$F_E(x, y) = f_E^1(x, y) + f_E^2(x, y)$$

(4.3c)

See Appendix A for more details in the derivation of these terms. The $\frac{1}{2}$ factor pertains to the arbitrary choice of initial travel direction.

### 4.3 Agency investment

In order to be consistent with the previous chapter, we will need to compute similar costs for agency and users. Therefore, as in the previous part, we will compute the expected vehicular distance travelled by hour of operation, $Q$, and the vehicular time expended, $M$. These costs will be determined as cost per unit area per unit time, as follows. See Appendix A for details on the agency cost derivation (Result 13-14).

$$Q(x, y) \approx \frac{2}{H_l(x, y)} + \frac{2}{H_w(x, y)} + 2 \left[ \frac{1}{2} \frac{\partial k(x, y)}{\partial x} \right] + \frac{1}{2} \frac{\partial k(x, y)}{\partial y} + \frac{1}{3} (D_{\text{start}}(x, y) + D_{\text{end}}(x, y)) [2 * l(x, y) + 2 * w(x, y)]$$

(4.4)

$$M(x, y) = \frac{Q(x, y)}{v_c}$$

(4.5)

where

$$\frac{1}{v_c(x, y)} = \frac{1}{v} + \tau * \frac{D_{\text{start}}(x, y) + D_{\text{end}}(x, y) + e_T(x, y)}{Q(x, y)}$$

(4.6)

### 4.4 User costs

For the user costs, we will consider the same components as in Section 3.1.2 and 3.2.2. Therefore, we need to determine the waiting time, $W$, the expected in-vehicle travel time, $IVTT$, and the penalty associated with transfers. Those terms will be expressed in cost per unit area per unit time. See Appendix A for details on the user costs derivation (Result 15-16).
\[ W(x,y) = \frac{H}{2} [D_{\text{start}}(x,y) + D_{\text{s-transf}}(x,y) + D_{\text{d-transf}}(x,y)] \]  

(4.7)

\[ IVTT(x,y) = \] 

\[ \frac{1}{v} (F_E + F_W + F_N + F_S) + \frac{1}{4w} \left( \frac{(F_E + F_W) H w(x,y) l(x,y)/w(x,y)}{(F_N + F_S) H w(x,y)} \right) + \left( \tau + \frac{1}{l} \right) (D_{\text{start}}(x,y) + D_{\text{end}}(x,y)) \]  

(4.8)

\[ l(x,y) = (F_N + F_S) \] 

\[ e_T(x,y) = [D_{\text{s-transf}}(x,y) + D_{\text{d-transf}}(x,y)] \]  

(4.9)

4.5 Mathematical model

Similar to Section 3.1.3, agency’s monetary costs are converted to equivalent travel time. The optimization problem becomes the following:

\[ \text{Min} \quad \frac{1}{T_{\text{pax}}} \int_0^d \int_0^d Z(x,y) dx dy \]

s.t. \( l(x,y) = \frac{l_0}{Z(x,y)} \), for all \( (x,y) \)

\[ w(x,y) = \frac{w_0}{2k(x,y)} \], for all \( (x,y) \)

\[ l_0 = \frac{d}{N_1} \quad w_0 = \frac{d}{N_2} \]

\[ k(x,y) \in \{0,1,2 \ldots \} \], for all \( (x,y) \)

\[ N \in \{1,2 \ldots \left \lfloor \frac{d}{S} \right \rfloor \} \]

\[ H \geq 0 \]

The generalized cost, \( Z(x,y) = \frac{s_D}{\mu} Q(x,y) + \frac{s_{M}}{\mu} M(x,y) + W(x,y) + IVTT(x,y) + \theta e_T(x,y) \), is expressed per unit area per time. Note that \( M(x,y) \), when integrated over space, can be interpreted as the fleet size needed to run the system. Thus, the cost derived in this section is consistent with that in Section 3.1.2 and 3.2.2.
The design problem minimizes the expected generalized cost to the agency and to the users by determining optimal values of the decision variables $H$ and $l(x,y)$, for every neighborhood of $(x,y)$, or equivalently $H,l_0,w_0$ and $k(x,y)$ for every neighborhood of $(x,y)$. Recall that $k(x,y)$ is treated as a continuous function. Also note that the generalized cost function $Z(x,y)$ can be rewritten as follows.

$$Z(x,y) = \Phi(x,y) \cdot 2^{k(x,y)} + \Gamma(x,y) \cdot 2^{-k(x,y)} + \Pi(x,y) \cdot 4^{-k(x,y)} + \Psi(x,y) + \Lambda(x,y,\{k(x,y)\})_{y(x,y)}$$

where the functions of $(x,y)$ are

$$\Phi(x,y) = \left(\frac{S_0}{\mu} + \frac{S_M}{v\mu}\right) \left(\frac{2}{H}\left[l_0 + \frac{1}{w_0}\right]\right)$$

$$\Gamma(x,y) = \left(\frac{S_0}{\mu} + \frac{S_M}{v\mu}\right) \left(\frac{2}{3} [D_{start}(x,y) + D_{end}(x,y)]\right) (l_0 + w_0) + \tau (D_{start}(x,y) + D_{end}(x,y)) H \left[(F_N(x,y) + F_S(x,y)) l_0 + (F_E(x,y) + F_W(x,y)) w_0\right]$$

$$\Pi(x,y) = \frac{1}{3v} (D_{start}(x,y) + D_{end}(x,y)) H \left[(F_N(x,y) + F_S(x,y)) l_0^2 + (F_E(x,y) + F_W(x,y)) w_0^2\right]$$

$$\Psi(x,y) = \frac{S_M}{\mu} \tau \left(D_{start}(x,y) + D_{end}(x,y) + D_{d-transf}(x,y)\right) + \frac{H}{z} \left(D_{start}(x,y) + D_{d-transf}(x,y)\right) +$$

$$\frac{1}{v} (F_E + F_W + F_N + F_S) + \theta D_{d-transf}(x,y)$$

$$\Lambda(x,y,\{k(x,y)\})_{y(x,y)} =$$

$$\frac{1}{4v} \left(\frac{F_E(x,y) + F_W(x,y)}{w_0/l_0}\right) \left|\frac{\partial k(x,y)}{\partial x}\right| + \left(\frac{F_N(x,y) + F_S(x,y)}{l_0/w_0}\right) \left|\frac{\partial k(x,y)}{\partial y}\right| + \left(\frac{H}{z} + \theta + \frac{S_M}{\mu} \tau \right) D_{d-transf}(x,y) +$$

$$\left(\frac{S_0}{\mu} + \frac{S_M}{v\mu}\right) \left(\frac{1}{Hl_0} \left|\frac{\partial k(x,y)}{\partial x}\right| + \frac{1}{Hw_0} \left|\frac{\partial k(x,y)}{\partial y}\right|\right) 2^{k(x,y)}$$

$$Z'(x,y) = \Phi(x,y) \cdot 2^{k(x,y)} + \Gamma(x,y) \cdot 2^{-k(x,y)} + \Pi(x,y) \cdot 4^{-k(x,y)} + \Psi(x,y)$$

is obviously a monotonic function of $k(x,y)$, for any given $H,l_0,w_0$. Thus, there is a unique optimal solution $k^*$ that minimizes $Z'$ for each $(x,y)$. An upper bound of the optimal value of $Z(x,y)$ can be derived by plugging $k^*(x,y)$ in (4.10) and computing the partial derivative of $k^*$. The integration of the upper bound across the entire service region gives a feasible solution that is
also near-optimal if spacing transfers only contribute to a small part of the generalized cost (Ouyang et al., 2014). Therefore, we will implement an algorithm based on this near-optimal feasible solution to design the transit network.

**Step 1:** Initialize the triplet \( \langle H, l_0, w_0 \rangle \) and divide the service region into many small cells. Each cell represents a pair \((x, y)\) and the cells collectively cover the entire region.

**Step 2:** Compute the passenger rate functions \( D_{\text{start}}(x, y) \), \( D_{\text{end}}(x, y) \), \( D_{d-\text{transf}}(x, y) \) and \( D_{s-\text{transf}}(x, y) \), and the fluxes \( F_E(x, y) \), \( F_W(x, y) \), \( F_N(x, y) \) and \( F_S(x, y) \) for each cell.

**Step 3:** For each cell, find the \( k^*(x, y) \) that minimizes \( Z'(x, y) \) and compute the generalized cost by summing the local values \( Z(x, y) \) across all cells.

**Step 4:** Return to step 3 after perturbing the triplet \( \langle H, l_0, w_0 \rangle \) to \( \langle H + \Delta H, l_0 + \Delta l_0, w_0 + \Delta w_0 \rangle \) and use the gradient descent with respect to \( \langle H, l_0, w_0 \rangle \). If the gradient is within a specified tolerance, go to Step 5; otherwise, apply standard line search methods such as Newton’s method to find a new triplet \( \langle H, l_0, w_0 \rangle \) and go to Step 2.

**Step 5:** Round the optimal solution \{\( k^*(x, y), \forall (x, y) \}) to the nearest integer and implement the network design. Evaluate actual agency investment and user costs based on the final solution.

### 4.6 Numerical results

#### 4.6.1 System design

Similar to Chapter 3, the optimal grid transit network is designed for a square city with \( d = 10 \text{km} \) and for four distinct spatial demand distributions: (i) a mono-centric city; (ii) a twin-city; (iii) an asymmetric mono-centric city; and (iv) a commuter city. Demand distribution parameters are borrowed from Ouyang et al. (2014) and presented in Table 3.1. In order to reach the low-demand scenario, Equation (3.1) was divided by 5, for a base demand of 2,000 passengers/hr. See Section 3.1.4 for an illustration of the four marginal distributions. The parameters values used for these cases are displayed in Table 4.1. For comparison, most of them are similar to Section 3.1.4.

The algorithm presented in Section 4.5 was implemented into Matlab on a 3.3 GHz CPU and 8 GB memory desktop computer. For each case, the city was partitioned into cells of size 0.1
km by 0.1 km to run the algorithm. The results are displayed in Table 4.2 and illustrated in Figure 4.2. In all four cases, local tubes are part of the network tube layout and branching generally happens in regions with high demand density. Section 4.6.2 and 4.6.3 explore some key features of the network design with local tubes.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>s (km)</th>
<th>v (km/hr)</th>
<th>v_w (km/hr)</th>
<th>( \tau ) (s)</th>
<th>( S_M ) ($/veh-hr)</th>
<th>( S_Q ) ($/veh-km)</th>
<th>( \mu ) ($/hr)</th>
<th>( \Theta ) (hr)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Value</td>
<td>0.15</td>
<td>25</td>
<td>2</td>
<td>13</td>
<td>40</td>
<td>2</td>
<td>20</td>
<td>1/60</td>
</tr>
</tbody>
</table>

Table 4.1. Input parameters for the model

<table>
<thead>
<tr>
<th>Demand distribution</th>
<th>Mono-centric city</th>
<th>Twin cities</th>
<th>Asymmetric city</th>
<th>Commuter city</th>
</tr>
</thead>
<tbody>
<tr>
<td>( T_{pax} )</td>
<td>2,000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Decision variables</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H</td>
<td>0.24</td>
<td>0.21</td>
<td>0.23</td>
<td>0.23</td>
</tr>
<tr>
<td>( N_1 )</td>
<td>9</td>
<td>8</td>
<td>9</td>
<td>7</td>
</tr>
<tr>
<td>( N_2 )</td>
<td>9</td>
<td>8</td>
<td>9</td>
<td>7</td>
</tr>
<tr>
<td>Agency metrics</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \pi Q )</td>
<td>0.28</td>
<td>0.27</td>
<td>0.27</td>
<td>0.27</td>
</tr>
<tr>
<td>( \pi M )</td>
<td>0.24</td>
<td>0.23</td>
<td>0.23</td>
<td>0.23</td>
</tr>
<tr>
<td>User metrics</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>W</td>
<td>0.25</td>
<td>0.22</td>
<td>0.24</td>
<td>0.26</td>
</tr>
<tr>
<td>IVTT</td>
<td>0.18</td>
<td>0.11</td>
<td>0.20</td>
<td>0.42</td>
</tr>
<tr>
<td>( \theta )</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
</tr>
<tr>
<td>Generalized cost</td>
<td>Z</td>
<td>0.97</td>
<td>0.85</td>
<td>0.96</td>
</tr>
</tbody>
</table>

Table 4.2. Result summary for several heterogeneous demand distributions
4.6.2 Comparison with grid structure without branching

In this section, we compare the optimal designs found in Section 4.6.1 with the same design without branching. The grid structure without local tubes serves as a basis for comparison. Thus, allowing local tubes within the system significantly reduces the total system cost (-18%, -30%, -21%, -25% for the four cases). The main cost reduction is observed for the agency costs. Total vehicular distance travelled and total vehicular time expended are reduced by 32% and 36%, respectively, in the case of the mono-centric city. They are reduced by 50% and 53%, respectively, for the twin cities, by 39% and 38%, respectively, for the asymmetric city, and by 61% and 62% for the commuter city. Branching reduces the vehicular distance since it reduced the lateral movement of the bus. With branching, the width of the local tubes
is reduced so buses travel less to pick up and drop off passengers. To a lesser extent, in-vehicle travel time is also reduced in all four cases (-8%, -8%, -8%, -5%, respectively). However, one can note that the average waiting time and the expected number of transfer increase when branching is allowed. Branching indeed forces passenger to experience spacing transfers. Since waiting time is partly related to the average number of transfer, if the latter increases then the former will increase as well. In the mono-centric, twin cities, and asymmetric cities cases, the increase is limited (5%, 4%, 3%), but in the case of the commuter city, average waiting time increases by 11%. This is due to the fact that commuters will experience more transfers (+19%), in particular spacing transfers, since branching occurs in the region of the trip origins and in the region of trip destinations.

We can then conclude that branching significantly reduces the agency metrics (vehicular time travelled and vehicular distance expended) mainly by reducing the lateral distance that buses travel to provide door-to-door service, and then reduces the total system cost.

<table>
<thead>
<tr>
<th>Demand distribution</th>
<th>Mono-centric city</th>
<th>Twin cities</th>
<th>Asymmetric city</th>
<th>Commuter city</th>
</tr>
</thead>
<tbody>
<tr>
<td>T_{pax}</td>
<td>2,000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Decision variables</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H</td>
<td>0.24</td>
<td>0.21</td>
<td>0.23</td>
<td>0.23</td>
</tr>
<tr>
<td>(N_1)</td>
<td>9</td>
<td>8</td>
<td>9</td>
<td>7</td>
</tr>
<tr>
<td>(N_2)</td>
<td>9</td>
<td>8</td>
<td>9</td>
<td>7</td>
</tr>
<tr>
<td><strong>Agency metrics</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\pi_{QQ})</td>
<td>-32%</td>
<td>-50%</td>
<td>-39%</td>
<td>-61%</td>
</tr>
<tr>
<td>(\pi_{MM})</td>
<td>-36%</td>
<td>-53%</td>
<td>-38%</td>
<td>-62%</td>
</tr>
<tr>
<td><strong>User metrics</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IVTT</td>
<td>-8%</td>
<td>-8%</td>
<td>-8%</td>
<td>-5%</td>
</tr>
<tr>
<td>(\theta_{e})</td>
<td>0%</td>
<td>6%</td>
<td>6%</td>
<td>19%</td>
</tr>
<tr>
<td><strong>Generalized cost</strong></td>
<td>(-18%)</td>
<td>(-30%)</td>
<td>(-21%)</td>
<td>(-25%)</td>
</tr>
</tbody>
</table>

Table 4.3. Percentage change in cost metrics between grid structure without and with branching

We now compare the optimal design with and without branching. We test both systems, with heterogeneous and homogeneous tube layout, with the demand distributions presented in the previous cases. Table 4.4 displays cost metrics and optimal decision variable values for both systems in the four above-mentioned densities. Model A pertains the optimal grid system with branching, while model B pertains the optimal grid system without branching.
Note that in each case branching produces lower generalized costs. In the four cases, total generalized cost is reduced by 3%, 8%, 6%, and 2%, respectively. Those values are consistent with the results presented in Nourbakhsh and Ouyang (2012). As expected, there are some noticeable differences in the optimal N, for each of the four cases. Since k varies between 0 and 2, the optimal N in the grid system without branching can be seen as an average value of the tubes density within the system with branching. One can notice that the optimal decision variables for the grid system without branching (homogeneous system) are close to those found in Section 3.1 and they are identic for each of the four cases. A potential reason for the difference in optimal solutions is that extra-distance is not taken into consideration in this chapter. Due to rounding in the table, some cost metrics appear to be equal while they are only very close to each other. Thus, for each of the four cases, all cost metrics are reduced, except for the fourth case, whereby the users’ cost increases.

Overall, the total generalized costs are lower for the heterogeneous layouts than for homogeneous layouts, which highlights the value of adding local tubes in high-demand neighborhoods.

<table>
<thead>
<tr>
<th>Demand distribution</th>
<th>Mono-centric city</th>
<th>Twin cities</th>
<th>Asymmetric city</th>
<th>Commuter city</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>A</td>
<td>B</td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>Decision variables</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H</td>
<td>0.24</td>
<td>0.25</td>
<td>0.21</td>
<td>0.25</td>
</tr>
<tr>
<td>N_1</td>
<td>9</td>
<td>19</td>
<td>8</td>
<td>19</td>
</tr>
<tr>
<td>N_2</td>
<td>9</td>
<td>19</td>
<td>8</td>
<td>19</td>
</tr>
<tr>
<td>Agency metrics</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>\pi_Q</td>
<td>0.28</td>
<td>0.29</td>
<td>0.27</td>
<td>0.29</td>
</tr>
<tr>
<td>\pi_M</td>
<td>0.24</td>
<td>0.25</td>
<td>0.23</td>
<td>0.25</td>
</tr>
<tr>
<td>User metrics</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>W</td>
<td>0.25</td>
<td>0.25</td>
<td>0.22</td>
<td>0.26</td>
</tr>
<tr>
<td>IVTT</td>
<td>0.18</td>
<td>0.18</td>
<td>0.11</td>
<td>0.11</td>
</tr>
<tr>
<td>\theta_e</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
</tr>
<tr>
<td>Generalized cost</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>\text{improvement}</td>
<td>-3%</td>
<td>-8%</td>
<td>-6%</td>
<td>-2%</td>
</tr>
</tbody>
</table>

Table 4.4. Optimal design summary for several heterogeneous demand distributions
4.6.3 Sensitivity analysis

The sensitivity analysis is performed on the total number of passengers in the system for a mono-centric city. Table 4.4 presents the cost metrics in each case and Figure 4.3 displays the optimal design for a wide range of values (500 pax/hr to 10,000 pax/hr).

We only increase the total number of passenger, that is, we multiply formula (3.1) by a factor accordingly to obtain the expected number of passengers. Thus, the ratio of the maximum local trip demand over the minimum local trip demand remains constant. Basically we shift up the shapes of the demand in Figure 3.2. The increasing number of passengers does not affect the branching since the value of \( k \) is always less than 2 \( (k \in \{0,1,2\}) \).

<table>
<thead>
<tr>
<th>Demand distribution</th>
<th>Mono-centric city</th>
</tr>
</thead>
<tbody>
<tr>
<td>( T_{pax} )</td>
<td>500</td>
</tr>
<tr>
<td>Decision variables</td>
<td></td>
</tr>
<tr>
<td>( H )</td>
<td>0.37</td>
</tr>
<tr>
<td>( N_1 )</td>
<td>6</td>
</tr>
<tr>
<td>( N_2 )</td>
<td>6</td>
</tr>
<tr>
<td>Agency metrics</td>
<td></td>
</tr>
<tr>
<td>( \pi_Q )</td>
<td>0.45</td>
</tr>
<tr>
<td>( \pi_M )</td>
<td>0.36</td>
</tr>
<tr>
<td>User metrics</td>
<td></td>
</tr>
<tr>
<td>( W )</td>
<td>0.39</td>
</tr>
<tr>
<td>( \text{IVTT} )</td>
<td>0.18</td>
</tr>
<tr>
<td>( \theta_e )</td>
<td>0.02</td>
</tr>
<tr>
<td>Generalized cost</td>
<td></td>
</tr>
<tr>
<td>( Z )</td>
<td>1.40</td>
</tr>
</tbody>
</table>

Table 4.5. Result summary for different demand levels
Figure 4.3. Sensitivity analysis for different demand levels

(a) $T_{pax} = 500 \text{ pax/hr}$

(b) $T_{pax} = 1,000 \text{ pax/hr}$

(c) $T_{pax} = 2,000 \text{ pax/hr}$

(d) $T_{pax} = 5,000 \text{ pax/hr}$

(e) $T_{pax} = 10,000 \text{ pax/hr}$
4.7 Conclusion

This chapter presented a continuum approach for designing heterogeneous tube layouts in response to a spatially heterogeneous passenger demand. The framework used in this chapter was provided by Roundy (1985) as the power-of-two scheme. The grid network is composed of a set of main bus tubes, with larger width, along with a grid of local bus tubes. Main tubes can split locally into smaller tubes, so-called local tubes, to better serve regions with higher trip demand. The method used in this chapter allows main tubes to be seamlessly aligned with local tubes. Closed-form expressions were formulated to derive local agency investment and users’ costs. A constrained optimization model was formulated and solved using a localized optimization model.

Numerical results were presented using several travel demand distributions (mono-centric city, twin cities, asymmetric city, commuter city) and showed that the heterogeneous network allows to reduce all generalized costs for the cases studied in this thesis, in comparison to the homogeneous grid network. The major impact of branching is the reduction of lateral distance, since the local width of the bus tubes is reduced. Next, a sensitivity analysis is performed with respect to the passenger demand level. It is found out that the branching does not depend on the demand level but rather on the ratio of the maximum local demand over the minimum local demand.

Finally, the proposed network configuration displays benefits to flexible transit system, since better service can be provided in any region with higher demand level, and not only within the central region as in Chapter 3. Given the results in Chapter 3, this feature could be used within a hybrid network structure. However, this approach comes with some limitations. Because of branching, passengers experience more transfers in average. Since transfers are to be avoided, coordinated bus schedule have to be implemented to reduce the waiting time at transfer points.
CHAPTER 5
FUTURE RESEARCH

This chapter provides potential extension to this thesis along with some future research topics related to public transit design. Ideas are organized as follow. Section 5.1 describes future extensions to the systems designed in this thesis. Section 5.2 presents an alternative transit system design. Section 5.3 takes other modes of flexible transit into account. Section 5.4 provides some ideas on transit system design that are not related to the previous sections.

5.1 Flexible hybrid system with local tubes

In Chapter 3, we proved that hybrid transit structure was beneficial over grid system under heterogeneous travel demand since it only provides double coverage in the central region, which is usually the region with the highest demand rate. Peripheral region is expected to have lower demand rate and therefore tube density is lower in that region. However, double coverage is allowed only in the central region, but we want to be able to provide higher service wherever demand density is higher. Chapter 4 provides guidelines to design a grid system with local branching. The transit system is designed such that tube density varies with space to provide higher local service in zones with higher demand. Hence, higher coverage would be provided, regardless whether the higher trip demand rate is in the central region or not.

Hence the next step would be to design local branching within a hybrid transit system; see Figure 5.1. This would provide even more flexibility to the system since we would have more control on the local service level. Several features could be discussed regarding this potential transit system: (i) the location of the central region could be a decision variable to fit any heterogeneous demand density, (ii) whether branching affects the design of the peripheral region or not. (i) would consist of offsetting the central region location such that the double-
coverage service region is not necessarily at the center of the service region. This would obviously be beneficial for demand distribution such as asymmetric city or commuter city. (ii) would allow branching to happen also in some zones of the peripheral region, in case moderate demand regions overlap with the peripheral region.

To put it in a nutshell, the ultimate transit system would have the following features: (i) flexible-route tubes, (ii) hybrid structure, and (iii) local branching. Formulations provided in this thesis could serve as a basis to design such a system, and as benchmark to assess the level of improvement provided by each feature, with respect to fixed-route system, grid structure, and system with no local branching, respectively.

![Figure 5.1. Hybrid structure with local tubes](image)

Most of the cost metrics can be adapted from the mathematical derivations provided in this thesis. However, a complexity in the in-vehicle travel time that needs to be addressed is how to account for the fact that passengers within the peripheral region have to travel through the central region. Cost metrics within the central region are the same as for the grid structure presented in this thesis.
5.2 Semi-flexible transit system

Since Quadrifoglio and Li (2009), Li and Quadrifoglio (2010) and Nourbakhsh and Ouyang (2012) demonstrated the advantages of flexible transit over fixed-route transit under low-to-moderate demand level, it could be beneficial to design a transit system that features both modes. Under heterogeneous trip distribution, demand level may vary significantly over space in such a way that the service region includes both low-to-moderate and high demand levels. Hence transit design would consist of a fixed-route system in regions with higher demand rate and a flexible-route system in lower demand region. Some forms of this system were implemented in the U.S., for instance in Hampton, VA, Fort Worth, TX, or Ottumwa, IA.

Aldaihani et al. (2004) designed a semi-flexible grid transit system consisting of a fixed-route system associated with on-demand vehicles that service predetermined zones; see Figure 5.2. Since we proved in this thesis that hybrid structure is beneficial over grid structure for flexible transit, we could implement a hybrid transit system with fixed routes in the higher density central region and flexible routes in the peripheral region; see Figure 5.3. A potential future research topic is to adapt the work done in this thesis with Smith (2014) to design a hybrid transit system under heterogeneous demand. However, one issue addressed by Smith (2014) would be how to ensure a constant headway for the whole system whereas bus trip lengths in the peripheral region are stochastic.

![Figure 5.2. Semi-flexible grid transit network (Aldaihani et al., 2004)](image-url)
Also, several transportation modes can be experimented for the fixed-route service region. Daganzo (2010a) and Smith (2014) provide a comparative analysis on the different transportation modes (bus, Bus Rapid Transit, Metro) to design a fixed-route hybrid transit system. Given that buses will be operating within the flexible-route service zone, we could allow a different mode within the fixed-route central region.

5.3 Flexible modes

As mentioned in Section 2.1, there are several flexible service modes. This thesis only designed transit systems for a zone route mode, which is statistically the most common method (Transportation Research Board, 2004). One could also investigate the benefits of using different modes, such as route deviation or point deviation, which are hybrid modes including some aspects of fixed-route service and zone route service, in terms of passenger density. Daganzo (1984) designed a checkpoint dial-a-ride system while Zhao and Dessouki (2008) designed a route deviation service. However, those transit systems only consist of a single bus tube. A potential future research would be to extend them to a global network, as presented in this thesis.
5.4 Other

5.4.1 Service reliability

Since flexible transit is supposed to provide more flexibility to the users, but also is highly dependent on the demand distribution, since its routes are not predetermined and route lengths are stochastic. As a result, service level can vary significantly with the demand density. Service reliability then becomes a major issue. Then we could use a service reliability-based model (Lo et al., 2013), to cover the stochastic trip demand up to a certain specified reliability, or a headway-based approach (Daganzo, 2009), to maintain bus schedule by adding timed checkpoints for instance. One could adjust the parameters in the given model to take service reliability into account but it would certainly add more complexity to the model.

Since flexible transit usually operates in low-density regions, people are likely to use individual automobiles to commute, since public transportation may not provide convenient service in those regions. Thus fare collection system could be investigated and incorporated within the transit design model. It could have a major impact on the number of commuters using the system and then on the viability of the transit system.

Also critical occupancy can be another issue since flexible transit usually consists of small buses or vans. Thus, vehicle capacity is very limited. For instance, a solution would be to incorporate a capacity constraint in the network design model (Imam, 1998).

5.4.2 Environmental impact

Additional environment-related features such as energy consumption, gas emissions and pollution could also be taken into consideration, since environmental impact of transportation systems is now becoming a major issue (Greene, 2006). However, environmental aspects are rarely considered when designing transit networks. Although public transportation has the potential to reduce the ecological footprint of transportation systems by aggregating passenger trips, transit vehicle characteristics and operations tend to balance that statement. For instance, transit vehicles are larger and therefore released larger gas emissions during operations, stop and get going way more often than individual vehicles, to pick up and drop off passengers. Some literature accounts for those concerns (Saka, 2003, Dessouky et al.,
2003, Diana et al., 2007, Beltran et al., 2009, Griswold et al., 2013). Saka (2003) proved that the optimal spacing obtained by taking environmental aspects into account is much larger than the actual bus spacing in urban areas. Dessouky et al. (2003) included environmental metrics to design a flexible transit system. They found out that the environmental impact of transit systems can be significantly reduced while slightly increasing the other costs. Diana et al. (2007) investigated the environmental impact of several transportation modes including flexible-route transit for various road networks and demand densities. Beltran et al. (2009) provide a transit design model including a constraint on the availability of green vehicles. Griswold et al. (2013) designed a transit network and incorporated a constraint on the total gas emissions and environmental costs. The increasing concern on the greenhouse gas emissions combined with the significant environmental impact of transportation gives credit to those models. Thus the previously described modeling frameworks are of high interest.

5.4.3 Spatially and temporally heterogeneous demand

As mentioned in Section 2.4, demand can be spatially and temporally heterogeneous. Newell (1971), Hurdle (1973), Clarens and Hurdle (1975), and Chang and Schonfeld (1991) provided the basis on temporally heterogeneous demand. Thus, it would be interesting to design a transit system under spatially heterogeneous demand, considering several time periods and investigate the sensitivity of the optimal design and operation when the demand can change with respect to both space and time.
CHAPTER 6
CONCLUSIONS

This thesis investigated different features in the design and operation of a flexible-route transit system under heterogeneous trip demand. Prior to the analysis, Chapter 2 presented related literature on (i) flexible transit characteristics, (ii) existing flexible transit systems (iii) transit network design, (iv) flexible-route system design, (v) continuum approximation methods, (vi) spatially and temporally heterogeneous passenger demand, and (vii) solution methods for transit design problems.

Chapter 3 provided the formulations to the total system cost including the agency’s investment and users’ costs for both a grid and a hybrid transit structure (Daganzo, 2010a) with spatially heterogeneous demand. Chapter 3 also presented numerical results for the design and operating frequency in mono-centric city, twin cities, asymmetric city, and commuter city, with a low-to-moderate demand level (2,000 pax/hr). Results for the grid system showed that the grid system design and operating frequency do not depend on the demand distribution but rather on the total demand. Hybrid system provides more flexibility to the system, by reducing double coverage to the central region, where demand density is expected to be higher. For the results presented, hybrid structure is preferred for each demand distribution, since it allows a cost decrease up to 12% with respect to the grid system. The total system cost decreases more when the demand becomes more concentrated (mono-centric city, twin cities, asymmetric city).

Chapter 4 presented a method to improve grid system, by allowing the insertion of local tubes within region with higher passenger demand. Roundy (1985) and Ouyang et al. (2014) provide a detailed framework for the method used in this thesis. Similar to Chapter 3, agency’s investment and users’s costs were derived to build a model that could allow a heterogeneous tube layout in the transit system. Closed-form approximate expressions were formulated to express the total system cost and a localized optimization algorithm was used.
to obtain global optimal network design. Numerical results are provided for mono-centric city, twin cities, asymmetric city, and commuter city cases. They show that heterogeneous tube layout networks produce lower generalized costs than do homogeneous grids. Chapter 4 also performed a sensitivity analysis and it was found out that branching does not depend on the total passenger demand.

Finally, Chapter 5 presented some future research ideas to follow up the work of the present thesis: (i) hybrid structure design with local branching, (ii) semi-flexible transit network including both flexible-route and fixed-route transit features, (iii) other flexible modes implementation, (iv) service reliability aspects, (v) environmental impact, and (vi) spatially and temporally heterogeneous passenger demand.
REFERENCES


City of Winnipeg Transit Department, 1997. Evaluation of DART (Dial-a-Ride Transit). Memorandum for submission to the Committee on Works and Operations.


New York, NY.


Sarzynski, A., Wolman, H.L., Galster, G., Hanson, R., 2006. Testing the conventional wisdom about land use and traffic congestion: The more we sprawl, the less we move?


Vaughan, R., 1986. Optimum polar networks for an urban bus system with a many-to-many


APPENDIX A

PROOFS

Result 1. The hourly total vehicle distance is given by (3.3)

\[
Q \approx \frac{1}{H} \sum_{k=1}^{2N} \left[ 2 \left( d + \frac{1}{3} (H) \left( \frac{d}{N} \right) \left( \frac{1}{2} \int (D_{\text{start}} + D_{\text{end}}) \right) \right) + \sum_{i=1}^{\infty} (i - 1)(s/2) \left( \sum_{(x,y) \in T_k} P_{x,y}(i) \right) \right]
\]

where

\[
P_{x,y}(i) = \left( \frac{(D_{\text{start}}(x,y) + D_{\text{end}}(x,y))}{2} \right)^i \left( e^{-[D_{\text{start}}(x,y) + D_{\text{end}}(x,y)]H/2} \right) \frac{i!}{i!}
\]

Proof. The total vehicle distance per hour is given by the product of the number of bus tubes, the expected travel distance of one round trip, and the inverse of the bus headway. The expected travel distance per bus round trip is divided in three parts: longitudinal distance, lateral distance, and extra detour distance.

The longitudinal distance is obviously the square side length and the lateral distance is

\[
\frac{1}{3} \ast \left( \frac{d}{N} \right) \ast \#\text{pass} = \frac{1}{3} \ast \left( \frac{d}{N} \right) \ast H \ast \frac{1}{2} \ast \left( f_{T_k} (D_{\text{start}} + D_{\text{end}}) \right).
\]

Note that a bus has to travel this lateral distance twice for each passenger (pick-up and drop-off).

Because of the street spacing s, the bus must travel extra distance if it needs to visit more than one passenger before advancing through one street block in the longitudinal direction. The expected longitudinal distance to visit each additional passenger is \(s/2\). The number of passengers in the tube block area follows a Poisson distribution with mean \(\frac{\lambda_j H}{2}\), considering double coverage, where \(\lambda_j\) is the number of passengers per time in the tube block. Therefore, the probability of having \(i\) passengers in a tube block is \(P_j(i) = \left( \frac{\lambda_j H}{2} \right)^i e^{-\frac{\lambda_j H}{2}}/i!\). For more
accuracy we divide each tube block into small cells, where we assume that the demand is homogeneous.

**Result 2.** The expected number of transfers is given by

\[
\eta_T = 1 - \frac{\lambda^{(0)}}{T_{pax}}
\]

where \( \lambda^{(0)} = \sum_{i=1}^{2N} \left[ \mathbb{T}_{T_i} \mathbb{T}_{T_i} \delta \right] - \sum_{i=1}^{N^2} \left[ \mathbb{T}_{C_i} \mathbb{T}_{C_i} \delta \right] \)

**Proof.** One trip may include 0 or 1 transfer, depending on the location of the origin and destination. Thus, the expected number of transfers is equal to the conditional probability of having one transfer. The conditional probability of having one transfer is \( \lambda^{(1)}/T_{pax} \). The number of passengers experiencing exactly one transfer is given by \( \lambda^{(1)} = T_{pax} - \lambda^{(0)} \), where \( \lambda^{(0)} \) is the number of passengers experiencing zero transfer, that is, if both the origin and destinations are in the same tube. We need to subtract the number of passengers whom origin and destination are in the same square, since it is counted twice in that case.

**Result 3.** The average bus speed satisfies (3.5)

\[
\frac{1}{v_c} \approx \frac{1}{v} + (\tau)(T_{pax})(\frac{2 + \eta_T}{Q})
\]

**Proof.** We assume that during each stop, the bus pick up or drop off exactly one passenger (except at transfer points). The total number of stop is the double of the total number of passengers \( T_{pax} \) plus the number of passengers experiencing a transfer. The bus needs to overcome time over distance \( 1/v \) and stops for \( \tau \) time per passenger during a round trip with average distance \( QH/2N \). On average, each bus serves \( T_{pax}/2N \) passengers per time.

Hence, \( \frac{QH/2N}{v_c} = \frac{QH/2N}{v} + (\tau)(T_{pax})(H)(2 + \eta_T)/2N \).
**Result 4.** The expected waiting time per passenger is given by (3.6)

\[ W = \frac{H}{2} (1 + e) \]

**Proof.** The waiting time is the sum of the time spent at the origin waiting for the bus, and at the transfer station. Therefore, the expected waiting time is obtained by multiplying the expected number of transfers by the waiting time per transfer and adding it to the waiting time at the origin. Each waiting time (at the transfer point, at the origin) is approximately \(\frac{H}{2}\). □

**Result 5.** The expected in-vehicle travel distance, \(E\), and travel time, \(IVTT\), per passenger are

\[ E \approx \frac{QH}{6N} \text{ and } IVTT = \frac{E}{v_c}, \text{ respectively.} \]

**Proof.** The expected in-vehicle travel distance is obtained by multiplying the ratio of the expected total distance over the expected longitudinal distance by the expected in-vehicle longitudinal distance per passenger trip.

The expected longitudinal distance per passenger trip is \(\frac{d}{3} + \frac{d}{3} = \frac{2d}{3}\), while the ratio of the expected total distance over the expected longitudinal distance during one bus round trip is \(\frac{QH}{2d}\).

Therefore, \(E \approx \frac{2d}{3} \times \frac{QH}{2d} = \frac{QH}{6N}\).

Then the total expected travel time for each passenger trip is given by \(IVTT = \frac{E}{v_c}\). □

**Result 6.** The vehicle occupancy is given by (3.9)

\[ O = \frac{HT_{pax}}{2N} \]
Proof. There are $N$ equatorial bus tubes that cross between the northern and southern halves. On average there are $\frac{HT_{pax}}{4}$ passengers traveling along each of these tubes. Similar to Daganzo (2010a), we multiply by a safety factor of 2. This results in the critical occupancy. $\square$

Result 7. The hourly total vehicle distance and fleet size are given by (3.17)

$$Q = \frac{1}{h} \left[ 2 \left( 2NaD + \frac{2aD}{3} \lambda_{c-p}H + \frac{aD}{3} (\lambda_{c-p} + \lambda_{p-c})H + 2N(1 - \alpha)D + l_p(\lambda_{c-p} + \lambda_{p-c})H + 2l_p\lambda_{p-p}H \right) + L_{extra} \right]$$

$$M = \frac{Q}{v_c}$$

where $P_j\{i\}$ is given in (3.1) and $l_p$ is given by

If $D^2(1 + \alpha)^2 \left[ \frac{[\lambda_{c-p} + \lambda_{p-c} + \lambda_{p-p}]}{(1-\alpha)^2D^2} \right] H \geq 2N^2$, 

$$l_p = \frac{(1 + \alpha)D}{6N} + \frac{2N^3}{H^2} \left[ \frac{\lambda_{c-p} + \lambda_{p-c} + \lambda_{p-p}}{(1-\alpha)^2D^2} \right] - \frac{2N^5}{H^3} \left[ \frac{\lambda_{c-p} + \lambda_{p-c} + \lambda_{p-p}}{(1-\alpha)^2D^2} \right]$$

otherwise,

$$l_p = \frac{N}{D(1 + \alpha) \left[ \frac{\lambda_{c-p} + \lambda_{p-c} + \lambda_{p-p}}{(1-\alpha)^2D^2} \right] H}$$

Also, $L_{extra}$ is given by

$$L_{extra} = \sum_{k=1}^{2N} \left[ \sum_{i=1}^{\infty} (i - 1) \left( \sum_{(x,y) \in T_k} P_{x,y}(i) \right) \right]$$
Proof. The total vehicle distance per hour is given by the sum of the vehicle distance within the central region and within the peripheral region: \( Q = Q_c + Q_p \). Each term includes three parts, as in Result 1: longitudinal distance, lateral distance, and extra-distance.

The round-trip longitudinal distance is \( 2\alpha D \) in the central square and \( 2(1 - \alpha)D \) in the peripheral region, for each tube. The lateral distance is computed from the number of passengers served during one headway. The average lateral distance in the central region is to pick up one passenger is \( \frac{\alpha D}{3N} \) and the lateral distance in the peripheral region to pick up one passenger is \( l_p \). Since buses travel lateral distance to pick up and to drop off passengers, for each passenger, depending on his origin and destination, we can compute the lateral distance traveled accordingly (one for the pick-up and one for the drop-off).

\[
l_p \text{ is derived in Nourbakhsh and Ouyang (2012): } l_p = \begin{cases} \frac{w}{3} + \frac{b^2}{w} - \frac{b^3}{3w^2} & \text{if } b < w \\ \frac{b}{w} & \text{if } b \geq w \end{cases}
\]

where \( w \) is the local tube width, \( w = \frac{(1+\alpha)D}{2N} \), and \( b \) is the lateral offset of the tube between two consecutive passengers, \( b = \frac{\lambda_c(1) + \lambda_{p-c} + \lambda_{c-p} + \lambda_p(1) + \lambda_{p-p}}{D(1+\alpha)} \frac{N}{(1-\alpha)^2D^2} \), if we assume that demand is approximately uniform in the peripheral region. Since we are considering a low demand region, and the demand is even lower in the peripheral region, we can make that assumption. This yields to the formula displayed above.

The extra distance \( L_{\text{extra}} \) is computed as in Result 1. We divide the service region within a set of small cells. Then we consider each small cell (0.1km×0.1km) where we compute the travel demand. Then we compute the extra-distance as if the demand was homogeneous in the whole service region. Finally, we scale it down to the size of the cell to obtain the extra-distance of one particular cell. By adding all extra-distances we can compute the total extra-distance experienced by the fleet within the service region. □

Result 8. The expected number of transfers is given by (3.20)

\[
e_T = \frac{\lambda_{c-c}^{(1)} + \lambda_{p-c}^{(1)} + \lambda_{c-p}^{(1)} + \lambda_{p-p}^{(1)}}{T_{\text{pax}}} + 2\frac{\lambda_{p-p}^{(2)}}{T_{\text{pax}}}
\]
**Proof.** One trip may include 0, 1 or 2 transfers. Similar to Result 2, we compute the mean of the probability of having k transfers, \( e_T = \sum_{k=0}^{2} k P(X = k) \). □

**Result 9.** The expected in-vehicle travel distance and expected in-vehicle travel time are given by (3.21) and (3.22)

\[
E = \frac{Q_c H / 2N}{2aD} \left[ \frac{\lambda_{c-c}}{T_{pax}} \frac{2}{3} aD + \frac{\lambda_{p-c} + \lambda_{c-p}}{T_{pax}} \frac{5}{6} aD + \frac{\lambda_{p-p}}{T_{pax}} \frac{11}{12} aD \right] + \frac{Q_p H / 2N}{2D (1-\alpha)} \left[ 2 - 3\alpha + \alpha^2 \right] \frac{D}{3} \]  

\[T = \frac{E}{v_c}\]

**Proof.** Similar to Result 5, the expected in-vehicle travel distance is obtained by multiplying the ratio of the expected total distance over the expected longitudinal distance by the expected in-vehicle longitudinal distance per passenger trip: \( E = \rho_c E(R_c) + \rho_p E(R_p) \). \( E(R_p) \) is derived in Result 5 in Nourbakhsh and Ouyang (2012). \( E(R_c) \) is the weighted longitudinal distance traveled for passengers travelling from Central to Central, between Central and Peripheral, and from Peripheral to Peripheral. The expected longitudinal distances for these cases are \( \frac{2}{3} aD \), \( \frac{5}{6} aD \), and \( \frac{11}{12} aD \), respectively. The corresponding probabilities are \( \frac{\lambda_{c-c}}{T_{pax}} \), \( \frac{\lambda_{p-c} + \lambda_{c-p}}{T_{pax}} \), and \( \frac{\lambda_{p-p}}{T_{pax}} \). Then,

\[
E(R_c) = \frac{\lambda_{c-c}}{T_{pax}} \frac{2}{3} aD + \frac{\lambda_{p-c} + \lambda_{c-p}}{T_{pax}} \frac{5}{6} aD + \frac{\lambda_{p-p}}{T_{pax}} \frac{11}{12} aD
\]

\[
E(R_p) = (2 - 3\alpha + \alpha^2) \frac{D}{3}
\]

Similar to Result 5, \( \rho_c = \frac{Q_c H / 2N}{2aD} \) and \( \rho_p = \frac{Q_p H / 2N}{2D (1-\alpha)} \). This yields the formula for \( E \), and the total expected travel time for each passenger trip is given by \( T = \frac{E}{v_c} \). □
Result 10. The vehicle critical occupancy is given by (3.23)

\[ O = \frac{H}{N} \max \left\{ \frac{\max(\lambda_{p-c} \lambda_{c-p}) + \lambda_{p-p} + \lambda_{p-p}}{2}, \frac{\lambda_{c-p}}{4}, \frac{\lambda_{c-p} + \lambda_{p-c} + \lambda_{p-p}}{2} - \frac{\lambda_{p-p}}{8} \right\} \]

Proof. The formula is essentially the same as Result 8 in Smith (2014).

Result 11. The expected northbound flux per time-distance of onboard passengers passing through \((x,y)\) is given by (4.3)

\[ F_E(x, y) = f_E^1(x, y) + f_E^2(x, y) \]

Proof. The same as in Ouyang et al. (2014). Only the eastbound rate is developed but westbound, northbound and southbound fluxes can be obtained using symmetry by swapping the variables accordingly. Recall that

\[
\begin{align*}
    f_E^1 &= \frac{1}{2} \int_{x_1=0}^{x} \int_{x_2=x}^{d} \delta(x_1, y, x_2, y_2) dy_2 dx_2 \bigg| dx_1 \\
    f_E^2 &= \frac{1}{2} \int_{x_2=x}^{d} \int_{y_1=0}^{x} \delta(x_1, y_1, x_2, y) dy_1 dx_1 \bigg| dx_2.
\end{align*}
\]

Result 12. The approximate rates per unit area per time that passengers perform directional transfers and spacing transfers at \((x, y)\) are, respectively,

\[
\begin{align*}
    D_{d-transf}(x, y) &= \frac{1}{2} \int_{\bar{x}=0}^{d} \int_{\bar{y}=0}^{d} \left[ \delta(\bar{x}, y, x, \bar{y}) + \delta(x, \bar{y}, \bar{x}, y) \right] d\bar{x} d\bar{y} \\
    D_{s-transf}(x, y) &\approx \frac{1}{2} \left( f_E^2(x, y) + f_W^1(x, y) \right) \left[ \frac{\partial}{\partial x} k(x, y) \right]^- + \frac{1}{2} \left( f_E^1(x, y) + f_W^2(x, y) \right) \left[ \frac{\partial}{\partial x} k(x, y) \right]^+ \\
    &+ \frac{1}{2} \left( f_W^2(x, y) + f_S^1(x, y) \right) \left[ \frac{\partial}{\partial y} k(x, y) \right]^- + \frac{1}{2} \left( f_W^1(x, y) + f_S^2(x, y) \right) \left[ \frac{\partial}{\partial y} k(x, y) \right]^+.
\end{align*}
\]

Proof. The same as in Ouyang et al. (2014).
Result 13. The expected vehicular distance travelled by hour of operation is given by (4.4)

\[ Q(x, y) \approx \frac{2}{H_l(x, y)} + \frac{2}{H_w(x, y)} + 2 \left[ \frac{1}{2} \frac{1}{H_l(x, y)} \left| \frac{\partial k(x, y)}{\partial x} \right| + \frac{1}{2} \frac{1}{H_w(x, y)} \left| \frac{\partial k(x, y)}{\partial y} \right| \right] + \frac{2}{3} w(x, y) \left[ \frac{D_{start}(x, y) + D_{end}(x, y)}{2} \right] + \frac{2}{3} l(x, y) \left[ \frac{D_{start}(x, y) + D_{end}(x, y)}{2} \right] \]

**Proof.** The first two terms are the same as in Nourbakhsh and Ouyang (2014) and are the longitudinal distance in each H traveled by the four directional buses per unit area and the lateral distance traveled by each local bus (near bifurcation point). The additional distance is the lateral distance traveled to pick up and drop off passengers in the four directions. At \((x, y)\), we assume that half of the passenger travel in the E-W direction, and the other half in the N-S direction. Regarding the E-W direction the expected lateral distance to be traveled in a rectangle with sides of 1 and \(w(x, y)\) around \((x, y)\) is \(\frac{1}{3} w(x, y) \left[ \frac{(D_{start}(x, y) + D_{end}(x, y))}{2} w(x, y) \right]\). Thus, the expected lateral distance for E-W tubes at \((x, y)\) is \(\frac{2}{3} w(x, y) \left[ \frac{(D_{start}(x, y) + D_{end}(x, y))}{2} \right] \). Similar result is obtained for N-S tubes. \(\square\)

Result 14. The expected vehicular time expended can be expressed as follows.

\[ M(x, y) = Q(x, y) \frac{1}{v_c(x, y)} = Q(x, y) \left[ \frac{1}{v} + \frac{\tau (D_{start}(x, y) + D_{end}(x, y) + e_T(x, y))}{Q(x, y)} \right] \]

**Proof.** The locally approximated commercial speed is derived as in Result 4a to compute the expected vehicular time expended. \(\square\)

Result 15. The local out-of-vehicle waiting time at stops is given by (4.7)

\[ W(x, y) = \frac{H}{2} \left( D_{start}(x, y) + D_{d-transf}(x, y) + D_{s-transf}(x, y) \right) \]
Proof. We assume that bus headways are maintained without major variation and since they cannot be synchronized across the intersecting tubes, given the randomness of the bus travel path, wait time at all stops is on average $H/2$. □

Result 16. The in-vehicle travel time at $(x, y)$ is given by (4.8)

$$IVTT(x, y) = \frac{1}{v} (F_E + F_W + F_N + F_S) + \frac{1}{4v} \left( \frac{(F_E + F_W) \partial w}{w(x,y)/l(x,y)} + \frac{(F_N + F_S) \partial w}{l(x,y)/w(x,y)} \right) + \left( \tau + \frac{1/2w(x,y)}{v} \right) \left( \frac{D_{start}(x,y) + D_{end}(x,y)}{4} \right) * H * w(x,y) * (F_E + F_W)$$

Proof. The first two terms are similar to Ouyang et al (2014). The first one pertains to the time that onboard passenger travelling in all four directions collectively spend while the buses move one unit longitudinal distance at the cruising speed. The second one captures the travel time to make a spacing transfer near convergence and bifurcation points, where the lateral distance is approximately half of the tube width. The last term accounts for the lateral distance travelled to provide door-to-door service. Regarding E-W tubes, the time to serve each passenger within a rectangle with sides of 1 and $w(x,y)$ is approximately $\tau + \frac{1/2w(x,y)}{v}$. The number of onboard passengers that experience this door-to-door service time is $(F_E + F_W)w(x,y)$. Therefore the time that onboard passenger spend while providing lateral door-to-door service is $\left( \tau + \frac{1/2w(x,y)}{v} \right) \left[ \frac{D_{start}(x,y) + D_{end}(x,y)}{4} \right] w(x,y) (F_E + F_W) * w(x,y)$. Thus the time spent for lateral movement within a unit square around $(x, y)$ is $\left( \tau + \frac{1/2w(x,y)}{v} \right) \left[ \frac{D_{start}(x,y) + D_{end}(x,y)}{4} \right] w(x,y) (F_E + F_W)$. Similar result is obtained for tubes in the N-S hemisphere. □