THE COORDINATED DECENTRALIZED PARATRANSIT SYSTEM:

DESIGN, FORMULATION, AND HEURISTIC

A Dissertation

by

CHUNG-WEI SHEN

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

May 2012

Major Subject: Civil Engineering

The Coordinated Decentralized Paratransit System: Design, Formulation, and Heuristic

THE COORDINATED DECENTRALIZED PARATRANSIT SYSTEM:

DESIGN, FORMULATION, AND HEURISTIC

A Dissertation

by

CHUNG-WEI SHEN

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Approved by:

Chair of Committee,	Luca Quadrifoglio
Committee Members,	Sergiy Butenko
	Dominique Lord
	Bruce Wang
Head of Department,	John Niedzwecki

May 2012

Major Subject: Civil Engineering

ABSTRACT

The Coordinated Decentralized Paratransit System: Design, Formulation, and Heuristic.

(May 2012)

Chung-Wei Shen, B.B.A., National Cheng Kung University, Taiwan;

M.S., National Taiwan University, Taiwan

Chair of Advisory Committee: Dr. Luca Quadrifoglio

This dissertation investigates the different organizational structures of paratransit services that cover large regions. A paratransit service is demand-responsive, shared-ride transit service using vans or small buses. It is characterized by the use of vehicles that do not operate on a fixed route or a fixed schedule. The paratransit route and schedule are arranged from a user-specified origin to a user-specified destination, and at a user-specified time.

To retain productivity by focusing on shorter trips within a denser area, some larger systems have outsourced operations to more than one contractor, with each contractor responsible for the service zone to which their vehicles have been assigned. This service design is called a "zonal structure" or a "zoning approach."

The zoning with transfer system coordinates vehicles' schedules at various transfer locations. The schedule coordination of inter-zonal mechanisms of transportation likely reduces trip costs by increasing the ridesharing rate and lowering the number of empty return miles.

This study first presents the exact formulation for a coordinated decentralized paratransit system in order to compare its productivity and service quality with independent decentralized and centralized strategies. The formulation is then proven to work correctly, and the results of the computational experiments of small scale instances are shown to demonstrate that the proposed coordinated system is superior to independent decentralized systems in terms of passenger miles per vehicle revenue mile.

In the second section, this study develops an insertion-based heuristic method in order to compare the performances of different operational designs when applied to a large-scale system. In an experiment utilizing Houston's demand-responsive service data, we compare the productivity and service levels among three organizational structures: zoning with transfer, zoning without transfer, and no-zoning designs.

The results indicate that zoning with transfer can provide significant benefits to paratransit operations that manage zoning structure; however, the no-zoning strategy used by Houston METRO (a relatively low-density region) performs better on average in terms of efficiency. This study concludes that the zoning with transfer method can be proven to be a productive organizational structure.

ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to my advisor, Dr. Quadrifoglio, for his guidance, caring, and patience, and for always providing me with an excellent atmosphere in which to perform this research. I would also like to thank my committee members, Dr. Butenko, Dr. Lord, and Dr. Wang, for their guidance throughout the evolution of this dissertation.

Additionally, I would like to thank my parents and elder sister. They have always supported me and encouraged me with their best wishes.

Finally, I would like to thank my wife, Pei-Fen. She has always been there, cheering me up, standing by me through the good times and the bad.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER I

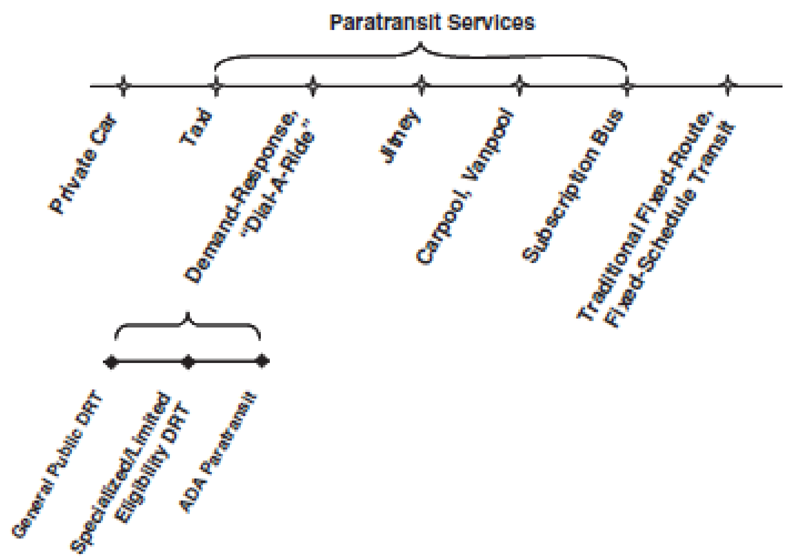
INTRODUCTION

Since the passage of the Americans with Disabilities Act (ADA) in 1990, public transit operators have been required to provide disabled passengers with a level of service comparable to that offered to regular passengers. Such service is often provided via a fixed-route bus system with proper handicap accessibility, or via ADA paratransit services. Paratransit systems are shared-ride flexible services with no fixed routes or schedules, that pick up and drop customers off at desired locations and within specified time windows. ADA paratransit systems are a demand-responsive transit (DRT) of service offering disabled customers better service than fixed-route transit systems because they provide curb-to-curb/door-to-door service and flexible schedules (Figure 1) (KFH Group et al., 2008). The scheduling/routing of paratransit systems is commonly known as the dial-a-ride problem (DARP). Each passenger is transported by a ridesharing vehicle from a specific origin to a specific destination, at a desired departure or arrival time. DARP is a subclass of vehicle-routing problems with pickups and deliveries (VRPPD), commonly faced in the transportation of goods or persons. The volume of vehicle miles traveled by paratransit services increased tremendously after passage of the ADA in 1990 (Figure 2). Paratransit services, however, are extremely

---

This dissertation follows the style of Transportation Research Part B: Methodological.

costly to operate, despite their ridesharing characteristic (Figure 3).



From: "Guidebook for Measuring, Assessing, and Improving Performance of Demand-Response Transportation." KFH Group, Transportation Research Board, 2008.

Figure 1 ADA paratransit is a type of paratransit service

In 2008, paratransit ridership made up only 1.8 percent of public transit ridership, but 13.3 percent of the total operating costs of public transit in the United States (2010 Public Transportation Fact Book, 2010). The productivity of paratransit services, with respect to passenger trips per revenue hour, is steadily decreasing (Figure 4). Hence, an improvement in productivity that would not sacrifice service quality would be very desirable, and is a much-needed goal for this industry.

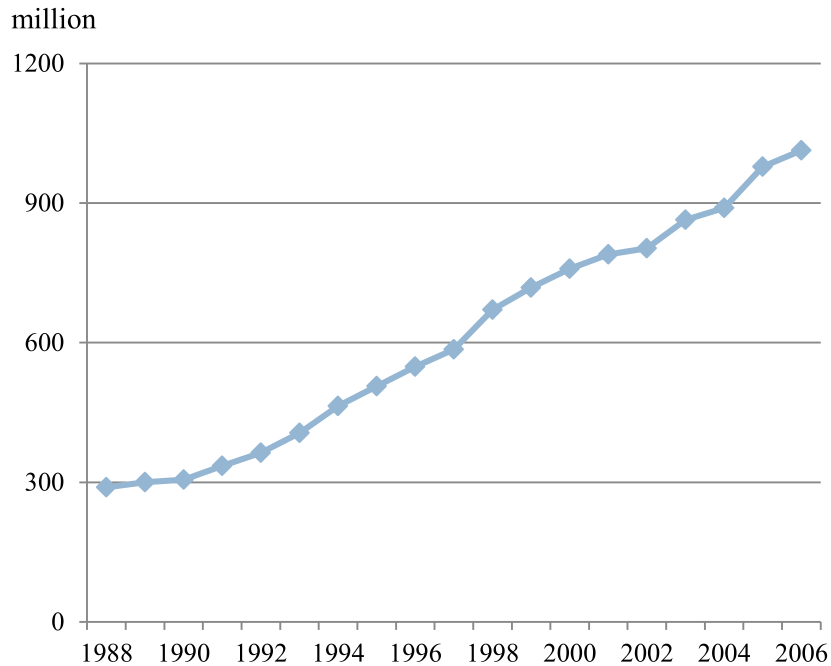


Figure 2 Vehicle total miles for paratransit system

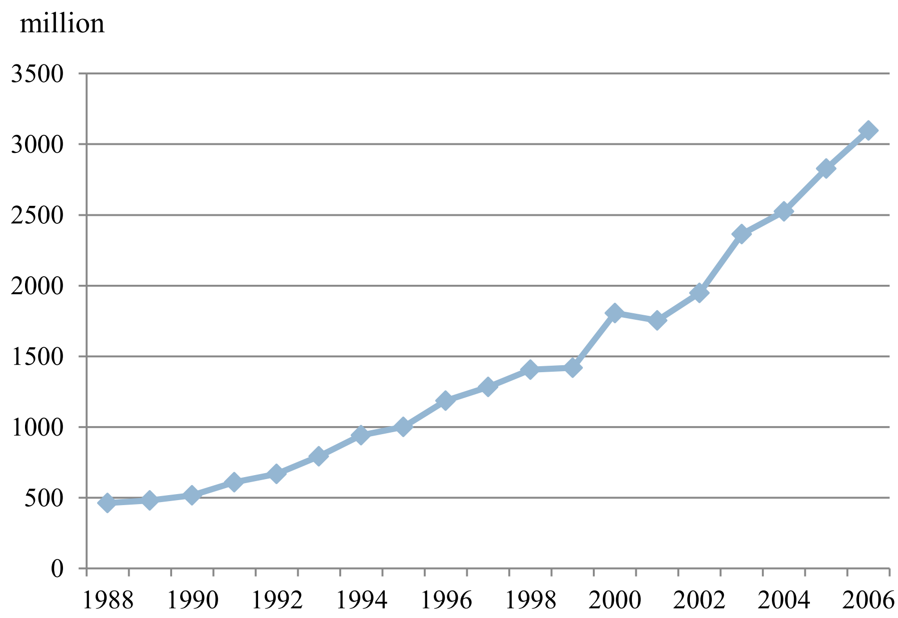


Figure 3 Operating expense for paratransit system

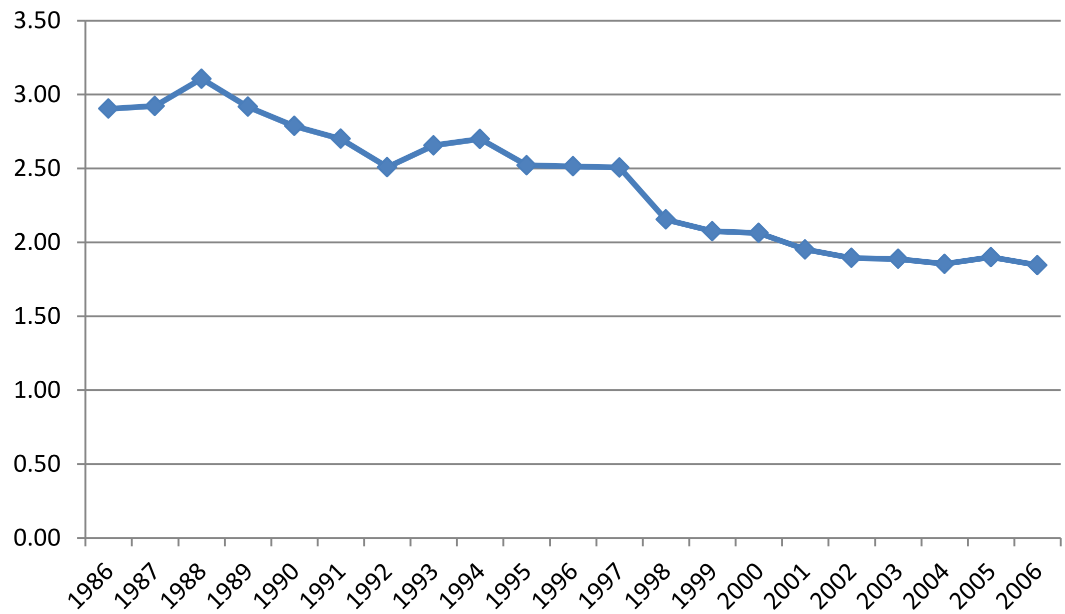


Figure 4 Passenger trips per vehicle total hour

A decentralized zoning strategy is a more practical method of operating paratransit systems, due to their ever-sprawling and ever-expanding service areas. When utilizing a decentralized strategy, service providers independently operate within their designated zones and only cross into other regions to drop off their inter-zonal customers. For example, Metropolitan Transit System in San Diego employs a decentralized strategy and divides its service area into four regions (Figure 5). In contrast, a centralized strategy considers a whole single region to be served by one designated provider (as is the case in Houston, for instance). A decentralized strategy better fits locations where there is more than one regional center; a centralized strategy better fits places where there is one compact center. Because of increasing urban sprawl, a decentralized strategy has become more popular, even in those cases where there is only one regional center.

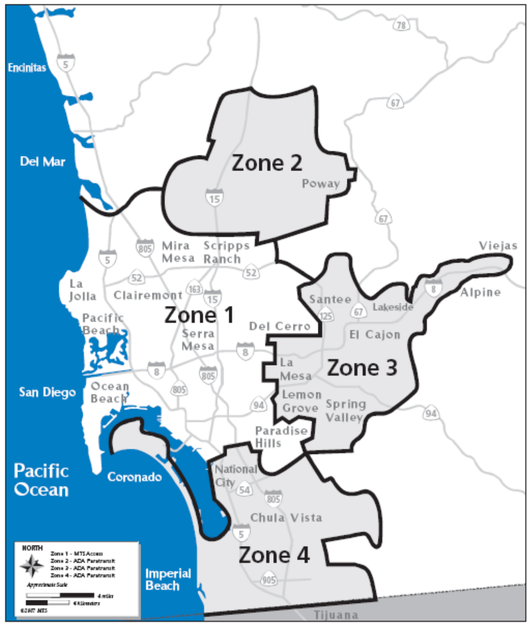


Figure 5 Service regions in the San Diego area

Utilizing a decentralized strategy, however, will likely reduce the productivity of the system; in fact, it has been found that when additional geographical constraints (zone boundaries) are added to the system, the scheduling solution cannot be improved intuitively because the set of feasible routing solutions is reduced. This is because the total vehicle empty backhaul miles (defined as the miles driven by a vehicle with no customers on board, excluding the first and last trip segments to and from the depot) increases as compared to when the centralized strategy is used. Quadrifoglio et al. (2008)

found the operating choices offered by a decentralized strategy to have a significant impact on the performance of certain demand-responsive transit services.

The first part of this dissertation provides a more in-depth analysis of the problem discussed above, comparing and contrasting centralized and decentralized (zoning) strategies when used by paratransit services. Centralized strategies treat the entire service area as a single zone; in areas using decentralized strategies, multiple zones are defined and managed independently in an effort to downgrade the operational complexity of paratransit services, particularly for large metropolitan areas. Zoning paratransit systems has several advantages. First, for service providers, smaller zones are easier to manage and control. In addition, drivers prefer to be assigned to smaller, more familiar zones, rather than to larger ones. Smaller zones also help reduce the effort to generate feasible schedules and routes, and deliver customers better quality at a higher on-time rate. Adopting the decentralized zoning strategy, however, likely increases the total number of assigned vehicles and empty trip miles, also referred to as empty trip miles, when compared to the centralized strategy. This increase remains likely even when additional geographical constraints are added to the system, because improvements in travel time cannot be made via route scheduling. In either case, despite the advantages of zoning appearing intuitively apparent, it is difficult to quantify the negative effects of zoning on scheduling solutions. Therefore, this research hopes to help planners and operators make more informed decisions and trade-offs amongst the various alternative organizational solutions, such as the centralized and decentralized approaches.

A coordinated decentralized system is a common operating practice adopted by many agencies (see paratransit in Chicago and the Twin Cities, as well as rural agencies around Dallas, for example) which requires customers to switch vehicles at preset transfer points to complete their inter-zonal trips. This practice is attracting more and more attention from transit providers because of its perceived potential to significantly reduce operating costs, mainly by reducing empty backhaul miles and increasing rideshare rates. However, the system requires synchronization between operators and increases customer discomfort, which is particularly undesirable for elderly and/or disabled riders (but is certainly more tolerable for healthy customers, thus increasing the potential benefit of adopting this operating practice for regular demand-responsive services within certain service areas). In this dissertation, the term coordination refers to the switching of inter-zonal passengers through the arrangement of the vehicle routes and schedules at specific transfer locations.

When considering the adoption of a particular coordination strategy, one should first quantify the trade-offs between the various pros and cons of that strategy. As discussed in the next section, the coordination of and integration between demand-responsive transit systems and fixed-route transit systems have been investigated in the literature; however, coordination among independent decentralized paratransit systems, especially in terms of exact formulation and heuristic solutions approaches, has been given little attention.

In an effort to fill this apparent need, this dissertation will:

- Quantify the productivity and service quality of decentralized paratransit systems, as opposed to alternative strategies currently used.
- Provide an innovative formulation of a generalization of the classic static DARP, adding the flexibility provided by considering the transfer option.
- Develop a heuristic algorithm able to handle the synchronization of vehicles between various zones.

This dissertation is divided into six chapters. Chapter 2 reviews the relevant literature on DARP. Chapter 3 includes a decision analysis between the centralized and decentralized zoning strategies, focusing particularly on the respective trade-offs. Chapter 4 contains a description of the formulation of the generalization of the DARP used for this research, adding the flexibility provided by considering the transfer option. Chapter 5 introduces a heuristic able to solve a practical-sized DARP with transfers. Chapter 6 ends with a conclusion and a set of recommendations for future research.

CHAPTER II

LITERATURE REVIEW

In this chapter the literature of two main fields is reviewed: performance models for DAR services, and formulations and algorithms for DARP.

The scheduling and routing for classic paratransit systems is known as the DARP, in the common terminology used for the study of Vehicle Routing Problems (VRP); the DARP without passenger ride time constraints are denoted by the term Vehicle Routing Problem with Pickup and Delivery (VRPPD). After Wilson and Sussman (1971) first introduced their real-time algorithms for DARP, this problem became a frequent object of study; researchers mainly have focused on developing a heuristic algorithm because of the benefits believed to be obtainable from its combinatorial characteristic.

Analytic analysis and simulation, categorized by their use of tools applicable to the evaluation of the performance of practical strategies of system design, are two major methods commonly used for this type of research. The approximate analytic model of a demand-responsive transportation system was first developed by Daganzo (1978). Daganzo provides a simple model for estimating the average total time (including both wait and ride times) in this type of system. Fu (2003) provides an analytic model predicting fleet size and quality-of-service measurements. Diana et al. (2006) proposes analytic equations to calculate fleet size in a square service area. Li and Quadrifoglio (2009) develop an analytic model which determines the optimal service zone for a feeder transit service. The analytic model is easier to use when performing a parametric analysis of a given system; however, the model makes it difficult to build a close-form expression, especially when we consider time-window constraints, irregular service areas, and non-uniform distributions of the origins and destinations of requests.

In contrast to the analytic models, simulation methods have been applied to the evaluation of performance measures, especially the effects of various system designs and stochastic event analyses of dial-a-ride systems. Wilson et al. (1970) develop a computer-aided routing system (CARS) which establishes relationships between performance parameters and different scheduling algorithms. Xiang et al. (2008) develop a simulation which evaluates the influence of different stochastic factors. In order to evaluate operational improvements from the application of automatic vehicle location technology, Fu (2002) applies a simulation model to the analysis. Shinoda et al. (2004) develops a simulation method which compares the performance of dial-a-ride systems to fixed-route bus systems. Quadrifoglio et al. (2008) considers the impact of specific operating practices and time-window settings on zoning strategies currently used by demand-responsive transit providers.

Several papers have surveyed the performance of dial-a-ride systems. Wilson and Hendrickson (1980) summarize earlier models that predict the performance of flexibly-routed transportation systems. McKnight and Pagano (1984) explore the service quality of DARP by investigating 42 service providers in the U.S. Paquette et al. (2009) conclude that further study is needed for a better understanding of the trade-offs among costs, operational policies and quality in dial-a-ride systems.

Comparatively, performance evaluations of practical strategies (such as the effects of zoning strategies on DARP) have received meager attention. The size of the service area is one key factor that affects the productivity of DRT. In general, the larger the service area, the longer the trip length, and thus DRT will not always be able to serve consistently a given number of passengers in a specified amount of time (KFH Group et al., 2008). The impact on the productivity of the different area sizes was first studied by Wilson et al. (1970). They demonstrate that the number of vehicles used is proportional to the size of the service area. Chira-Chavala and Venter (1997) adopt the data provided by the Outreach Paratransit Service in Santa Clara County, California, and observe that longer trip lengths contribute to an increase in the number of empty trip miles in an expanding service area.

In addition, a large area usually means more dispersed trips. Large service areas with dispersed trip patterns, which translate to lower demand densities, make it difficult to achieve the most beneficial effects of ride-sharing. On average, larger service areas mean more dispersed origin and destination points than those enjoyed by more compact service areas. In low-density areas, DRT systems have a lower productivity level than those systems that function in municipal areas (Ellis and McCollom, 2009).

To retain productivity by focusing on shorter trips within denser areas, some larger systems have outsourced operations to more than one contractor, with each contractor responsible for the service zone to which their vehicles have been assigned. This service design is called a zonal structure or a zoning approach. Adjacent zones generally have no overlapping or shared buffer areas. The zoning approach is attractive not only because it creates more manageable pieces of work, but more importantly because it establishes an ongoing spirit of competition throughout the contract term (Lave and Mathias, 2000). Zonal demand-responsive services are also used for dispatching, as well as for fare determination purposes (Burkhardt et al., 1995).

Coordination of paratransit services increases not only efficiency and productivity, but also mobility. From the analysis performed by Burkhardt et al. (1995), it can be concluded that around $700 million per year could be generated by transportation providers in the United States after the implementation of a successful coordination system. The consolidation of inter-zonal transportation would likely reduce trip costs by achieving higher ridesharing rates and lower numbers of empty return miles (Cook et al., 2003). Malucelli et al. (1999) present a flexible collective transportation system. They suggest that a future study might deal with allowing passengers to transfer from one vehicle to another. Häll et al. (2009) introduce the integrated DARP, where some parts of a journey could be carried out by a fixed-route service. Aldaihani and Dessouky (2003) propose a system that integrates fixed routes within a pick-up and delivery problem (PDP). An integer programming formulation for cooperative PDP with time windows was analyzed by Lin (2008). Lin concludes that the cooperative strategy could achieve savings in both total cost and vehicles used, if one assumes that all delivery locations are identical, transfer is only allowed at the last pickup location of the returning vehicle, and vehicle capacity is unlimited.

Paratransit services using a transfer system can be analyzed as a generalization of DARP. The transfer of passengers will always require more than one vehicle to fulfill a trip; therefore, the spatial and temporal synchronization constraints will, by necessity, be imposed on more than one vehicle. A schedule delay in one vehicle route may necessitate a change in all other routes. Therefore, this problem is computationally difficult even when simply trying to develop a heuristic algorithm. Shang and Cuff (1996) provide a concurrent heuristic approach to solve the PDP with transfer issue, using as an example a Health Maintenance Organization. They show that their proposed heuristic performs better than the HMO's scheduling heuristic, according to the overall lower number of delays, total travel hours, and total number of vehicles. However, this paper considers neither excess passenger travel times nor vehicle capacity constraints. Cortes et al. (2010) study a PDP with transfers issue through the process of Mixed Integer Programming (MIP). They find that the transfers permit a higher level of efficiency over the total vehicle travel time. Due to the complexity of the problem, this solution can handle only very small instances, which are maximized at six customers. They suggest further development of the transfer application for the strategic design and planning of paratransit systems.

Because this study develops an insertion heuristic for a rarely investigated design, we first surveyed the literature pertaining to the heuristic method broadly employed throughout paratransit services. We now review research related to an application similar to the CZPS.

Paratransit services are a part of the more general Demand-Responsive Services (DRS), where vehicles pick up and deliver customers at their desired locations within specifically defined time windows (either for pick-up or delivery, or both). Such services are also known as Dial-A-Ride Services (DARS). The DARP is an application of the Pickup and Delivery Problem with Time Window (PDPTW) for door-to-door transportation services, which in turn is a generalization of the Vehicle Routing Problem (VRP). Unlike DARP, PDPTW does not consider maximal travel time constraints to ensure a minimum level of service. Berbeglia et al. (2007) provide a comprehensive survey of PDPTW. DARP preserves a certain quality of service for passengers by imposing maximum ride-time constraints, a consumer protection that PDPTW does not include. The most recent surveys on DARP and PDPTW are presented by Cordeau and Laporte (2007), and Berbeglia et al. (2007), respectively.

The insertion method is a popular method used to generate routes and schedules that resolve vehicle routing problems. There are quite a few pieces of research that investigate PDPTW by applying an insertion heuristic. Use of the insertion method can be divided into categories according to how customers are "inserted" into vehicles: a sequential method maintains one route at a time, and a parallel method maintains more than one route at a time. Recently, several extensions based on basic insertion schemes have been proposed. Diana and Dessouky (2004) propose a regret-insertion heuristic for solving DARP. Alternatively, Lu and Dessouky (2006) present a new insertion-based construction heuristic that solves pick-up and delivery problems with time windows.

Meta-heuristics offer a promising method for improving solution quality. Cordeau and Laporte (2003) apply a Tabu search to the multi-vehicle DARP. Jorgensen et al. (2006) adopt the classic cluster-first, route-second approach where a genetic algorithm is used to assign customers to vehicles and a routing heuristic is used to solve independent routing problems. Other research addressing meta-heuristic algorithms used to solve variant pick-up and delivery problems include Zachariadis et al. (2009), Nanry and Wesley Barnes (2000), and Li and Lim (2001). The solution quality is dominated by extensive testing on the settings of the parameters.

There is little PDPTW literature that considers cooperative strategies. Mitrovic-Minic and Laporte (2006) propose a two-phase heuristic for PDPTW with trans-shipment problems. In this situation, vehicles are allowed to drop their loads at trans-shipment points, allowing other vehicles to carry those loads to the final delivery locations. An integer programming formulation for cooperative PDPs with time windows was analyzed by Lin (2008). This research concludes that a cooperative strategy could achieve savings in both total cost and vehicles used, if the strategy assumes that all delivery locations are identical, transfers are allowed only at the last pickup location of the returning vehicle, and vehicle capacity is unlimited. Cortes et al. (2010) demonstrate the efficiency of certain transfer designs for the PDP in certain small instances. However, due to the limitations of extremely complex characteristics, only very small problems can be solved to an optimal degree (i.e., six customers, two vehicles and one transfer point). They suggest ideas for further development of the transfer application for the strategic design and planning of the PDPTW.

2.1 Findings

Based on the above review, the existing research dealing with paratransit operating designs is still limited, and a proper decision analysis of the trade-offs between centralized and decentralized strategies has yet to be performed. The first part of this dissertation seeks to address gaps in the literature associated with zoning strategies and productivity analyses based on real paratransit demand data provided by METROLift of Houston, Texas. Because an analytic investigation of the problem is very difficult to develop without drastic approximations, a simulation approach is used to investigate it here. We compare the current centralized strategy with hypothetical but plausible decentralized scenarios that we developed according to the demand-distribution characteristics and by following METROLift's suggestions. Through simulations and statistical comparison methods, the performances of zoning strategies are analyzed.

To the best of our knowledge, the published body of research regarding cooperative solutions to DARP is still quite limited and the effects attributable to the addition of transfers between independent decentralized systems have not yet been fully studied. An especially significant gap is evidenced by the fact that no heuristic method has been proposed to solve problems of a practical size. The contribution this dissertation will make is that it will fill research gaps that have been neglected in the past.

CHAPTER III

PERFORMANCE OF ZONING SYSTEM

This chapter is organized into four sections. Section one introduces the ADA paratransit system in Houston. Section two builds the simulation model and develops zoning strategies. Section three describes the performance analysis of the simulation outputs. Section four ends with a summary of this chapter.

3.1 Data Analysis

Houston is the fourth most-populated city in the nation (trailing only New York, Los Angeles and Chicago), and it is the largest city in the southern United States. The dial-a-ride services provided in the Houston area, collectively called METROLift, are offered by the Metropolitan Transit Authority of Harris County. People with disabilities have the right to access this service. Figure 6 shows the map of the service area covered by METROLift. The rough distances from east to west and from north to south are both approximately 30 miles. The fare for a single ticket is $1.15 per ride. The operating hours are 5 a.m. to 11 p.m. from Monday to Friday, 7 a.m. to 12 a.m. on Saturdays, and 7 a.m. to 11 p.m. on Sundays and holidays. All trips need to be scheduled one day in advance. Once customers make a reservation, the schedule operator gives the customer the estimated scheduled pick-up times. These times can change plus or minus 20 minutes, for a resulting 40-minute time window (other US cities typically use a 20- or 30-minute window). Over 5,000 trips are made through this service during weekdays, and 1.44 million annual trips were provided by METROLift in 2007 (APTA 2009). The system has two depots; one is for the van provider and another is for the sedan provider. The vans can accommodate up to four wheelchairs or 10 separate ambulatory persons. Taxi cabs can accommodate up to one wheelchair or four ambulatory persons. During weekdays, the average total number of scheduled vehicles is 274 per day, including 138 vans and 136 taxi cabs. Currently, no specific zoning strategy is employed by METROLift.

In the following subsections, we analyze the actual demand data released by METROLift, including the distribution of pick-up/drop-off locations and the distribution of requested pick-up times. These distributions are then used to generate the input data for the simulation model.

METROLift Service Area

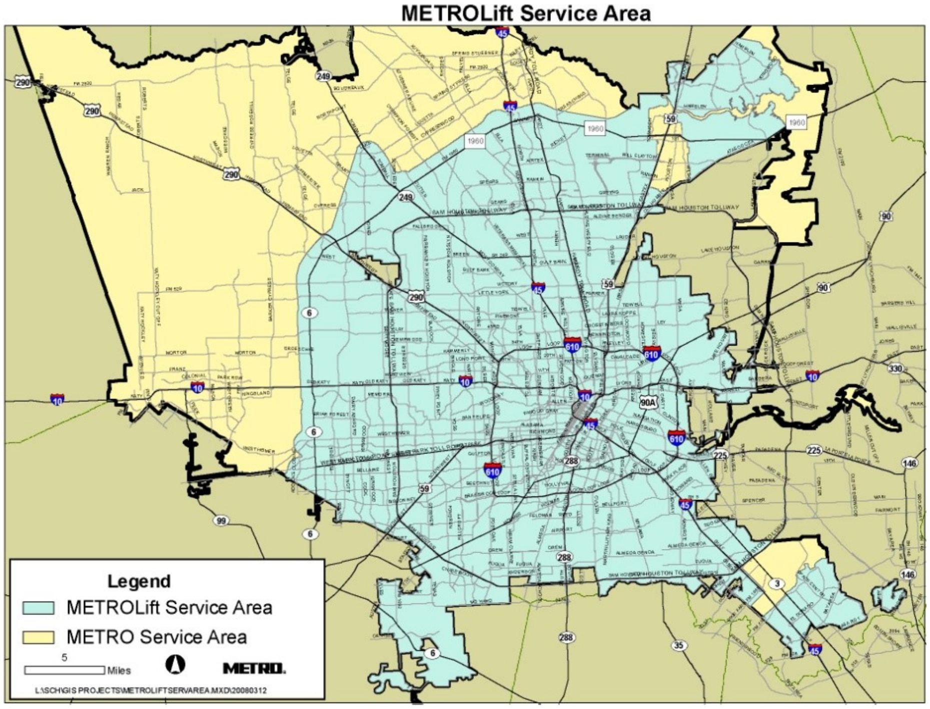


Figure 6 Service area for METROLift

3.1.1 Pick-up and drop-off locations

We use weekday travel data as the reference for location distribution. Figures 7 and 8 show the distributions of pick-up and drop-off locations. Each square in the figures represents a one-by-one mile area. Over 90 percent of requests are for roundtrips. The pick-up and drop-off locations are spread throughout the whole service area, but both contain an identical high-demand density area. Through the inspection of trip requests that travel to and from this high density area, we found that there are medical institutions within this area. The requested pick-up locations for these medical institutions are scattered across the whole area. This distribution can be seen as a single-core demand pattern.

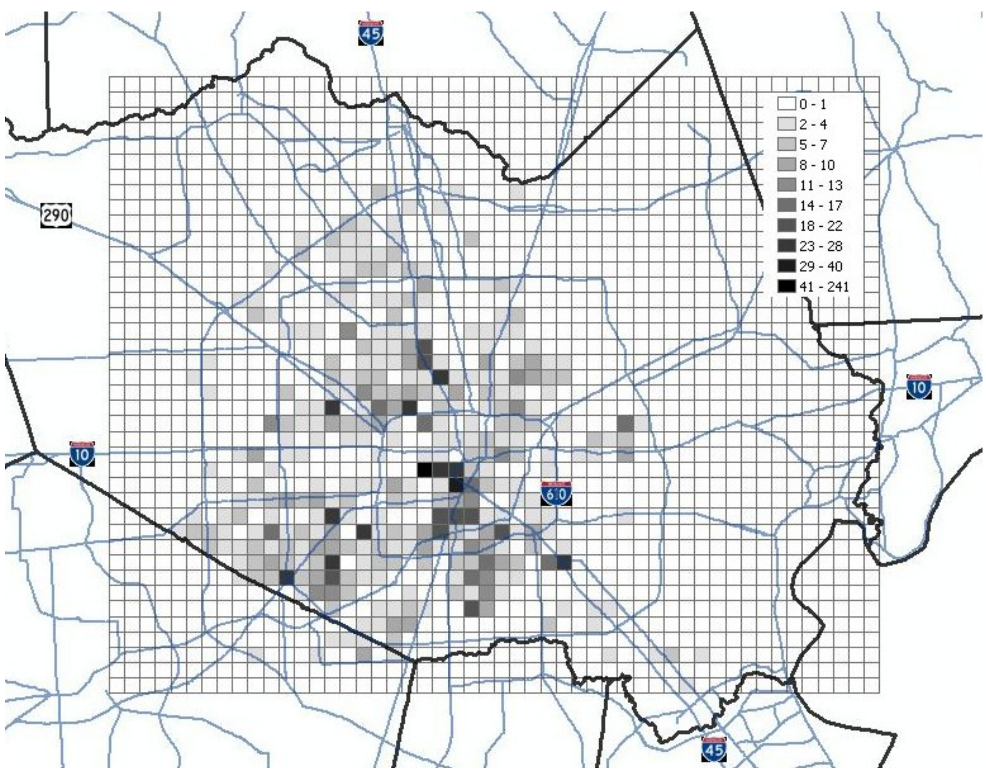


Figure 7 Distribution of pick-up locations

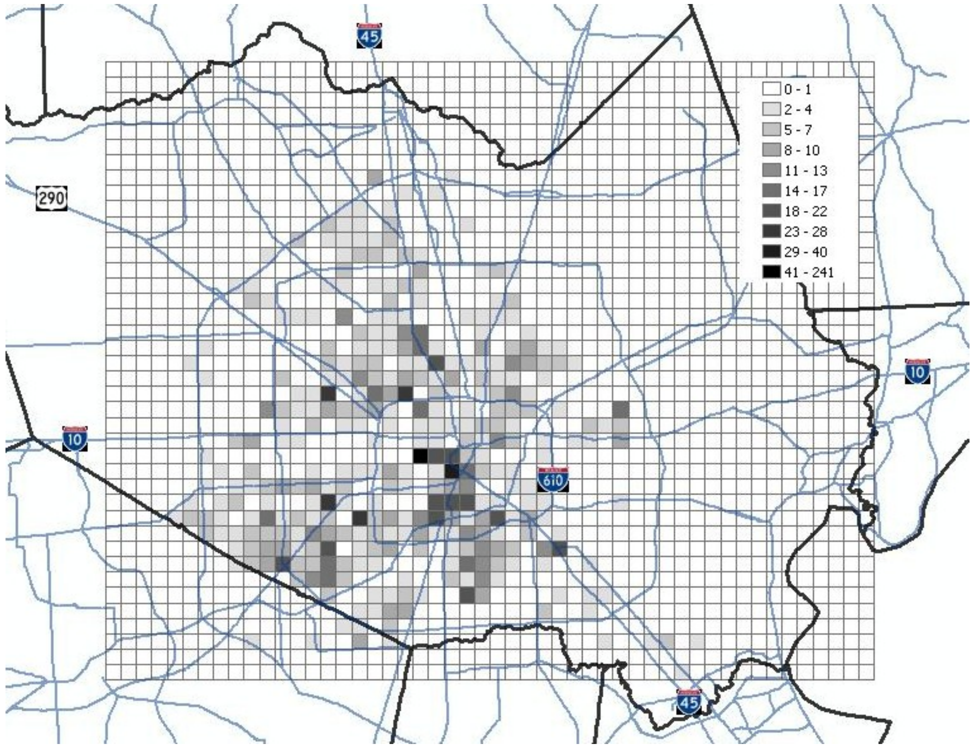


Figure 8 Distribution of drop-off locations

3.1.2 Pick-up time distribution

The distribution of requested pick-up times is shown in Figure 9. The cumulative percentage curve shows that over 90 percent of the requested pick-up times lie between 6 a.m. to 7 p.m. The morning peak hour is from 7 a.m. to 8 a.m.; the afternoon peak hour is from 3 p.m. to 4 p.m. The dial-a-ride service peak hours are more concentrated than those of other transportation systems, and the peak hours are slightly earlier, especially the afternoon peak hour. This might be due to the opening hours of most medical institutions. The requested pick-up times within the high-density area are concentrated during morning peak hours, and the pick-up times from the high-density area are concentrated during afternoon peak hours. This special time and location travel pattern must separately be reproduced in the simulation input data in order to emulate this specific demand pattern. We describe in detail this procedure in the customer-generation section.

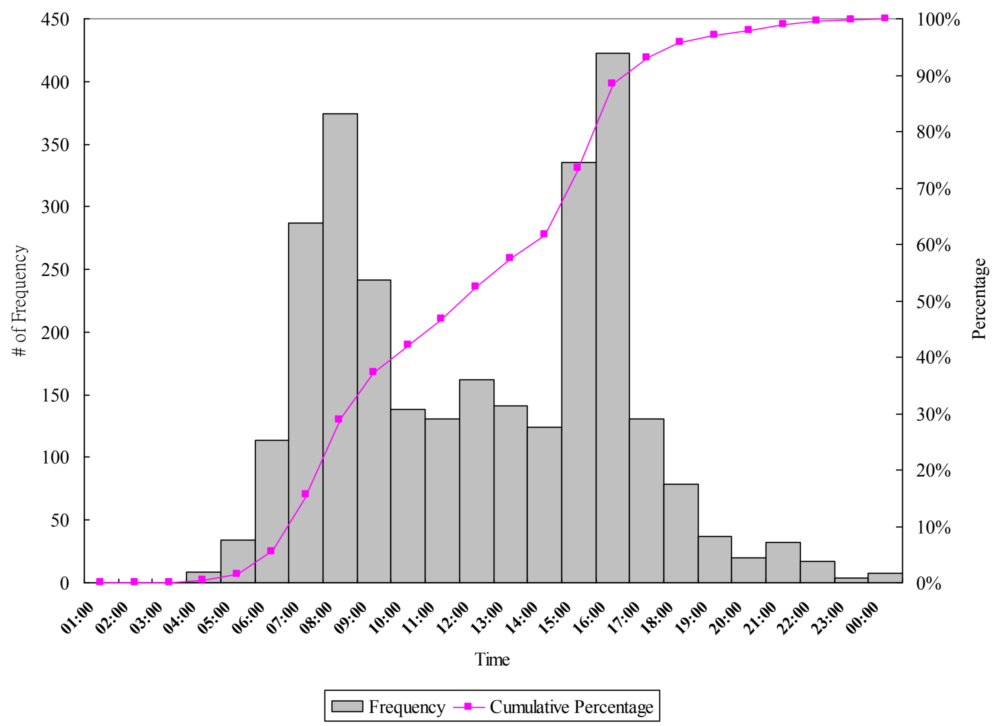


Figure 9 Distribution of requested pick-up times and cumulative percentages

3.2 Simulation Model

In this section, we present the simulation model and the zoning scenarios. First, the network assumptions are described, followed by an overview of the customer-generation method, the setting of simulation parameters, algorithm scheduling, and zoning-scenario development.

3.2.1 Network assumptions

The simulation area covers the pick-up/drop-off locations shown in the data analysis section. The Manhattan (rectilinear) distance is used to calculate the travel distance between each pair of points. For example, $\mathrm{A}(x_1, y_1)$ and $\mathrm{B}(x_2, y_2)$ represent either the pick-up or drop-off points, respectively. The travel distance between A and B can be calculated as $|x_1 - x_2| + |y_1 - y_2|$. This calculation implies that the network is arranged in a grid pattern. This estimated travel distance was verified to be reasonably close to the actual travel distance by Quadrifoglio et al. (2008). We regard the system as a deterministic case, so the travel time between two points is only a matter of travel distance and vehicle speed.

3.2.2 Customer generation

For each customer, the generation of a trip requires the following information: pick-up and drop-off location, requested pick-up time, the number of passengers, and whether a wheelchair-accessible vehicle is needed. Because more than 90 percent of requests are roundtrips, we assume that the origin of the inbound trip is the destination of the outbound trip and the destination of the inbound trip is the origin of the outbound trip. For each request, its pick-up and drop-off locations are generated independently from the pick-up and drop-off distribution. First, the pick-up and drop-off one-square-mile areas are chosen using one random number stream; another random number stream is used to determine the coordinates within the chosen one-square-mile area. The above procedure avoids generating pick-up and drop-off locations from only one specific point within the area.

There are two major sets of requests when the corresponding pick-up times are sampled. If the destination of a request is within the high-density area, the pick-up time of the outbound trip is sampled from the time distribution built by requests whose destinations are within the high-density area. If the origin of a request is within the high-density area, the pick-up time of the inbound trip is sampled from the time distribution built by requests whose origins are within the high-density area. The pick-up time of the inbound trip needs to be later than the pick-up time of the outbound trip, plus any direct travel time. If the destination and origin of a request are not within the high-density area, the cumulative distribution in Figure 9 is used to generate the requested pick-up time.

Furthermore, because the pick-up and drop-off locations are independently generated, the pick-up and drop-off points might lie within the same square-mile area, which is infeasible in reality. Therefore, if the generated drop-off location is the same as its pick-up location, a new drop-off location will be produced.

3.2.3 Parameters setting

The simulation model will use the following system parameters:

- Vehicle speed: 20 miles/hour
- Ambulatory passenger: boarding time = 1 minute; disembarking time = 1 minute
- Wheelchair passenger: boarding time = 6 minutes; disembarking time = 4 minutes
- Maximum deviation time: 40 minutes plus the requested pick-up time

● Maximum ride-time factor: customers have different parameters, according to their direct travel distances (the ratio of actual ride time divided by direct ride time)

● Size of available fleets: unlimited for vans

● Van capacity: up to 4 wheelchairs or 10 ambulatory persons

3.2.4 Scheduling algorithm

A sequential insertion algorithm is used to schedule the dial-a-ride services. The concept of the insertion algorithm is explained in the following paragraph.

The trips are ranked by ascending requested pick-up time. We insert trips one route at a time into each zone. All unassigned trips search for feasible insertions with the minimal extra travel distance. During the procedure of searching for feasible insertions, four constraints are taken into consideration. First, for each customer the drop-off time should always be later than the corresponding pick-up time. Second, the unassigned trips can only be assigned into the time slots within their pick-up and drop-off time windows. Third, after inserting the new trip, we check whether this insertion will violate the successive assigned customers' time windows. Finally, the vehicle capacity has to be satisfied in the process of inserting the unassigned trips. When each unassigned trip is inserted into a feasible position, the trip is marked as "Assigned;" otherwise it will be marked "Unassigned." If there are any "unassigned" trips left after one run, it means the existing routes cannot accommodate any unassigned trips; the existing routes are then moved to the set of generated routes. Afterwards, new empty routes are generated and the remaining unassigned trips are checked by the same insertion procedure until all trips are assigned to a route. In our algorithm, we allow both non-empty and empty-load vehicles to wait at pick-up locations before the ready service time. This assumption can increase the possibility of feasible insertions when operating the algorithm. The scheduling algorithms are coded in C++ and run on an Intel Core Due 2GHz processor. The pseudo-code of the algorithm is as follows:

Insertion Algorithm

```
begin
  while (there still are unassigned trips)
    for each depot, generate one empty route from it
    for each unassigned trip do
      check all feasible insertions where the consequence constraints, time-window
      constraints, and capacity constraints are not violated
      if (there is at least one feasible insertion) then
        select the insertion that minimizes the additional travel distance for the
      existing route
        insert the unassigned trip
        update the schedule of the inserted route and delete trips from unassigned lists
      if need be
      end if
    end for
  end while
end
```

3.2.5 Zoning scenarios

Dividing the whole service area into smaller zones can be achieved through various rules. The rules include adopting natural boundaries such as existing major highway corridors, administrative zones, the perimeter of the predefined service area, and depot locations within the service area. In a zonal-based design, trips in which a customer's pick-up and drop-off locations belong to different zones are called "inter-zonal" trips; otherwise, trips are understood as "intra-zonal." The method of accommodating inter-zonal trips into the routing schedule determines the operational type. In this study, service providers only pick up customers whose origins are within that service provider's service zone. In other words, inter-zonal trips are served by providers according to the origins associated with those trips. Therefore, for an inter-zonal roundtrip, the return trip must be made by another provider, which means that the customer is required to make two different reservations.

The key to determining service zones is to accommodate a high volume of intra-zonal trips, while balancing the percentage of inter-zonal trips within each zone. We set the zoning scenarios in Houston, and consequently found an extremely high-frequency square area containing major medical institutions where many trips began and terminated. It is roughly situated in the gravity center of the demand distribution and also the geographic center of the entire service area.

In order to better understanding the spatial-temporal characteristics of an ADA paratransit trip in the Houston area, this study created Kernel density maps for the above data in five different time periods: midnight to 6 a.m., 6 a.m. to 9 a.m., 9 a.m. to 2 p.m., 2 p.m. to 5 p.m., and 5 p.m. to midnight (Figure 4(a),(b),(c),(d), and (e)). Kernel density mapping is one of the most common methods of defining high density areas, because it details smooth and continuous probability targets within the study area (Chainey et al., 2002). The basic premise is to calculate the density of each point instead of showing the actual location of each point. The density value is highest when the distance from the point is zero, and the density decreases when the distance increases. See Equation (3.1) for the detailed calculation of the Quartic Kernel Density function (Silverman, 1986).

$$K(u) = \sum_{d<\tau} \frac{3}{\pi\tau^2}(1-\frac{d^2}{\tau^2})^2 \tag{3.1}$$

Where,

K: Kernel density value;

d: The distance from event; and,

τ: Bandwidth

From these figures, it is very easy to define the high density area of trip endpoints by colors. Darker colors represent higher trip frequencies. Figure 10 shows how those distributions change from day to night. Before 6 a.m., the origins of the ADA paratransit trips were spread out, but their corresponding destinations were concentrated in and near the central hospital area. Between 6 a.m. and 9 a.m., the trips conspicuously lead to hospital areas. In the afternoon peak hours (from 2 p. m. to 5 p.m.) and in the evenings, customers appear to be heading home from the central hospital area. The map indicates that the trips made by ADA paratransit services are highly directional during morning and afternoon peak hours.

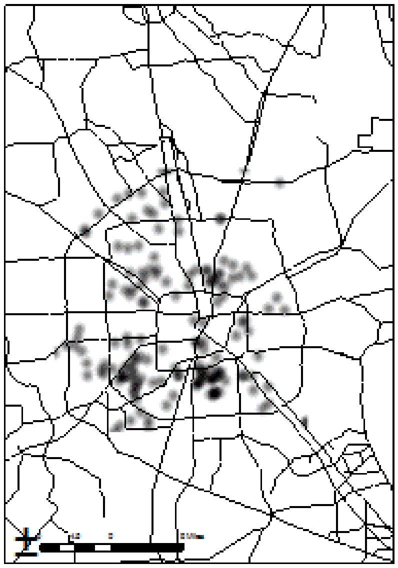
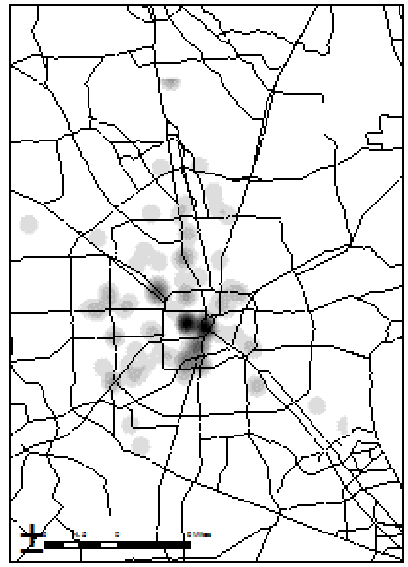
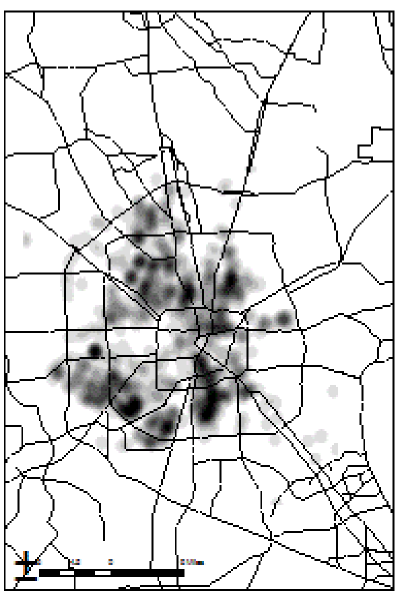
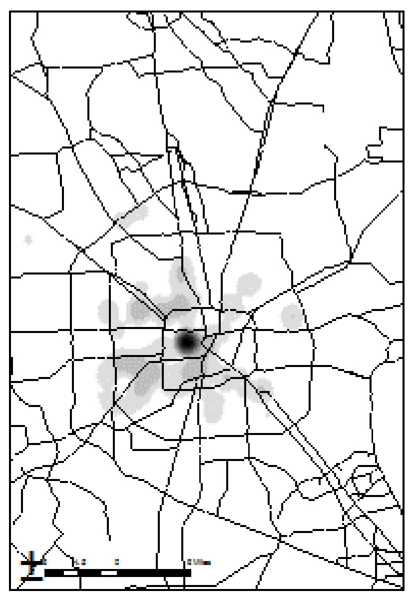
Time	Pick-up locations	Drop-off locations
0-6		
6-9		

Figure 10 Pick-up and drop-off clusters

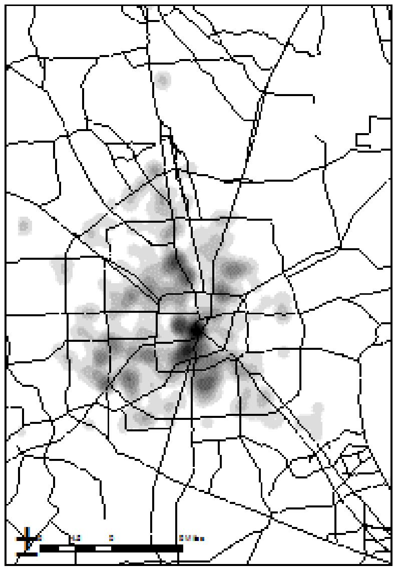
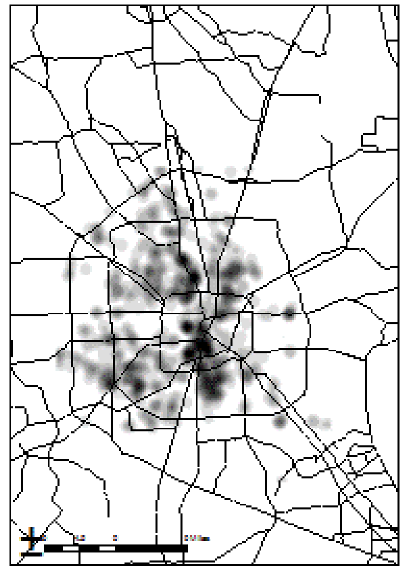
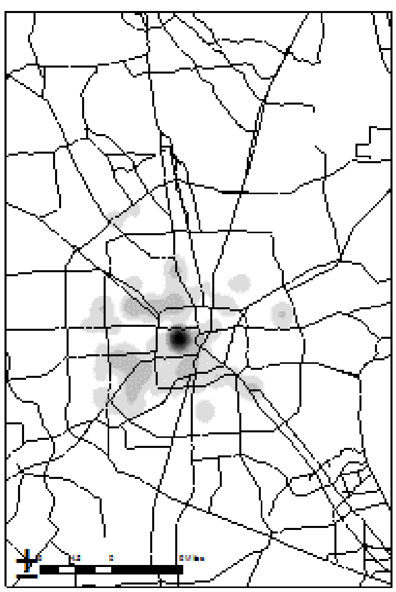
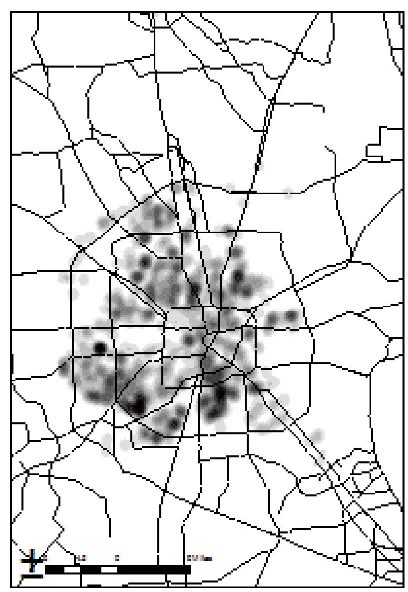
Time	Pick-up locations	Drop-off locations
9-14		
14-17		

Figure 10 continued

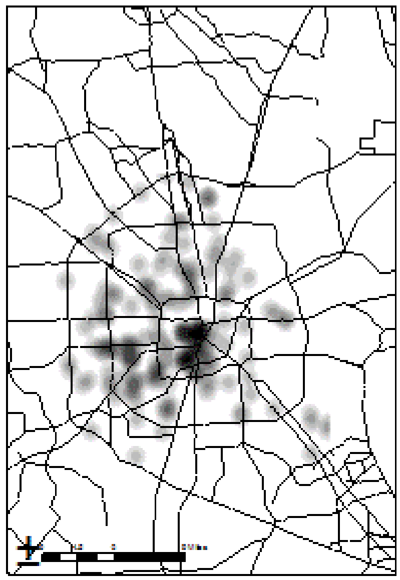
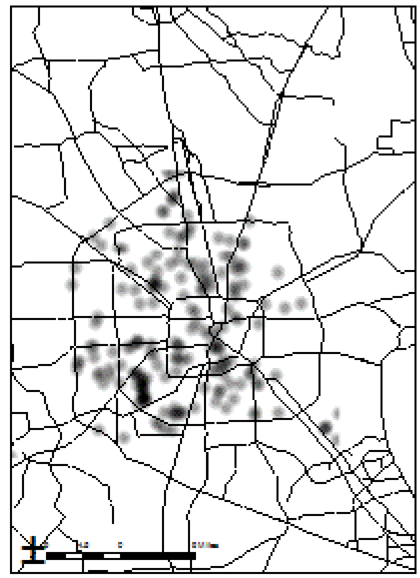
Time	Pick-up locations	Drop-off locations
17-24		

Figure 10 continued

After investigating the distributions of the customers to and from this high frequency area, it can be seen that trips traveling to and from this area are scattered evenly throughout the service region. Therefore, each zone must include a common place that people regularly visit. This square area should not be arranged into any single zone but is suitable to serve as a break or center point, with the boundary lines radiating from this square area. An advantage to this design is that one can avoid unbalanced percentages of inter-zonal trips for each zone if the area is only covered by one specific zone. According to the above approach, three zoning scenarios can be introduced: North/South, East/West, and NorthEast/NorthWest/SouthEast/SouthWest (Four Zones).

For each zoning scenario, we arrange the customers that lie within the breakpoint square area into different zones. The number of customers to and from the breakpoint square area is then categorized within the zones according to the proportion of demand requests of each zone. Table 1 summarizes the intra-zonal and inter-zonal percentages for each zone according to a corresponding zoning scenario. For zoning cases, each zone assumes one depot in the center of its zone.

Table 1 Pick-up and drop-off percentages between zones

	Drop-off			
Pick-up	NorthWest	NorthEast	SouthWest	SouthEast
NorthWest	59%	16%	18%	7%
NorthEast	34%	39%	13%	14%
SouthWest	14%	6%	64%	15%
SouthEast	9%	9%	30%	52%
Pick-up	North	South	East	West
North	74%	26%	—	—
South	19%	81%	—	—
East	—	—	57%	43%
West	—	—	23%	77%

In addition to the above three zoning scenarios, we have attempted to increase the percentage of intra-zonal trips by introducing an overlapping centered core district to create a four-zone strategy. In this case, every zone includes the core area, which could be the trip concentration center (a case similar to that of Boston's paratransit structure). In a scenario where there is a common core zone, whichever carrier brings the rider to the core zone also carries him or her back to their origination point. In such a situation, our experiments show that approximately 66 percent of the trips would be intra-zonal. We explain the effects of inter-zonal trips on the performance of paratransit systems below, in the section on analysis and comparison of zoning strategies.

3.3 Performance Analysis

In this section, we describe the simulation results based on the demand data and zoning strategy mentioned above. The performance measurements are designed first to evaluate the performance of each zoning strategy. We then utilize a statistical technique which involves multiple comparisons in order to analyze and compare alternative zoning strategies.

3.3.1 Performance measurements

We investigate the performance of zoning strategies from the dual perspectives of efficiency and service quality. With regards to efficiency, the number of routes is the most direct indicator for comparing alternative strategies for DARP. We organize the total travel distance of each assigned vehicle into three categories: vehicle travel miles to and from the depot, travel miles with no passengers on board from first pick-up to last drop-off location, and travel miles with passengers on board from first pick-up to last drop-off location.

First, the vehicle travel miles to and from the depot are known as “deadhead miles.” In practice, the METROLift does not take into account this distance when calculating their revenue miles. Second, the travel miles with no passengers on board prior to the first pick-up location and after the last drop-off location are termed “empty trip miles” in this analysis. For the operator, smaller amounts of empty trip miles are preferable, as productivity decreases with larger amounts of empty trip miles. Third, the travel miles with passengers on board can be calculated by subtracting deadhead miles and empty trip miles from the total number of miles traveled.

We also explore other useful measurements, such as passenger miles and passenger miles per total number of miles. Passenger miles are the sum of miles traveled multiplied by the number of customers on board for each travel segment. We also provide the "passenger trips per vehicle revenue hour," because in practice this figure is the most commonly used index for comparing the productivity of service for each zoning strategy.

In addition to performance measurements from the perspective of productivity, we also analyze the zoning strategies from the perspective of quality of service. From this perspective, the deviation time and ride time are our primary concerns, and not fare level. The deviation time is the time difference between the requested pick-up time and the actual pick-up time. In this experiment, the actual ride time of customers cannot exceed K times the direct ride time, due to a maximum ride-time factor corresponding to the direct travel distance.

3.3.2 Analysis and comparison of zoning strategies

The performance of alternative zoning scenarios is compared via 10 replications by simulation. In order to increase the simulation's statistical efficiency and validation, this study applies the variance-reduction technique (i.e., we synchronize a random number across different configurations on a particular replication). This procedure can help to obtain greater precision with fewer simulation replications. All pairwise confidence intervals were built to calculate certain important performance measurements for all other strategies. Table 2 shows the average results of 10 replications for each zoning strategy; the unit of time is in minutes. Here, we use numbers 1 through 5 to represent four scenarios: No Zoning (i=1), North/South (i=2), East/West (i=3), Four zones (i=4), and Four zones with core overlap (i=5).

It is worth mentioning that although our simulation contains some assumptions to simplify the actual case, the number of routes generated and obtained from the simulation is very close to the real number provided by METROLift (approximately a 5% error) for the no-zoning cases (the strategy currently adopted by METROLift). This serves as a validation of our model and its associated assumptions.

Table 2 Measurements of zoning strategies

Scenario	# of routes	Total miles	Deadhead miles	Empty miles	Passenger miles	Passenger miles/total mile	Passenger trips/revenue hour	Average deviation time*	Average Passenger Ride Time*	Revenue Hours
No-zoning	289	50,473	8,129	6,846	60,617	1.20	1.55	23.1	38.6	3226
North/South	323	52,372	6,413	9,700	60,525	1.16	1.45	22.6	38.4	3443
East/West	326	54,727	7,086	11,004	60,505	1.11	1.39	22.5	38.4	3588
Four zones	355	57,141	6,833	13,460	60,481	1.06	1.33	21.6	38.4	3757
Four zones Overlap	334	54,013	5,865	11,583	60,747	1.12	1.37	21.7	38.6	3653

* Time in minute

In order to examine whether the measurements are significantly different among the different zoning strategies, we construct all pairwise confidence intervals for five measurements: number of routes, deadhead miles, empty trip miles, passenger trips per vehicle revenue hour, and average deviation time. Because there are ten paired comparisons among five strategies, we set each individual interval at a level of 99.5 percent (1-0.05/10) to achieve a 95 percent overall confidence, according to the Bonferroni correction. In Table 3, the number represents the confidence intervals of differences $\mu_{i_2} - \mu_{i_1}$ for each measurement, for all i_1 and i_2 between 1 and 5, with $i_1 < i_2$. The numbers with asterisks in Table 3 indicate those intervals missing zero (i.e., those pairs of strategies that have significantly different numbers of assigned vehicles).

Using comparisons to each zoning case included in our simulation, we illustrate the savings according to total number of routes generated, total miles, and the empty trip miles. The deadhead miles and average deviation time increase from a zoning strategy to a no-zoning strategy. The no-zoning strategy generates the highest passenger trips per revenue hour. The passenger miles and average passenger ride-time remain almost the same in all scenarios.

Table 3 All pairwise confidence intervals of measurements

(a) Number of Routes

Paired-t		i_2			
		2	3	4	5
i_1	1	34.6 ± 8.38*	37.2 ± 5.94*	66.5 ± 7.05*	45 ± 9.14*
	2		2.60 ± 8.40	31.9 ± 5.23*	10.4 ± 9.09*
	3			29.3 ± 8.73*	7.80 ± 7.70*
	4				-21.5 ± 7.32*

(b) Deadhead Miles

Paired-t		i_2			
		2	3	4	5
i_1	1	-1716.4 ± 223.35*	-1043.8 ± 155.17*	-1296.3 ± 183.38*	-2264.6 ± 236.01*
	2		672.6 ± 248.77*	420.1 ± 207.40*	-548.2 ± 217.77*
	3			-252.5 ± 286.49	-1220.8 ± 184.70*
	4				-968.3 ± 282.25*

(c) Empty Trip Miles

Paired-t		i_2			
		2	3	4	5
i_1	1	2854.6 ± 290.04*	4158.2 ± 432.27*	6614 ± 256.06*	4737.8 ± 304.44*
	2		1303.6 ± 365.51*	3759.4 ± 179.08*	1883.2 ± 201.28*
	3			2455.8 ± 322.82*	579.60 ± 240.85*
	4				-1876.2 ± 167.37*

(d) Passenger Trips per Vehicle Revenue Hour

Paired-t		i_2			
		2	3	4	5
i_1	1	-0.098 ± 0.017*	-0.156 ± 0.029*	-0.219 ± 0.019*	-0.181 ± 0.022*
	2		-0.058 ± 0.029*	-0.121 ± 0.019*	-0.083 ± 0.021*
	3			-0.063 ± 0.021*	-0.025 ± 0.019*
	4				0.038 ± 0.018*

(e) Deviation Time

Paired-t		i_2			
		2	3	4	5
i_1	1	-0.537 ± 0.196*	-0.582 ± 0.193*	-1.487 ± 0.168*	-1.455 ± 0.270*
	2		-0.045 ± 0.202	-0.951 ± 0.170*	-0.918 ± 0.273*
	3			-0.905 ± 0.197*	-0.873 ± 0.291*
	4				0.032 ± 0.257

* denotes a significant difference

Although the total number of routes between the North/South and East/West strategies are not significantly different, the empty trip miles in the North/South zoning strategy is 12 percent lower than in the East/West zoning strategy. The passenger trips per revenue hour in the North/South zoning strategy is 4 percent higher than in the East/West zoning strategy. By introducing overlap in the four-zone case, savings are shown in the number of routes generated (6%) and empty miles (14%); thus, the passenger trips per revenue hour increase by 3 percent.

From the perspective of quality of service, the average deviation time should be as small as possible. All pairwise comparisons for the average deviation times of zoning strategies are shown in Table 3(e); we conclude that the Four Zone strategy significantly decreases the average deviation time by 6.5 percent as compared with the centralized no-zoning strategy. The comparison between the North/South and East/West strategies reveals minimal differences. We infer, then, that the Four Zone strategy geographically groups the pick-up points into considerably smaller zones as compared to the two-zone cases and the no-zoning case. The scheduling algorithm based on a minimization of extra insertion distance will help to reduce the deviation from the desired pick-up time. Additionally, an increase in the number of assigned vehicles in the Four Zones strategy also helps to decrease the average deviation time.

The effect of increasing the intra-zonal percentage is evident in the decrease of empty miles. In Figure 11(a), we plot the percentage of intra-zonal trips and empty miles for each scenario. When the intra-zonal percentage increases from 53 percent (Four-Zone strategy) to 67 percent (East/West zoning strategy), empty trip miles decrease by 18 percent, and empty trip miles decrease significantly – by 12 percent – when the intra-zonal percentage increases from 67 percent to 77 percent (as in the North/South zoning strategy). On average, for each percentage increase in intra-zonal percentage, the empty miles decrease by approximately 140 miles. This trend is validated by the performance of the Four-Zone overlap scenario. With an almost equal percentage of intra-zonal percentages, empty miles are very similar in the East/West and Four-Zone overlap scenarios.

In Figure 11(b), we provide the empty miles and passenger trips per revenue hour for each zoning scenario. There is an obvious negative correlation between these two measurements. When the empty miles decrease by one thousand, the passenger trips per revenue hour increase by 0.0035. Therefore, it is desirable to build a relationship between the intra-zonal percentage and the passenger trips per revenue mile. Figure 11(c) shows that the passenger trips per revenue mile improve by about 0.005 when the intra-zonal percentage increases by one percent. This positive correlation can be used by planners or managers when designing zoning policies.

Although the above results are based on the specific context of Houston, we believe that adding zoning constraints decreases productivity in general. The actual negative effects that might occur in other cities with different demand configurations require further study.

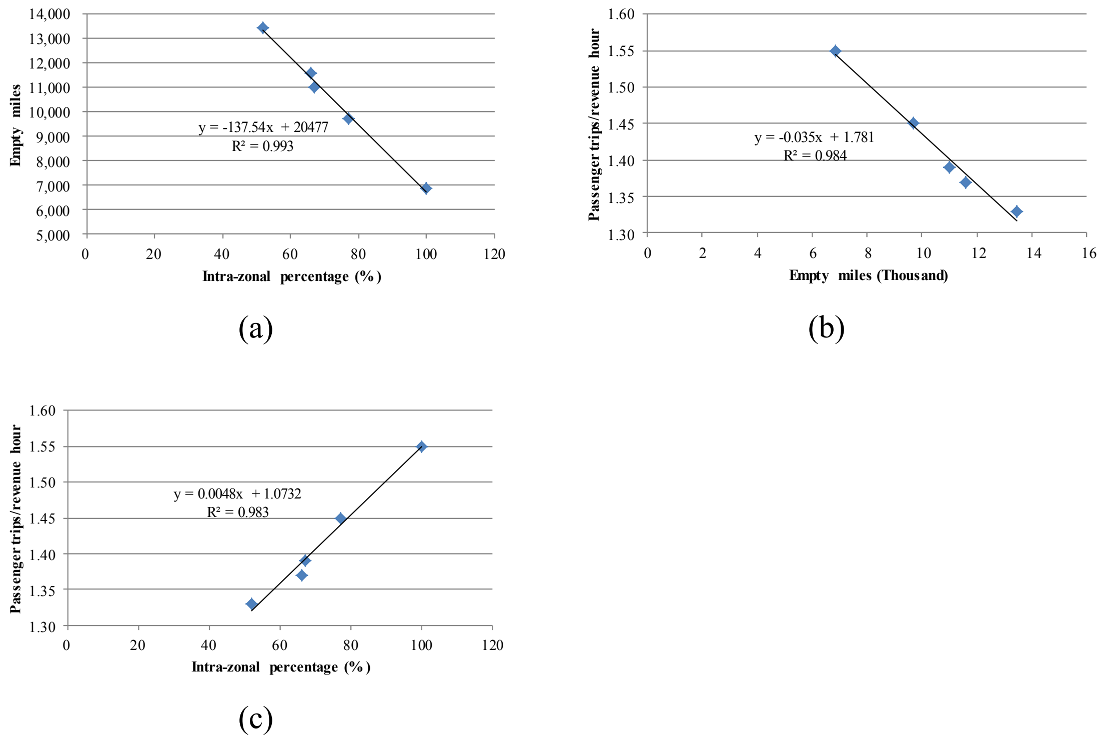


Figure 11 Trend of productivity among scenarios

3.4 Summary

In this chapter, we investigated productivity and quality of service vis-à-vis zoning strategies for ADA paratransit systems, looking at both centralized and decentralized tactics. Four zoning strategies were developed according to the distribution of pick-up and drop-off locations in Houston, Texas. A simulation model was introduced and this model can be applied to other systems (with modifications to the configuration settings).

Through simulation and statistical comparison methods, the effects of zoning strategies on ADA paratransit systems has been analyzed. From the perspective of productivity, the centralized strategy generates the smallest number of total routes and empty trip miles. With regards to the centralized no-zoning strategy, the low number of empty trip miles helps to increase the passenger trips per vehicle revenue hour. Regarding quality of service, decentralized zoning strategies decrease the average deviation time for customers. Customers' scheduled ride times remain unchanged in both the centralized and decentralized strategies.

Although we utilize the specific context of Houston, the simulation results of the performance measurements are expected to be similar in other contexts, especially in those with a one-trip concentration area. This is because the addition of the zoning constraints and pick-up restrictions reduces the number of available feasible solutions and can only degrade the overall optimal solution by increasing the number of total routes generated and decreasing passenger trips per vehicle revenue hour (as compared to the centralized strategy). However, the degree of degradation of the zoning strategies depends upon the actual demand distribution and design of the service zone.

This chapter has demonstrated the quantification of how productivity and quality of service vary with alternative centralized and decentralized zoning strategies under a demand distribution similar to that of Houston. It provides a blueprint for future investigations regarding the introduction of transfers in the zoning of paratransit services.

CHAPTER IV

FORMULATION OF TRANSFER DESIGN

In this part of the dissertation, a formulation is developed in an effort to determining the optimal performance of the transfer design. The main purpose of this formulation is to provide a rigorous optimization model. We compare the productivity and level of service of the transfer design with the zoning-without-transfer design and the centralized design.

4.1 Problem Description and Key Assumptions

In this section, we introduce the design and key assumptions of the coordinated decentralized paratransit system, adopted from actual operating paratransit services. These assumptions identify the scope of the problem and provide the basis for the following formulation.

Within the service area, a number of requests are made; each request has a specific pick-up and a drop-off location, as well as a time window identified for both of these locations. Requests will be categorized into two sets according to the pick-up and drop-off locations: inter-zonal requests and intra-zonal requests. The requests whose pick-up and drop-off locations belong to different zones are considered inter-zonal requests; the requests whose pick-up and drop-off locations belong to the same zone are considered intra-zonal requests. Figure 12 illustrates the coordinated decentralized paratransit system. Two zones are generated by boundaries; transfer points are located on the boundaries between contiguous zones. Vehicles with limited capacity return to the depots from which they depart. Under the transfer restrictions of the decentralized strategy, vehicles can only travel within their designated service zone. Vehicles picking up inter-zonal passengers need to stop at a transfer location to drop them off. Vehicles dropping off inter-zonal passengers will need to stop at a transfer location to pick them up.

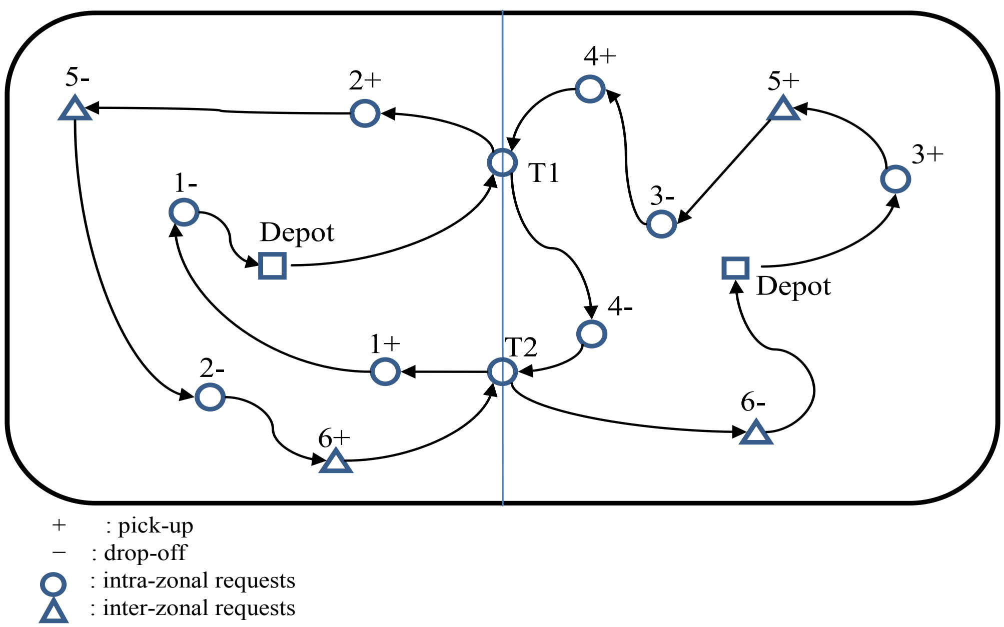


Figure 12 Example of coordinated decentralized system

At every pick-up or drop-off location, we identify only one operation: either the loading or unloading of passengers. At transfer nodes, vehicles may either load or unload passengers, or do both. To capture the differences between these operations, we generate two corresponding nodes (the load node and unload node) at each transfer location for each transfer request.

When a vehicle visits the transfer node, it either loads or unload passengers according to the node's characteristic. Figure 13 shows how vehicles enter and leave the transfer location. A pair of pick-up and drop-off nodes at each transfer location is generated.

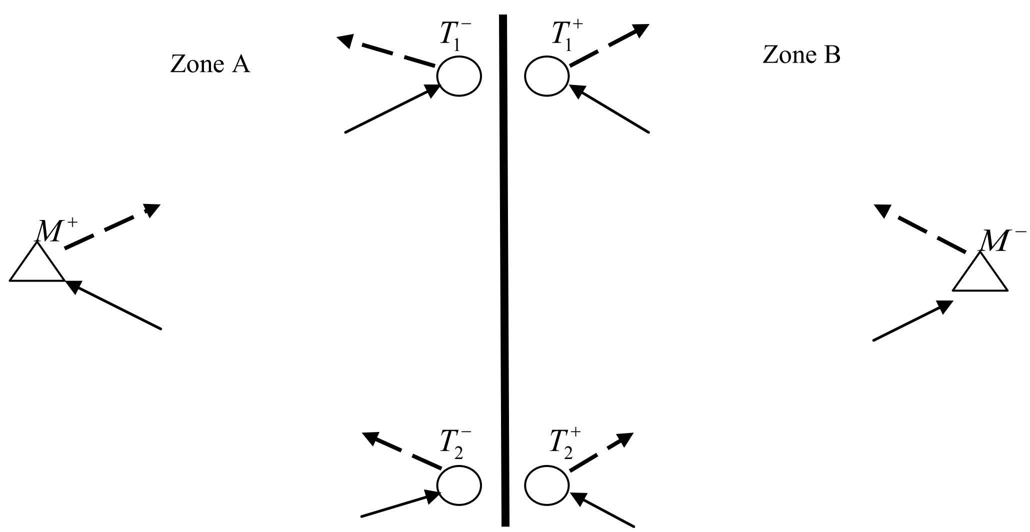


Figure 13 Transfer mechanism representation

Manhattan distances are used to calculate the symmetrical travel distances between any pair of points. Such estimated travel distances were verified to be close to the actual travel distances by (Quadrifoglio et al., 2008). For example, A (x_1, y_1) and B (x_2, y_2) represent either the pick-up or drop-off point, respectively. The travel distance is calculated as $|x_1 - x_2| + |y_1 - y_2|$. This calculation implies that the network is arranged in a grid pattern. We also assume no traffic jams in the system; the travel time between any two points is only a matter of travel distance and vehicle speed. This assumption may not allow for the consideration of precise travel times between two given points, but this does not alter the results of our performance comparison. The link distances and speeds are inputted into the model and can easily be updated with more accurate values, when available.

We assumes that each inter-zonal requester can only switch vehicles once at any one particular transfer location, and each intra-zonal requester does not switch vehicles in order to complete their trip. Concerning customer discomfort, more than one transfer is considered undesirable, if not unreasonable (Balog et al., 1997). In fact, in practice, passengers of paratransit services in Chicago and Boston are assured no more than one transfer per trip (both use a coordinated decentralized system).

All requests are known in advance, which means our problem is in a static mode. This is quite reasonable, as nearly 90 percent of paratransit customers book their rides at least a day in advance, allowing for static scheduling of the service (generally performed the night before the operation). This dissertation focuses on the productivity and service quality of the transfers design on the decentralized strategy, rather than the feasibility of accommodating all unexpected events such as the absence of customers, breakdowns of vehicles, cancellations of requests, and so on. The last assumption we make is that vehicles are allowed to wait at a pick-up node if they arrive before the earliest pick-up time, even with passengers on board.

4.2 Model Formulation

This problem is formulated as a mixed-integer programming (MIP) problem. For clarity of index and notation, we assume only two zones; the numbers of intra-zonal and inter-zonal requests are the same between the two zones in the following context. The formulation is made applicable to a multiple zones case by adding corresponding sets that operate under the assumption that exactly one transfer is allowed for every inter-zonal customer. Every request consists of pick-up and drop-off locations and the corresponding time windows. Within each zone, there are two types of requests. Each type is shown below:

$Z = \{1, ..., z\}$: Set of zones.

$N_Z = \{1, ..., n_Z\}$: Set of intra-zonal requests in zone z.

$M_z = \{1, ..., m_z\}$: Set of inter-zonal requests whose pick-up nodes are in zone z.

The node sets within each zone include:

N_z^+: Set of pick-up nodes for requests in N_z.

N_z^-: Set of drop-off nodes for requests in N_z.

M_z^+: Set of pick-up nodes for requests in M_z.

M_z^-: Set of drop-off nodes for requests in M_z.

$V(i)$: Set of nodes that are within the same zone as node i.

$T = \{1, ..., t\}$: Set of transfer locations for M_z.

S_z: Set of all nodes within zone z.

For each request r belonging to M_z, the set of drop-off nodes is generated at each transfer location in the same zone as node $i \in M_z^+$; for each request r belonging to M_z, the set of pick-up nodes is generated at each transfer location in the same zone as node $i \in M_z^-$. The corresponding notations are:

$Z(i)$: Zone of node i.

$d(r)$: Set of generated drop-off nodes at transfer location for request $r \in M_z$.

$p(r)$: Set of generated pick-up nodes at transfer location for request $r \in M_z$.

$t(r)$: Set of generated paired pick-up and drop-off nodes at transfer locations for request $r \in M_z$.

Let K_z be the set of vehicles in zone z. Every vehicle leaves from and arrives to the same depot and has a capacity Q_k, and each node $i \in S_z$ is associated with a load q_i. Every arc (i, j) is associated with a routing cost c_{ij}^k and a travel time t_{ij}^k for vehicle k. For each node $i \in S_z$, e_i and l_i represent the earliest and latest times at which service may begin, and d_i is the service duration at node i. F denotes the fixed cost of assigning each extra vehicle. W represents the maximum passenger wait times at transfer locations. G is a sufficiently large constant (usually noted as '*M*', which would, however, conflict with other notations in this research). Three decision variables are introduced in our formulation. First, the binary variable x_{ij}^k equals 1 if vehicle k uses link (i,j), and otherwise equals 0. Second, for each node i and each vehicle $k \in K_z$, let B_i^k be the time variable at which vehicle k begins service (either pick-up or drop-off) at node i. Third, Q_i^k is the load variable of vehicle k after visiting node i. The following is the formulation of the coordinated decentralized paratransit system:

$$Min \sum_{z\in Z}\sum_{i,j\in S_z}\sum_{k\in K_z} c_{ij}^k x_{ij}^k + F\sum_{z\in Z}\sum_{k\in K_z}(1 - x_{0,2n_z+2m_z+2tm_z+1}^k) \quad (4.1)$$

Subject to

$$\sum_{k\in K_z}\sum_{j\in V(i)} x_{ij}^k = 1 \quad \forall z \in Z,\ i \in (N_z^+ \cup N_z^- \cup M_z^+ \cup M_z^-) \quad (4.2)$$

$$\sum_{k\in K_z}\sum_{j\in V(i)} x_{ji}^k - \sum_{k\in K_z}\sum_{j\in V(i)} x_{ij}^k = 0 \quad \forall z \in Z, i \in S_z \quad (4.3)$$

$$\sum_{j\in S_z} x_{0j}^k = 1 \quad \forall z \in Z, k \in K_z \quad (4.4)$$

$$\sum_{i\in S_z} x_{i,2n_z+2m_z+2tm_z+1}^k = 1 \quad \forall z \in Z, k \in K_z \quad (4.5)$$

$$\sum_{k\in K_{z(j)}}\sum_{i\in(V(j)\backslash d(r))}\sum_{j\in d(r)} x_{ij}^k = 1 \quad \forall r \in M_z \quad (4.6)$$

$$\sum_{k\in K_{z(j)}}\sum_{i\in(V(j)\backslash p(r))}\sum_{j\in p(r)} x_{ij}^k = 1 \quad \forall r \in M_z \quad (4.7)$$

$$\sum_{k\in K_{z(i)}}\sum_{a\in V(i)} x_{a,i}^k = \sum_{k\in K_{z(j)}}\sum_{b\in V(j)} x_{j,b}^k \qquad \forall r \in M_z, \forall\, i,j \in t(r)\ and\ i \in d(r), j \in p(r) \quad (4.8)$$

$$\sum_{j\in V(i)} x_{ij}^k - \sum_{j\in V(i)} x_{j,i+n_z}^k = 0 \quad \forall z \in Z, i \in N_z^+, \forall k \in K_z \quad (4.9)$$

$$\sum_{j\in V(i)} x_{ij}^k - \sum_{j\in V(i)}\sum_{t=1...|T|} x_{j,i+2tm_z}^k = 0 \quad \forall z \in Z, i \in M_z^+, k \in K_z \quad (4.10)$$

$$\sum_{j \in V(i)} x_{ij}^k - \sum_{j \in V(i)} \sum_{t=1...|T|} x_{j,i+2tm_z}^k = 0 \quad \forall z \in Z, i \in M_z^-, k \in K_z \tag{4.11}$$

$$B_j^k \geq B_i^k + d_i + t_{ij} - G(1 - x_{ij}^k) \quad \forall z \in Z, i, j \in S_z, k \in K_z \tag{4.12}$$

$$e_i \leq B_i^k \leq l_i \quad \forall z \in Z, i \in S_z, k \in K_z \tag{4.13}$$

$$B_i^k + d_i \leq B_j^{k'} \leq B_i^k + d_i + W \quad \forall r \in M,\ \forall i, j \in t(r)\ and\ i \in d(r), j \in p(r), k \in K_{z(i)}, k' \in K_{z(j)} \tag{4.14}$$

$$B_{i+n_z}^k - (B_i^k + d_i + t_{i,i+n_z}) \geq 0 \quad \forall z \in Z, i \in N_z^+, k \in K_z \tag{4.15}$$

$$G(1 - \sum_{j \in S_z} \sum_{k \in K_z} x_{j,i+2tm_z}^k) + B_{i+2tm_z}^k - (B_i^k + d_i + t_{i,i+2tm_z}^k) \geq 0 \quad \forall z \in Z, i \in M_z^+, t = 1...|T|, k \in K_z \tag{4.16}$$

$$G(1 - \sum_{j \in S_z} \sum_{k \in K_z} x_{j,i+2tm_z}^k) + B_i^k - (B_{i+2tm_z}^k + d_i + t_{i,i+2tm_z}^k) \geq 0 \quad \forall z \in Z, i \in M_z^-, t = 1...|T|, k \in K_z \tag{4.17}$$

$$Q_j^k \geq Q_i^k + q_j - G(1 - x_{ij}^k) \quad \forall z \in Z, i, j \in S_z, k \in K_z \tag{4.18}$$

$$\max\{0, q_i\} \leq Q_i^k \leq \min\{Q_k, Q_k + q_i\} \quad \forall z \in Z, i \in S_z, k \in K_z \tag{4.19}$$

$$x_{ij}^k = \{0,1\} \tag{4.20}$$

$$B_i^k, Q_i^k \geq 0 \quad \forall z \in Z, i \in S_z, k \in K_z \tag{4.21}$$

The objective function (4.1) minimizes the total traveling cost plus the total vehicle fixed cost, where 0 and $2n_z + 2m_z + 2tm_z + 1$ represents the origin and destination depot in each zone, respectively. For example, if $x_{0,2n_z+2m_z+2tm_z+1}^k$ equals 1, this means vehicle k goes to the destination depot directly from the origin depot, and thus will not incur the fixed cost F. Constraint (4.2) guarantees that all pick-up and drop-off nodes, except those at transfer locations, must be visited exactly once, and constraint (4.3) ensures the flow conservation for all nodes. Constraints (4.4) and (4.5) guarantee that each vehicle route starts out and returns to the depot, respectively. Constraints (4.6) and (4.7) ensure that exactly one node is chosen from all possible transfer nodes for each inter-zonal request. For inter-zonal requests, constraint (4.8) defines the flow conservation of each pair of pick-up and drop-off nodes at each transfer location. These constraints ensure that vehicles can only pick up inter-zonal customers at the transfer locations to which those customers are delivered. Constraints (4.9) to (4.11) are pairing constraints; each paired request must be served by the same vehicle, where t is the number of transfer locations. Constraints (4.12) and (4.13) guarantee time consistency, and key constraint (4.14) ensures the that transfer times cannot exceed the maximum passenger wait times at the various transfer locations, ensuring an acceptable service level.

Constraints (4.15) to (4.17) are precedence consistency for inter-zonal and intra-zonal requests. The "Big *G*" element of constraints (4.16) and (4.17) is crucial because exactly one drop-off (or pick-up) node at each transfer location will be chosen for inter-zonal requests. The precedence constraints do not apply to those transfer nodes that are not chosen. In Figure 14, the solid arrows illustrate the optimal solution and the dotted arrows illustrate feasible but not optimal solutions. The number next to each arrow corresponds to travel times. The paired numbers in the parentheses are the arrival and departure times. The wait time at each transfer node is 10. Without the Big *G* element, the arrival time at the final destination would be incorrectly decided by the larger number (which is 60 in this case). However, the correct arrival time of M^- would be 50, as T_2 is not used for the transfer of this example's customer. Constraints (4.18) and (4.19) are capacity constraints.

The model presented above significantly grows in size with the number of inter-zonal and intra-zonal requests. Each inter-zonal request adds $2t$ nodes, where t is the number of transfer locations in both its pick-up and drop-off zones; each inter-zonal request also adds the corresponding arcs to the network resulting in a rapidly increasing solve time.

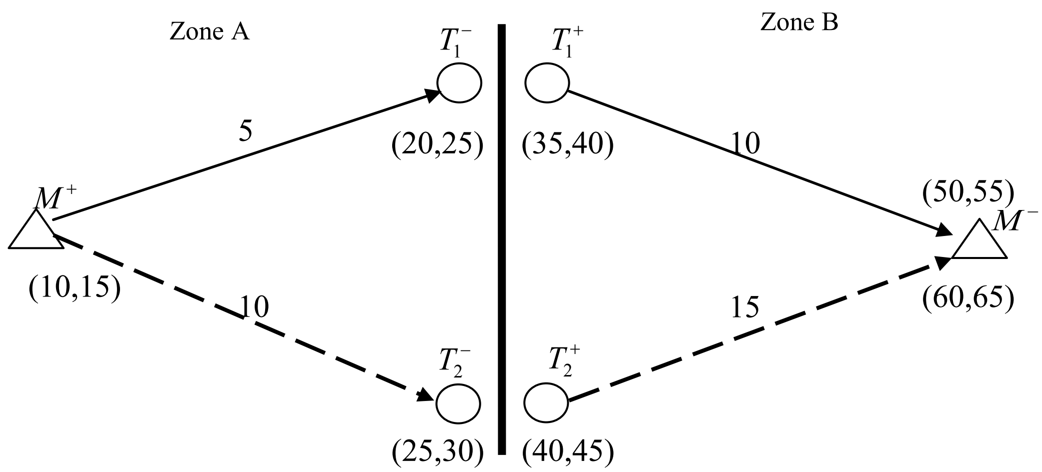


Figure 14 Example of inter-zonal request

We applied the following straightforward arc elimination rules to reduce the network size because these arcs cannot belong to a feasible solution:

No arc can go from the destination depot (node) to any node.

No arc can go to the origin depot (node) from any node.

No arc can go from any pick-up point to the destination depot.

No arc can go from the origin depot to any drop-off node.

No arc can go from the drop-off node to its corresponding pick-up node.

No arc can go between each pair of generated drop-off (pick-up) nodes at transfer location for the same request.

4.3 Parameters Setting

The optimization model will use the following system parameters:

Vehicles' speed: 30 miles/hour. (This converts the travel distance into travel time.)

Passenger load or unload time: 5 minutes.

Time-windows: 20 minutes plus the requested pickup time.

Maximum travel time factor: 2.5 (the ratio of maximum travel times divided by direct travel times).

Maximum passenger waiting times at transfer point: 10 minutes.

Number of available vehicles: unlimited.

Vehicle capacity: 5 persons.

4.4 Numerical Results Analysis

We have developed a strict formulation of a coordinated decentralized paratransit system, wherein we assume that transfer positions are fixed and known in advance. In this section, we compare the paratransit system performance among the independent decentralized, coordinated decentralized, and centralized strategies. The question then is whether the proposed transfer mechanism can improve upon the productivity of paratransit systems. The formulation was implemented by using ILOG OPL 6.3 and CPLEX 12.1. It was run on a 2.33 GHz Core2 Duo with 2 GB of memory. Experimental results conducted with CPLEX demonstrate the validation of the proposed innovative formulation.

4.4.1 Performance measurements

We investigate the performance of various centralized and decentralized strategies from the perspectives of productivity and service quality. With regards to productivity, the number of vehicles used is the most direct indicator for comparing the respective efficiencies of alternative strategies for DARP. Vehicle revenue miles are another measurement; we define vehicle revenue miles as the sum of the miles traveled from the first pick-up location to the last drop-off location for all vehicles used.

Passenger miles traveled are the sum of miles traveled multiplied by the number of customers on board for each travel segment. Passenger miles per vehicle revenue mile is one of the performance categories to measure the productivity of the transit system. This measurement is adopted by Safe, Accountable, Flexible, Efficient Transportation Equity Act: A Legacy for Users (SAFETEA-LU) to establish a provision in the FTA Urbanized Area Formula program to distribute funds to urbanized areas under a population of 200,000. For the operator, a lower number of empty backhaul miles are preferable because passenger miles per vehicle revenue mile increase when empty backhaul miles decrease.

In the course of performing this research, we analyzed the performance measurements of service quality for several different strategies. Regarding service quality, wait times and scheduled travel times are the primary concerns for customers (second only to fare level). Wait times are measured as the time differences between the requested pick-up times and the actual pick-up times. Again, the actual travel times of customers cannot exceed 2.5 times that of the direct travel times because of implementation of the maximum travel time factor for both the intra-zonal and inter-zonal requests.

4.4.2 Simulated instances and comparison of strategies

Each test case includes six requests in each zone, half of which are transfer requests. In all instances, the coordinates of the six pick-up nodes for each sub-zone are generated in the 10 × 20 square mile area (the expected direct travel time for trips within this area is 20 minutes); the entire service area is a 20 × 20 square mile are (the expected direct travel time for trips within this area is 26.7 minutes). Regarding the drop-off nodes for each set of six requests, three nodes are chosen within the same zone and the other three nodes are chosen from the adjacent zone. The simulated duration of the generated pick-up time is 150 minutes. For the independent decentralized scenario, transfers are not allowed and vehicles can only drop-off customers outside their designated service zone. There is one depot in each zone for decentralized cases. In the centralized strategy, no additional transfer and pick-up restrictions are added to the system and one depot is provided. The above two strategies adopt the formulations developed by Cordeau (2006). In the coordinated scenario, two transfer locations are available between the two zones. If we want to minimize both the total vehicle revenue miles and the number of vehicles used in this study, the location of depot does not affect the results of either the vehicle revenue miles or the passenger miles since deadhead miles (i.e., miles traveled when leaving or returning to the depot) are not included in the vehicle revenue miles.

After obtaining the preliminary results from the test cases, we found that a two-way passenger exchange at the transfer points largely decreases the number of empty backhaul miles, as compared to a one-way passenger feed. Therefore, we compare the performances of alternative centralized and decentralized scenarios through 15 instances where passengers are exchanged at transfer locations. To increase the experiment's statistical efficiency and validation, this research applies the variance-reduction technique to synchronize random numbers across the different configurations in each replication. This procedure helps obtain greater precision through the performance of fewer runs. Tables 4 and 5 show the computational results of 15 instances for each strategy. Here, the numbers 1 through 3 represent three scenarios: the Independent Decentralized $(i=1)$, Coordinated Decentralized $(i=2)$ and Centralized strategies $(i=3)$.

We observed that the centralized strategy has the lowest number of total vehicles used. This result is not beyond our speculation because the centralized strategy can be seen as pursuing global optimization (since no additional transfers or geographical constraints are added) under the objective of minimizing the total number of vehicles used, along with the total number of vehicle revenue miles.

Table 4 Productivity for the 15 replications

Run	Independent Decentralized				Coordinated Decentralized				Centralized			
	Vehicles used	Vehicle revenue miles	Passenger miles	Passenger miles/vehicle revenue mile	Vehicles used	Vehicle revenue miles	Passenger miles	Passenger miles/vehicle revenue mile	Vehicles used	Vehicle revenue miles	Passenger miles	Passenger miles/vehicle revenue mile
1	4	178.30	199.19	1.1171	4	161.90	207.62	1.2823	3	153.87	219.62	1.4273
2	4	190.37	171.43	0.9005	4	183.67	176.59	0.9615	3	217.70	229.86	1.0559
3	4	180.75	213.51	1.1812	4	170.82	218.68	1.2802	3	184.21	217.81	1.1824
4	5	170.06	206.47	1.2141	4	188.94	218.14	1.1545	3	201.35	270.47	1.3433
5	4	190.85	205.39	1.0762	4	184.68	223.54	1.2104	3	211.30	249.74	1.1819
6	4	223.69	273.69	1.2235	4	196.66	245.31	1.2474	4	190.81	239.46	1.2549
7	4	185.40	238.84	1.2883	4	175.23	231.59	1.3217	3	197.47	284.95	1.4430
8	5	210.89	240.92	1.1424	4	234.23	296.85	1.2673	4	206.67	248.26	1.2013
9	4	209.97	248.13	1.1818	4	194.43	235.87	1.2132	4	168.41	225.84	1.3410
10	4	221.44	262.14	1.1838	4	209.50	291.35	1.3907	4	190.75	254.06	1.3319
11	5	164.89	201.08	1.2195	4	170.10	225.75	1.3272	4	157.10	219.57	1.3976
12	5	177.86	203.52	1.1443	4	184.47	245.09	1.3286	3	183.91	225.10	1.2240
13	5	220.84	228.24	1.0336	5	203.60	234.07	1.1496	4	207.86	241.93	1.1639
14	4	186.15	211.03	1.1336	4	183.07	247.53	1.3521	4	153.80	213.57	1.3886
15	4	219.39	255.30	1.1637	4	199.59	251.68	1.2610	4	184.40	240.86	1.3061
Avg	4.33	195.39	223.93	1.1469	4.07	189.39	236.64	1.2498	3.53	187.31	238.74	1.2829

Table 5 Service quality for the 15 replications

Run	Independent Decentralized		Coordinated Decentralized		Centralized	
	Total waiting times	Average scheduled travel times	Total waiting times	Average scheduled travel times	Total waiting times	Average scheduled travel times
1	58.80	43.38	90.23	58.33	70.82	52.10
2	120.47	30.24	89.38	45.44	84.25	44.69
3	105.86	46.74	50.93	59.55	80.76	47.23
4	65.95	41.84	153.05	54.34	66.93	60.76
5	63.36	45.85	107.29	61.62	107.5	51.21
6	99.23	58.20	72.12	59.46	91.08	59.43
7	31.56	49.25	54.55	67.05	105.91	60.15
8	121.00	47.57	123.05	64.06	122.73	48.74
9	81.92	50.18	106.18	53.45	40.26	46.70
10	49.87	55.65	95.88	71.83	53.37	57.20
11	96.24	48.05	56.24	51.70	66.45	51.71
12	91.95	46.81	37.47	61.46	56.17	43.34
13	42.93	41.00	86.06	64.88	73.40	49.06
14	53.21	46.70	65.53	62.32	106.46	51.17
15	20.12	52.96	86.84	66.74	50.55	55.73
Avg	73.50	46.96	84.99	60.15	78.44	51.95

Unit: Minute

In order to examine whether the measurements are significantly different among the three strategies, we constructed all of the pairwise confidence intervals for the six measurements: the number of vehicles used, vehicle revenue miles, passenger miles, passenger miles per vehicle revenue mile, average customer wait times, and scheduled travel times. Because there are three pairs of comparisons among the three strategies, we set each individual interval at the level of 98.33 percent (1-0.05/3) to achieve a 95 percent overall confidence level, according to the Bonferroni correction. In Table 3, the number represents the confidence interval of the difference $\mu_{i_2} - \mu_{i_1}$ for each measurement, for all i_1 and i_2 between 1 and 2, with $i_1 < i_2$. The numbers with asterisks in Table 6 indicate those intervals missing zero (i.e., those pairs of measurements are significantly different under the corresponding strategies).

Table 6 All pair-wise confidence intervals of measurements

(a) Vehicles Used			
ed-t		i_2	
		2	3
i_1	1	-0.27 ± 0.31	- 0.80 ± 0.46*
	2		- 0.53 ± 0.35*

(b) Vehicle Revenue Miles			
red-t		i_2	
		2	3
i_1	1	- 6.00 ± 9.92	- 8.08 ± 17.01
	2		-2.08 ± 14.27

(c) Passenger Miles			
ed-t		i_2	
		2	3
i_1	1	12.72 ± 15.54	14.81 ± 20.32
	2		2.09 ± 22.85

(d) Passenger Miles per Vehicle Revenue Mile			
red-t		i_2	
		2	3
i_1	1	0.1029 ± 0.0533*	0.1360 ± 0.0553*
	2		0.0330 ± 0.0644

(e) Waiting Times			
ed-t		i_2	
		2	3
i_1	1	11.49 ± 31.11	4.94 ± 25.07
	2		- 6.54 ± 27.06

(f) Scheduled Travel Time			
ed-t		i_2	
		2	3
i_1	1	13.19 ± 4.24*	4.99 ± 4.41*
	2		- 8.20 ± 4.96*

In Table 6(a), the total number of vehicles used in the coordinated decentralized strategy is not smaller than the number used in the independent decentralized strategy. The number in the centralized strategy, however, is significantly different from the other two strategies. Tables 6(b) and 6(c) indicate that there is no evidence to show which strategy is better or worse than the other two strategies, in terms of vehicle revenue miles and passenger miles. The vehicle revenue miles and the vehicles used in each run among the three strategies are negatively correlated (see run 4). In the runs where the total vehicles used are equal among the three strategies, the coordinated decentralized strategy has a lower number of vehicle revenue miles but a higher number of passenger miles, as opposed to the independent decentralized strategy.

For passenger miles per vehicle revenue mile, Table 6(d) indicates that the value of the coordinated strategy is higher than that of the independent strategy; additionally, the value of the centralized strategy is higher than the same value in the independent case. However, the values in the coordinated and centralized strategies are not significantly different. On average, the coordinated strategy improves passenger miles per vehicle revenue mile by 9.0 percent, as compared with the independent strategy. The centralized strategy can improve passenger miles per vehicle revenue mile by 10.9 percent, as compared with the independent strategy. The transfer design can significantly increase the passenger miles per vehicle revenue mile due to a decrease in the empty backhaul miles (as compared with the independent decentralized strategy). The expected benefit of the transfer mechanism is a higher level of productivity, but in order to achieve this benefit, close coordination is required among the vehicles in use throughout the different zones.

To achieve a high level of service quality, it is desirable to keep wait times as small as possible. The coordinated strategy presents a slightly higher average value. Table 6(e) shows all pairwise comparisons for passenger wait times through all three strategies.

Another indicator is the scheduled travel times of each passenger (Table 6[f]). These values are significantly different for each of the three strategies. The coordinated strategy has the highest value, which is 28.1 percent higher than the value for the independent strategy and 15.8 percent higher than the value for centralized strategy, on average. This is expected because for the coordinated strategy, the inter-zonal passengers have to transfer vehicles at specific locations, which adds extra travel time. The total scheduled travel time, however, does not violate the maximum allowed ratio (travel time divided by direct travel time). Of course, inter-zonal passengers have to transfer, which could be a real burden in itself (especially for the disabled), regardless of the potential additional travel and wait times.

4.5 Summary of Results for Formulation

This chapter has presented the exact formulation of the coordinated decentralized paratransit system in order to compare its productivity and service quality with the independent decentralized and centralized strategies. The formulation has been proven to work correctly, and the results of the computational experiments of small-scale instances demonstrate that the proposed coordinated system is superior to the independent decentralized system in terms of passenger miles per vehicle revenue mile. The higher level of productivity over the coordinated strategy is achieved by a reduction of empty backhaul miles through close coordination among service providers. Based on the service quality, the proposed coordinated system increases the average scheduled travel time of passengers (as compared to the other two strategies). This outcome results from the inter-zonal passengers' increase in extra travel distances and extra service wait times for the transfer restriction. However, the maximum scheduled travel times are bounded by the acceptable maximum travel time factor. The passenger miles, vehicle revenue miles and passenger wait times do not show significant differences among the three strategies.

The exact solution approach to the proposed formulation is obviously constrained by problem scale, run time, and computer memory, given its combinatorial nature. However, results in terms of optimal solutions satisfactorily show a performance comparison between the different strategies.

CHAPTER V

HEURISTIC ALGORITHM

In this chapter, we develop a heuristic method for solving more realistic medium to large scale problems and carry out associated performance comparisons. Developed heuristics should pay particular attention to the coordination at transfer points and the arrangement of inter-zonal customers.

Due to the complexity of the formulation, we have shown that the exact method can only solve quite a small size of data, limiting the validity as it applies to the improvement of transfer design. Therefore, a heuristic is a more broadly applicable method for evaluating the performances of real-sized cases. As we mentioned earlier, developers of zoning strategies must decide how to accommodate trips that require customers to cross zones. Dealing with the inter-zonal trips, the zonal approach can be divided into two variations: (a) zoning without transfer, such as with the service provided in Los Angeles County, and (b) zoning with transfer, such as with the Chicago ADA paratransit service. In zoning without transfer, inter-zonal customers may not need to switch vehicles during their trips. Alternatively, zoning with transfer systems may require inter-zonal customers to switch vehicles.

Although the operational consolidation of providers appears to achieve economies of scale, the following may impede their coordination: (a) a user may have some concern that the current service levels will decrease; (b) the sponsoring agency may have doubts regarding whether there is a significant cost savings; and (c) the different jurisdictions within which component transportation systems operate may have different operational standards designed particularly to meet the local riders' needs (Lave and Mathias, 2000). To the best of our knowledge, there is no quantitative evidence to demonstrate the benefits and concurrent costs that occur from adopting a zoning with transfer design for a large-scale paratransit system.

In an experiment utilizing Houston's demand responsive service data, we compare the productivity and service levels among three organizational structures: zoning with transfer, zoning without transfer, and a no-zoning design. The zoning without transfer structure divides its service area into sub-zones, each zone with its own vehicle depot. The zonal service provider can only pick up customers whose pick-up location is within the service area; however, the provider is allowed to drop off customers outside of that area. Each provider is unaware of the state of the system in other zones. Alternatively, the no-zoning control system is a totally centralized system, which is the basic scenario describing paratransit services in general.

The rest of this chapter is organized into three sections. We first define the paratransit services of the zoning with transfer system, followed by a description of the demand data used. The computational experience of the algorithm is then outlined. Finally, we summarize the results of the simulation.

5.1 Transfer System

In this section we provide a description of the zoning without and zoning with transfer strategies and details of the scheduling procedure used.

Within a demand response service area, the service provider may subdivide the service area into zones. A zone is a geographical boundary. A list of customers will request a certain number of trips, defined by their location and time. In practice, each trip has a specific pair of scheduled pick-up and drop-off locations, as well as a desired pick-up (or drop-off) time for each pick-up (or drop-off) location. Each pickup and dropoff is considered a node in the system. All trips can be categorized as either inter-zonal trips or intra-zonal trips, as determined by the pick-up and drop-off locations. Trips with pick-up and drop-off locations in different zones are inter-zonal trips, and trips with pick-up and drop-off locations within the same zone are intra-zonal trips (see Figure 15).

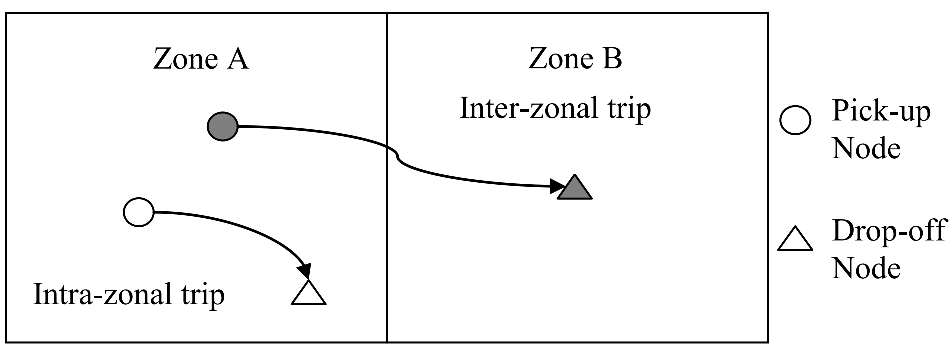


Figure 15 Categories of trip by zonal structure

For the zoning without transfer policy, zones are served and would be independently operated by different carriers. Figure 16 illustrates the characteristics of this policy. The pick-up location of each customer determines the zone and its service provider. Vehicles are, however, allowed to traverse zone boundaries in order to drop off inter-zonal customers.

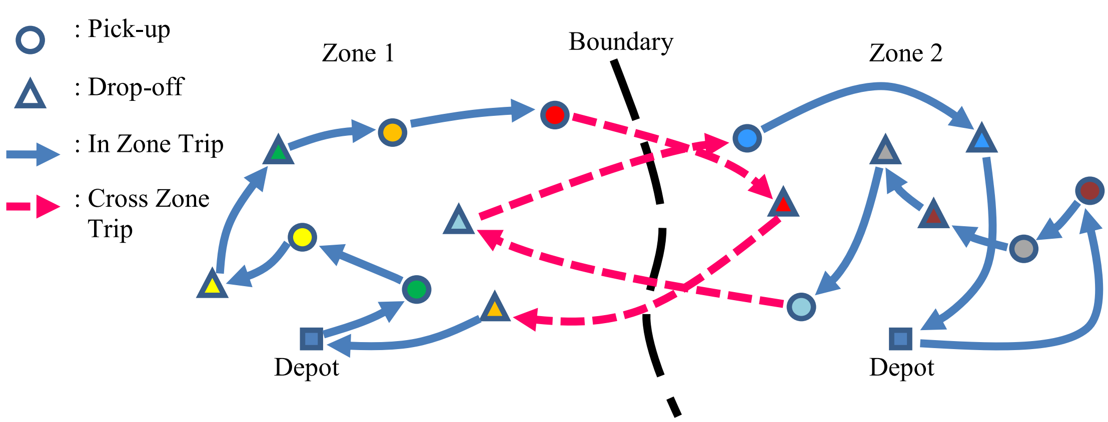


Figure 16 Zoning without transfer policy

In zoning with transfer control, inter-zonal passengers must transfer from one vehicle to another through given transfer locations in order to reach their final destination, while intra-zonal passengers do not switch vehicles to complete their trips. To highlight the loadings and unloadings at transfer locations, we generated two corresponding nodes (a load node and an unload node) at each transfer location, for each inter-zonal trip (i.e., when a vehicle visits the transfer node, it will either load or unload passengers according to the node's characteristic). Thus, the trip can be treated as two intra-zonal trips when schedules are coordinated, such that inter-zonal customers can switch vehicles at specific transfer locations (Figure 17).

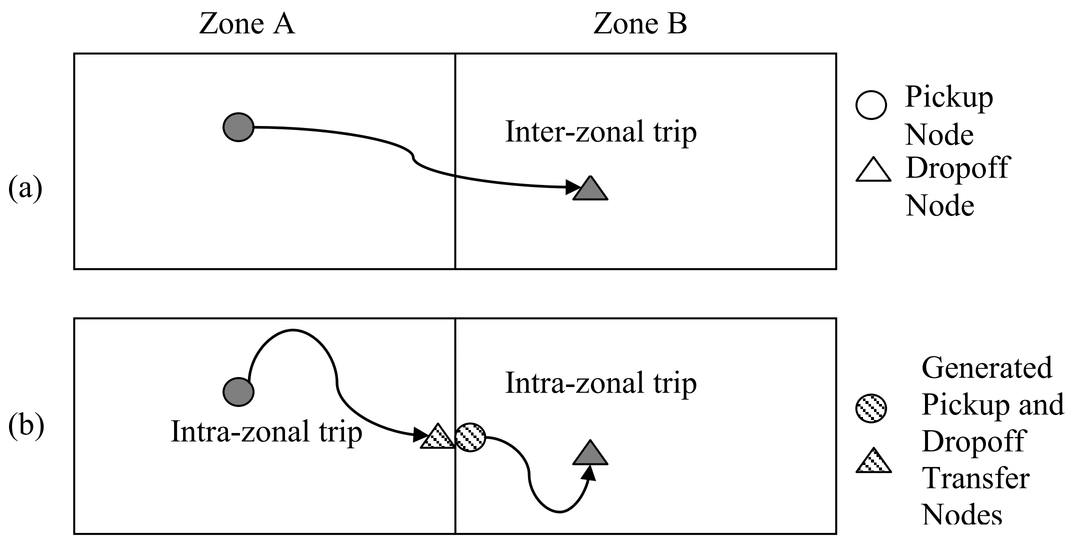


Figure 17 Example of generated intra-zonal travel trip

We assume that each inter-zonal trip allows passengers to switch vehicles only at particular transfer locations, and only once per trip. The transfer locations at which a vehicle might stop are typically on the boundaries between the various subzones. If passengers need to travel between zones that do not border one another, the transfer locations can be located within a shared buffer zone at a distance between the two zones. Concerning customer discomfort, more than one transfer might be undesirable, and in certain circumstances quite unreasonable. In practice, the passengers of the paratransit systems in both Chicago and Boston are assured, at most, one transfer (both systems use a coordinated zoning system).

For this study, we set hard time windows as follows: the earliest arrival time is ET_i and the latest departure time is LT_i $(i = 1, 2, ..., N)$, for both the pick-up and drop-off nodes. In the following context, "$+i$"("$-i$") denotes the point of pick-up (or drop-off) of customer i. The earliest vehicle arrival time is denoted as AT_i and the earliest vehicle departure time is denoted as DT_i. At pick-up nodes, the time gap between ET_i and LT_i denotes the width of a predefined pick-up time window. For example, one node may be a pick-up home address scheduled within a half hour window of time, between 6:45 a.m. and 7:15 a.m.

In many demand response scheduling systems' insertion algorithms, the objective is to minimize the vehicle travel distance while maintaining an acceptable level of service. In order to maintain such a service level, the ratio of maximum ride time (MRT_i) to direct ride time (DRT_i) needs to be within a specified value R, called a maximum ride time factor, for every customer. Therefore, the ET_{-i} and LT_{-i} of the drop-off node would be decided by the corresponding ET_{+i} and LT_{+i} of the pick-up node and R:

$$ET_{-i} = ET_{+i} + DRT_i$$

$$LT_{-i} = ET_{+i} + R \times DRT_i$$

if $LT_{-i} < LT_{+i}$, then

$$LT_{-i} = LT_{+i} + DRT_i$$

R can be a constant (such as in Los Angeles County) or an inverse function of the direct trip length (such as in Houston), in order to avoid extremely long maximum trips for already long direct journeys. Except in a case where the pick-up and drop-off vehicles arrive at the transfer location at the exactly same time, the earlier arriving vehicle must wait until another vehicle arrives (i.e., we do not allow customers to wait alone at transfer locations). We calculate the node distances based on the Manhattan distances used to calculate the symmetrical travel distances between any two pairs of nodes. The distance calculation implies that the network is arranged in a rectilinear grid pattern. We also assume that there are no traffic jams on the system, and the travel time between any two points is only a matter of the travel distance and vehicle speed. This assumption might not allow for a precise calculation of the travel time between two points, but it does not alter the results of our performance comparisons. The link distances and speeds are input into the model, and can easily be updated with more accurate values if and when those values become available.

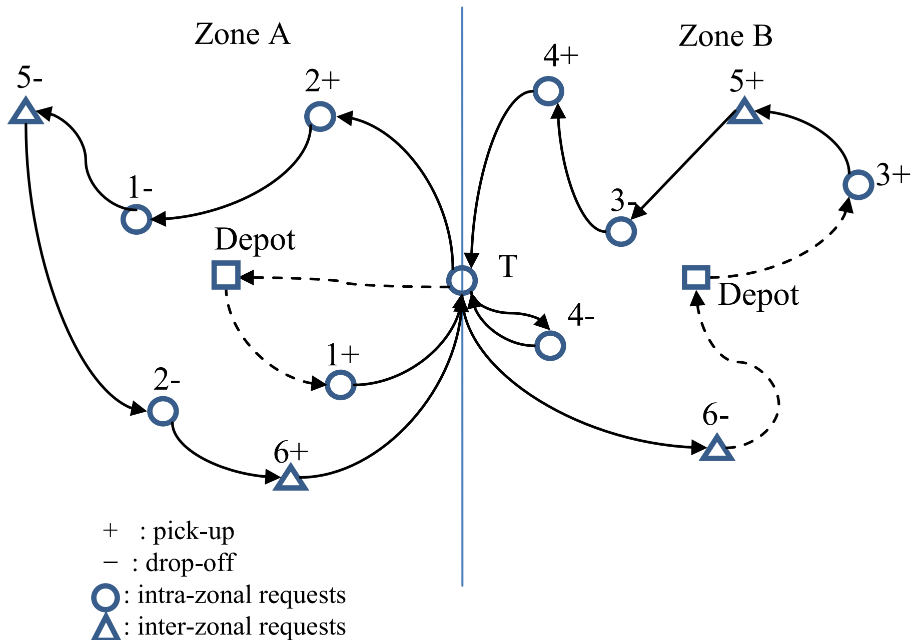
5.2 The Coordinated Zoning Dial-a-Ride Service

In this section, we introduce the idea of our proposed scheme followed by an example to illustrate the difficulties of building a feasible schedule. Our scheme combines at least two interactive paratransit systems where vehicles can only travel within their designated service areas. Basic assumptions employed in the proposed systems are also described.

Within our hypothetical demand response service region, the service provider must serve a large geographic area that may cover several neighboring cities or a single sprawling city. In response to such a problem, transit providers often subdivide service areas into component zones. A zone is a geographical boundary. It is desirable to keep vehicles and drivers near their home areas (i.e., this system will ensure that vehicles stay within their respective home zones when rides are requested). Service zones are often defined by the established boundaries of a city or other area which includes those places considered to be "local."

In our research problem, a trip is generated by a request that is composed of a pick-up and a delivery location, as well as a desired pick-up time. In other words, the customer needs to be served by a specific vehicle in order to satisfy their service need. Based on this framework, each trip can be categorized as an inter-zonal trip or an intra-zonal trip, determined by the locations of the pick-up and drop-off points. Trips with pick-up and drop-off locations in different zones are considered inter-zonal trips, while trips with pick-up and drop-off locations in the same zone are considered intra-zonal trips. Intra-zonal trips are guaranteed to be served without transfers; however, inter-zonal trips usually require a passenger to switch vehicles at some point between their origin and destination. Each inter-zonal trip asks passengers to switch vehicles at designated transfer points situated on the boundaries between the various zones. One inter-zonal trip can also be seen as two interactive trips with their schedules coordinated at the transfer locations. For each trip, customers have a desired pick-up time, and the ratio of maximum ride time to direct ride time is based on the length of direct ride time.

Figure 18 shows a simple example of a route plan with two zones and one transfer location between the zones. Requests 1 to 4 are intra-zonal trips, and requests 5 and 6 are inter-zonal trips. Requests 5 and 6 ask passengers to switch vehicles at transfer point T. For inter-zonal requests, at each transfer location we generate a drop-off node to represent the operation of disembarking customers, and a pick-up node to represent the operation of boarding customers. These nodes are defined as transfer nodes; nodes other than transfer nodes are non-transfer nodes. The synchronization between the various schedules of vehicles makes this system quite difficult to solve using traditional insertion methods. For instance, when schedules of routes between zones are synchronized, if the insertion of a customer delays the route then customers associated with that route may also be delayed. If we insert a new node, for instance node 7, before T(5-), it will not only delay the schedule of the following nodes in zone B but also delay the schedule of the nodes in zone A. This high number of interconnections means that the insertion of one customer might necessitate a recomputation for every other customer already inserted.



Route sequence
Zone A: Depot→1+→T(5+)→2+→1-→5-→2-→6+→T(6-)→Depot
Zone B: Depot→3+→5+→3-→4+→T(5-)→4-→T(6+)→6-→Depot

Figure 18 Example route plan

There are two main categories of this type of problem, based on whether the information necessary to form a solution is fully known in advance or obtained gradually over a certain span of time. In this research, we locate our problem within the first category (which is known as a static problem). In practice, the majority of paratransit services require customers to make reservations at least one day in advance. Same day reservations often can be requested, but not guaranteed. In solving the problem, we used Manhattan distances to calculate the symmetrical travel distances between any pairs of points. The result is that the calculated distance is approximately 1.4 times the Euclidean distance. We estimate the travel distance to be close to the actual travel distance established by Quadrifoglio et al. (2008). We also assume no traffic jams in the system;

the travel time between two points is figured only as a matter of travel distance and vehicle speed. This assumption does not allow us to consider the precise travel times between points, but that does not alter the results of our performance comparison (outlined below). The link distances and speeds are input into the model, and can be easily updated with more accurate values when available. We apply a drive-first waiting strategy for vehicles picking up or delivering customers. This strategy requires a vehicle to begin its drive as soon as possible; vehicles are allowed to wait at each pick-up point only when they arrive before the scheduled pick-up time. In order to provide a better level of service, inter-zonal customers remain in the vehicle to await the transfer vehicle necessary to complete their trip.

5.3 The Structure of Proposed Algorithms

In this section, we first define the systems and terminology used in our algorithm. Second, we provide a systematic description of our algorithm in a flow chart. Third, we analyze the complexity of our proposed algorithm.

5.3.1 Define the systems

The CZPS is modeled on a complete directed graph $G = (V, A)$ where V is the set of all vertices and A is the set of all arcs. For each arc (i, j), a non-negative travel time t_{ij} is considered. A total number of requests R is to be served by vehicles with a capacity of Q. At every pick-up node a number of passengers wait to be transported. Both pick-up and drop-off nodes are associated with a time window (ET_i, LT_i) where

ET_i is the earliest time and LT_i is the latest time. Each inter-zonal request $R_i = (i^+, i^-)$ may be split into two requests: $R_i' = (i^+, p^-)$ and $R_i'' = (p^+, i^-)$ where $p \in T$ and T is the set of transfer locations.

Before we introduce our proposed algorithm, we will summarize the terminology used:

Intra-zonal requests: Requests whose pick-up and delivery locations are within the same service area.

Inter-zonal requests: Requests whose pick-up and delivery locations belong to different service areas.

Transfers: Transfers are required for inter-zonal requests. Decisions about transfers are made before the insertion process. An inter-zonal request may be split into two intra-zonal requests which are synchronized to the arrival of vehicles.

Transfer locations: Transfer locations are meeting places where routes or zonal demand-responsive vehicles intersect so that passengers may transfer from one vehicle to another. Routes are often timed to facilitate transfers.

Transfer nodes: For each inter-zonal request, one must generate a drop-off node and a pick-up node at each transfer location. Such drop-off and pick-up nodes are referred to as transfer nodes.

Non-transfer nodes: Nodes other than transfer nodes are non-transfer nodes.

5.3.2 Structure of the algorithm

This section gives a general description of our insertion methodology. The new insertion-based heuristic makes use of the generic insertion framework of Solomon's sequential approach. This approach constructs one route at a time within each zone, inserting one customer at a time into the vehicle schedule until all requests are serviced. The sequential insertion implies that requests are inserted into the routes when the cheapest feasible slots are found. Therefore, the earlier build routes tend to have a higher number of requests. The sequential insertion also helps to decrease the time spent searching for feasible slots.

We sort all requests by the requested pick-up times; one empty route is then generated in each service zone. Each route starts from and ends at the depot. Intra-zonal requests search for the most feasible slots for the requested pick-up and delivery locations. Inter-zonal requests are handled by checking first requests first, and then moving on to second requests, if appropriate. In this algorithm the overall insertion procedure, from the first to the last unassigned request, is called one "round." Those requests that cannot be inserted into the schedule during a single round are copied to the list of unassigned trips. This insertion procedure requires that we maintain one route in each zone during each round.

During the search procedure, four constraints are taken into consideration. First, the arrival time, AT_i, of a vehicle at the pick-up (or drop-off) location i must be no later than $LT_{+1}\left(LT_{-i}\right)$. Second, for each request the pick-up node must be visited before the delivery node; this is also known as a precedence constraint. Third, the pick-up and delivery nodes for the same request in the same zone must be served by the same vehicle; this is also known as a pairing constraint. Finally, vehicle capacity is constrained when we search for feasible slots for unassigned trips. Figure 19 illustrates the algorithm procedure in a diagram. Detailed descriptions of the insertion and updated procedures are provided in the next section.

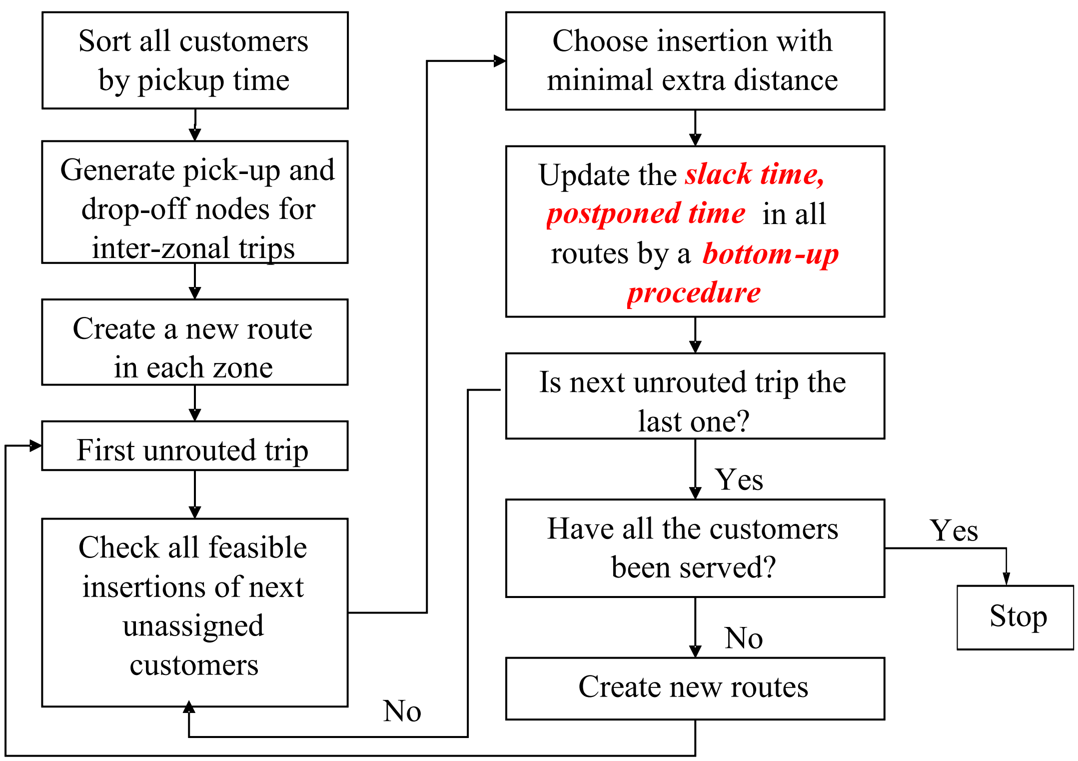


Figure 19 Flow chart of the algorithm

5.3.3 Complexity analyses

In this section, we first briefly summarize the conditions for time feasibility when inserting a customer proposed by Solomon (1987); next we provide a complexity analysis of our proposed algorithm. We assume that there are m nodes on a partially-constructed feasible route, $(i_0, i_1, i_2, \ldots, i_m)$, where i_0 and i_m are depots. If we want to insert a node, u, between the customers i_{p-1} and i_p, $1 \le p \le m$, the necessary and sufficient conditions for time feasibility are

$AT_u \le LT_u$, and

$AT_{i_r} + PF_{i_r} \le LT_{i_r}$, $p \le r \le m$ where PF_{i_r} is the push forward amount of time at node i_r and $PF_{i_r} = BT_{i_r}^{new} - BT_{i_r} \ge 0$.

Therefore, it is obvious that if $PF_{i_p} > 0$, some of the nodes $i_r, p \le r \le m$, could become unfeasible. We examined these nodes sequentially for time feasibility until we found the node where $PF_{i_r} = 0$ or $PF_{i_p} + AT_{i_p} > LT_{i_p}$. In the worst case, all the nodes $i_r, p \le r \le m$ were examined; therefore, the complexity of examining the feasibility condition is in linear time $O(m)$ for each insertion. Because the number of total possible inserted slots is $O(m)$, the total complexity of testing the insertion for a new node is $O(m^2)$. In our proposed model, the insertion of inter-zonal requests require the sequential insertion of two separate requests. We begin the procedure by inserting the pick-up location into a feasible slot, and then continue the procedure by checking all feasible slots for the delivery location of the first trip. We repeat the same procedure for the second trip. Therefore, if we use the push forward method, the complexity of the insertion for inter-zonal customers in the worst case is $O(m^8)$, which is inefficient for practical-sized problems. We propose maximal postpone time of a route location to make the feasibility check of inserting a node in constant time $O(1)$, such that the overall complexity of an inter-zonal customer insertion in routes with m nodes can be reduced to $O(m^4)$.

5.4 Computation of Postponed Time

5.4.1 Time windows for inter-zonal customers

When making reservations, customer must identify the origin and destination of the request, as well as the desired pick-up time. For request i, passengers desiring a particular pick-up location can specify the desired pick-up time ET_{i^+}, and the maximum passenger wait time WT determines the latest pick-up time LT_{i^+}. We use the following definition for the time window associated with the delivery location of request i for an intra-zonal request, where a and b are the parameters that are defined by the scheduler:

$$ET_{i^-} = ET_{i^+} + T_{i^+,i^-}$$

$$LT_{i^-} = \max\left\{ET_{i^+} + WT + b \times T_{i^+,i^-}, LT_{i^+} + T_{i^+,i^-}\right\}$$

When an inter-zonal request is split into two requests, $R_i' = (i^+, p^-)$ and $R_i'' = (p^+, i^-)$, the corresponding time windows for p^- are:

$$ET_{p^-} = ET_{i^+} + T_{i^+,p^-}$$

$$LT_{p^-} = \max\left\{ET_{i^+} + WT + b \times T_{i^+,p^-}, LT_{i^+} + T_{i^+,p^-}\right\}$$

The corresponding time windows for p^+ are:

$$ET_{p^+} = AT_{p^-} + ST_{p^-}$$

$$LT_{p^+} = ET_{p^+} + WT$$

When inserting p^+ into $R_i" = \left(p^+, i^-\right)$, the above time window setting helps to ensure that the wait time at the transfer point does not exceed the maximum allowable time.

The corresponding time windows for i^- are:

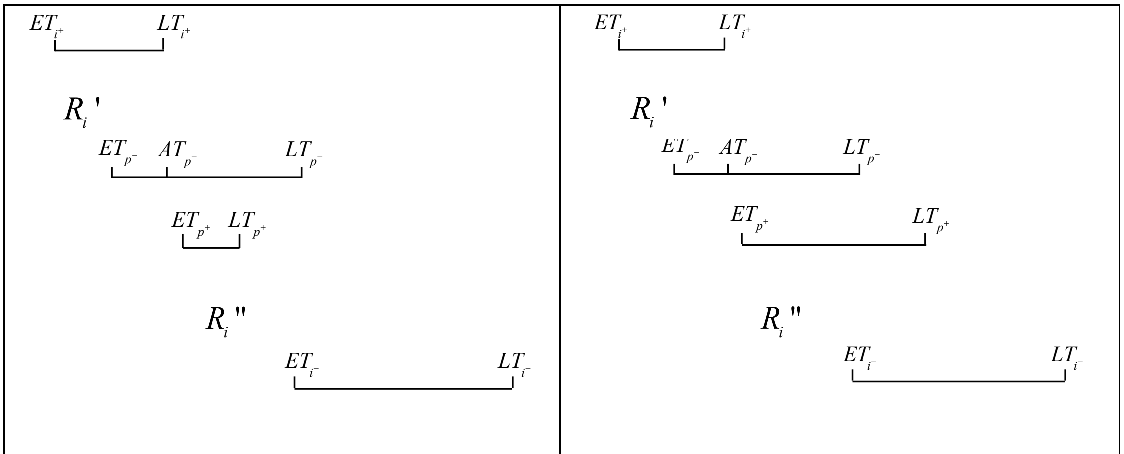
$$ET_{i^-} = ET_{p^+} + T_{p^+,i^-}$$

$$LT_{i^-} = \max\left\{ET_{p^+} + WT + b \times T_{p^+,i^-}, LT_{p^+} + T_{p^+,i^-}\right\}$$

After the insertion of an inter-zonal request, the latest pick-up time LT_{p^-} is updated by:

$$LT_{p^+} = LT_{p^-} + WT$$

In Figure 20, we illustrate how the above calculation is performed for two split but interconnected requests.



Time windows when insertion	Time windows after insertion

Figure 20 Time windows for two split requests

5.4.2 Postponed time

In our proposed zoning with transfer service, the heuristics pay particular attention to the coordination at transfer nodes for inter-zonal passengers. In addition to the four constraints mentioned earlier, vehicles with passengers onboard cannot idle at transfer nodes longer than a specific maximum vehicle idle time; vehicles without passengers onboard do not have restrictions on idle times at transfer nodes. The restriction on maximum idle time IT at a given transfer node helps to maintain an acceptable level of service for inter-zonal customers. For operators, allowing a certain amount of vehicle idle time can increase the possibility of obtaining feasible insertions, and thus increase the productivity of the service.

After inserting a new node into the vehicle schedule, the schedule of the nodes after the newly inserted node must be updated. Due to the transferring of inter-zonal customers between routes, a change in one route might make all other routes unfeasible.

This interdependence problem complicates the use of the standard insertion method. In order to ease the computational effort required to review the feasibility of inserting new nodes, we maintain two quantities for each assigned node: the vehicle wait time, WT_i, and the vehicle slack time , ST_i.

Vehicle wait time is denoted by the time difference between AT_i and ET_i if $AT_i < ET_i$. Vehicle slack time is the time difference between a maximum of $\{AT_i, ET_i\}$ and LT_i. Consider the case in which there is one node in the route, if $AT_i < ET_i$ at node i, $WT_i = ET_i - AT_i$ and $ST_i = LT_i - ET_i$. If $AT_i \geq ET_i$, $WT_i = 0$ and $ST_i = LT_i - AT_i$. Figures 21(a)–(b) offer a graphic illustration of the above two situations. At non-transfer location i, the summation of WT_i and ST_i is denoted by PT_i, the maximum postponed time, which is the maximum time interval that can be used for inserting new customers before this node. At the non-transfer node, the maximum slack time at each location is determined by the minimum of its maximum ST_i, or the PT_i of the next scheduled node. Figure 21(c) illustrates the calculation of PT_i for two consecutive non-transfer nodes, applying a bottom-up procedure. Considering the coordination of transfers, the maximum slack time of the corresponding pick-up and drop-off nodes at the transfer location will be the minimum PT_i of its following node and the slack time of the connecting vehicle at the transfer point.

We may assume, without loss of generalization, that v_i represents the node for the transfer drop-off location i, and v_j represents the node for the transfer pick-up location

j, and $AT_i < AT_j$. Except in a case where the pick-up and drop-off vehicles arrive at the transfer location at exactly the same time, the earlier arriving vehicle must wait until another vehicle arrives (i.e., we do not allow customers to wait alone at transfer locations). Therefore, at transfer node i the maximum postponed time would be the sum of WT_i, ST_i, and IT_i. After the passengers disembark, the drop-off vehicle is allowed to depart; the pick-up vehicle departs after the transfer passengers board. The above requirements can be summarized as:

$$\left|\max\left(AT_i, ET_i\right) - \max\left(AT_j, ET_j\right)\right| = IT \le IT_{\max}$$

$$DT_j = AT_j + \text{passenger boarding service time}$$

$$DT_i = AT_i + IT_i + \text{passenger disembarking service time}$$

We chose the minimum to fall between ST_i and ST_j, as the maximum slack time for both i and j. Using Figure 21(d) as an example, in the first step we calculate the ST_i of v_i and v_j using separate bottom-up updated procedures. We obtain the vehicle idle time at the transfer location as follows:

$$ST_i = \min\left\{LT_i - \max(AT_i, AT_j), PT_1\right\}$$

$$ST_j = \min\left\{LT_j - \max(AT_i, AT_j), PT_2\right\}$$

In the second step, the ST of the two transfer nodes is updated at the minimum level between ST_i and ST_j, and $ST_i = ST_j = \min(ST_i, ST_j)$.

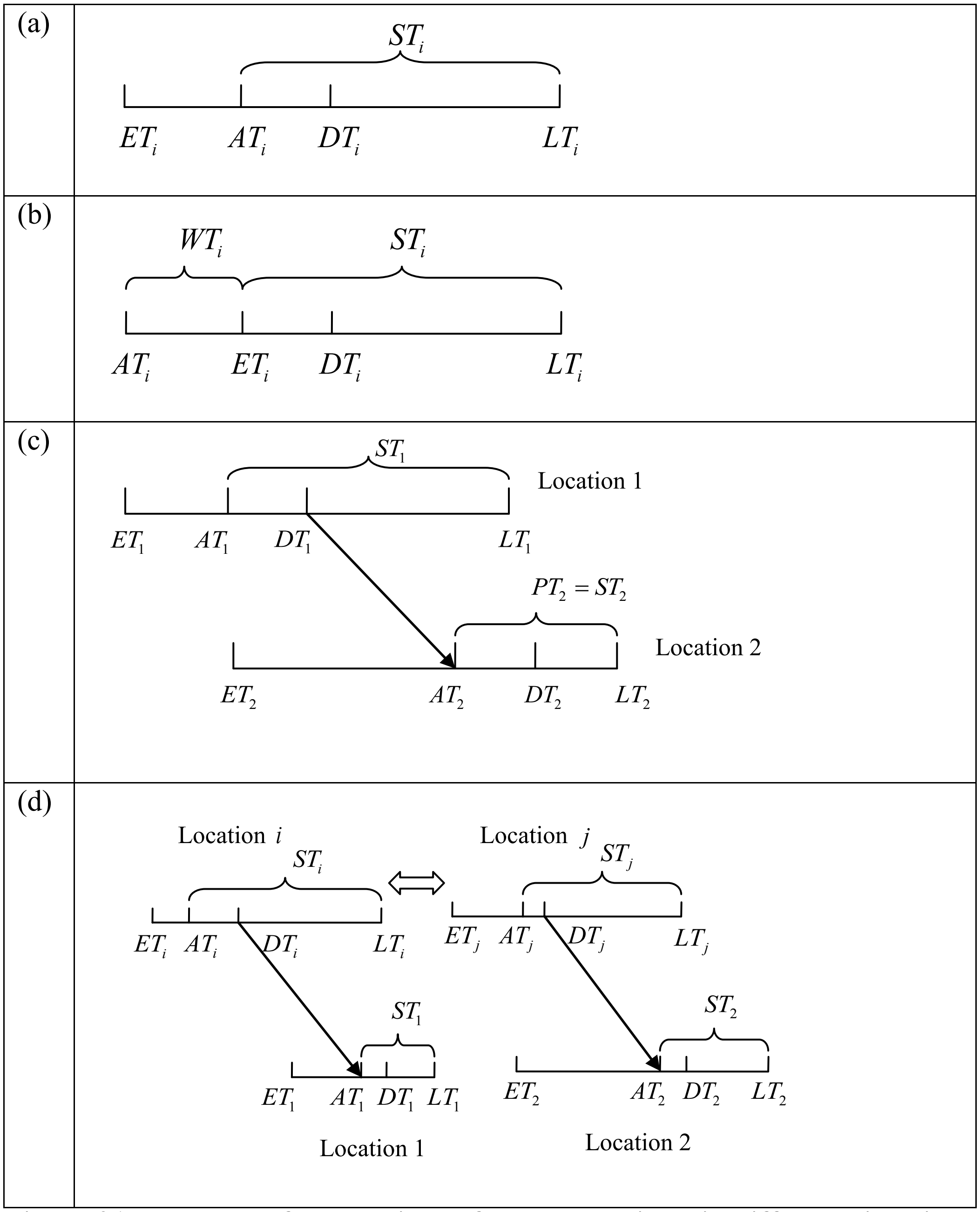


Figure 21 Examples of calculations of postponed times in different situations

5.4.3 Insertion of inter-zonal customers

In Figure 22(a), for example, the schedules of two routes are connected by the customer transfer from node A to node B. In the following procedure, we avoid situations that generate a cross-insertion between any two pairs of transfer customers in Figure 22(b). Such a cross-insertion would lead to an infinite loop when updating the arrival times of corresponding nodes. In order to make sure that the calculations of arrival and postpone times can be performed in linear time and without an infinite loop, we first prove that the cross-insertion is not feasible in our proposed method; we then give an illustration of how to update the arrival and postpone times.

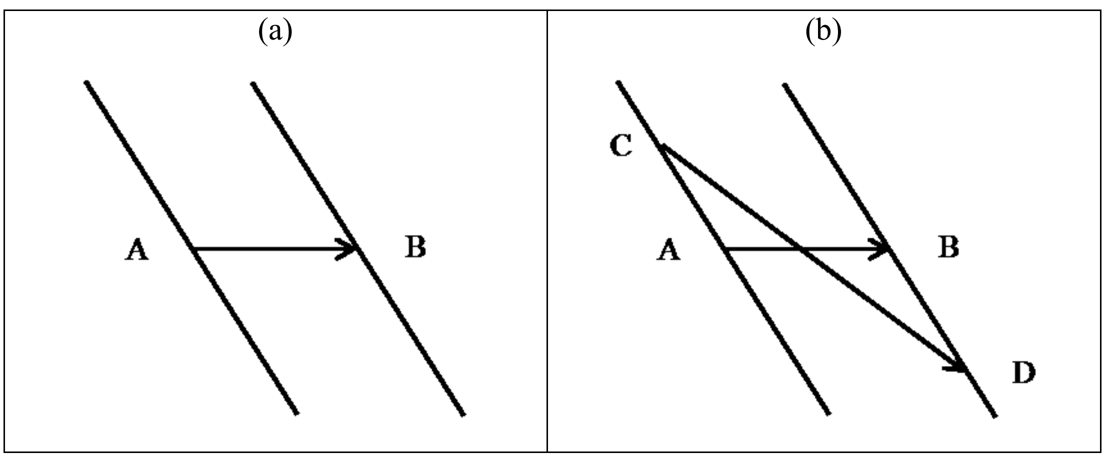


Figure 22 Cross-insertion example

First, it is necessary to introduce the definitions and observations listed below.

Definitions:

Downstream blocks of node i: The block where departure times must be later than the departure time of node i.

Upstream blocks of node i: The blocks where departure times must be earlier than the departure time of node i.

Observations:

(1) The departure times of the downstream blocks of B are always later than the departure times of nodes before both A and B.

(2) The departure times of blocks after A are not guaranteed to be later than the departure times of nodes after B, and vice versa.

Proposition: A necessary condition for feasible insertion of transfer requests is that the drop-off node and the pick-up node must both be in the blocks that do not connect nodes to the downstream or upstream blocks.

Proof:

Assume there is an insertion of a customer transferring from C to D, where D is the downstream of B and C is the upstream of A. Thus, the departure time of the nodes between C and A are necessarily later than the departure times of nodes between B and D, but this contradicts observation (1).

Therefore, in Figure 22 there is no feasible insertion that connects the downstream of node B to the upstream of node A, and vice versa. This implies that there is only one path traversing from a downstream node B to all its upstream nodes.

Figure 23 illustrates the search for feasible insertion blocks in a case with four vehicle routes. For an inter-zonal customer, we assume that the cheapest insertion has been found for the drop-off node A at a given transfer location. By applying a recursive traversal algorithm, it can be decided in linear time where the upstream and downstream blocks of node A are (indicated by solid lines in Figure 23). The allowable insertion blocks for corresponding pick-up nodes in other routes are indicated by dotted lines.

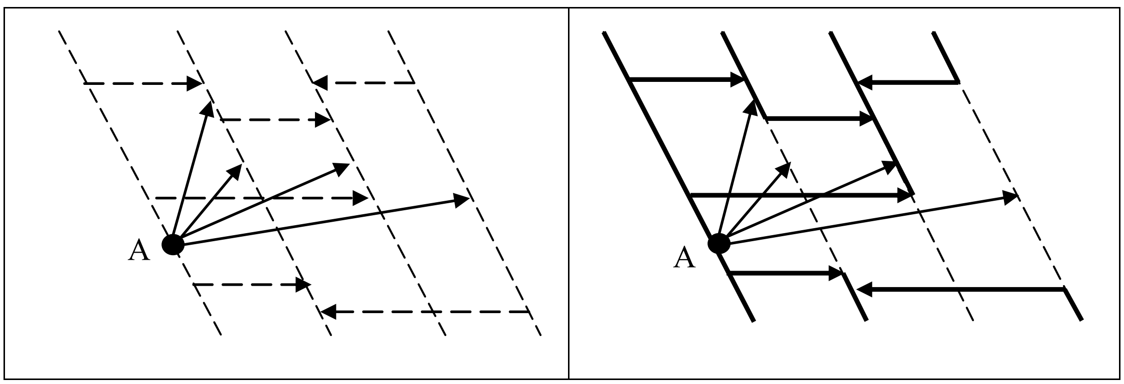


Figure 23 Search for feasible insertion blocks

5.5 Comparison vs. Optimality

We now provide an evaluation of the proposed insertion heuristic algorithm by comparing its performance against the optimality obtained in Chapter 4 by CPLEX.

Because in the formulation the objective is to minimize the number of vehicles used and the revenue miles, we might obtain a solution with a higher number of revenue miles but a smaller number of vehicles used. We discard runs with higher revenue miles in CPLEX than the heuristic when we calculate the average value. In the centralized case, the gap in vehicle revenue miles between the heuristic and the CPLEX solver is 19%, while in the coordinated decentralized case the gap is 26%.

The performance in the coordinated decentralized case seems inferior to the performance in the centralized case. This might be due to the fact that when dealing with the insertion of two consecutive requests from an inter-zonal customer, the predefined allowable pick-up time window is very narrow (10 minutes). In CPLEX, however, the time window for picking customers up at transfer points is relatively wide and thus retains the flexibility to find better solutions.

Table 7 Heuristic vs. optimality

	Coordinated Decentralized							
	Heuristic			CPLEX				
Run	Vehicles used	Vehicle revenue miles (a)	Optimal value (b)	Vehicles used	Vehicle revenue miles (a)	Optimal value (b)	Gap of (a)(%)	Gap of (b)(%)
1	4	194	2194	4	161.9	2162	0.20	0.01
2	4	237	2237	4	183.67	2184	0.29	0.02
3	6	254	3254	4	170.82	2171	0.49	0.50
4	4	232	2232	4	188.94	2189	0.23	0.02
5	6	226	3226	4	184.68	2185	0.22	0.48
6	6	258	3258	4	196.66	2197	0.31	0.48
7	4	239	2239	4	175.23	2175	0.36	0.03
8	6	226	3226	4	234.23	2234	-(0.04)	0.44
9	6	215	3215	4	194.43	2194	0.11	0.47
10	6	247	3247	4	209.5	2210	0.18	0.47
11	6	251	3251	4	170.1	2170	0.48	0.50
12	6	231	3231	4	184.47	2184	0.25	0.48
13	6	248	3248	5	203.6	2704	0.22	0.20
14	6	215	3215	4	183.07	2183	0.17	0.47
15	6	216	3216	4	199.59	2200	0.08	0.46
Avg	5.47	232.60	2966	4.07	189.39	2223	0.26	0.34
Std. Dev.	0.92	17.77	462	0.26	18.11	134	0.12	0.21

Table 7 continued

Run	Centralized							
	Heuristic			CPLEX				
	Vehicles used	Vehicle revenue miles (c)	Optimal value (d)	Vehicles used	Vehicle revenue miles (c)	Optimal value (d)	Gap of (c)(%)	Gap of (d)(%)
1	4	188	2188	3	153.87	1654	0.22	0.32
2	4	213	2213	3	217.7	1718	-(0.02)	0.29
3	4	241	2241	3	184.21	1684	0.31	0.33
4	4	210	2210	3	201.35	1701	0.04	0.30
5	5	201	2701	3	211.3	1711	-(0.05)	0.58
6	4	241	2241	4	190.81	2191	0.26	0.02
7	4	196	2196	3	197.47	1697	-(0.01)	0.29
8	5	186	2686	4	206.67	2207	-(0.10)	0.22
9	4	217	2217	4	168.41	2168	0.29	0.02
10	4	223	2223	4	190.75	2191	0.17	0.01
11	5	188	2688	4	157.1	2157	0.20	0.25
12	5	216	2716	3	183.91	1684	0.17	0.61
13	5	261	2761	4	207.86	2208	0.26	0.25
14	4	170	2170	4	153.8	2154	0.11	0.01
15	5	204	2704	4	184.4	2184	0.11	0.24
Avg	4.40	210.33	2410	3.53	187.31	1954	0.19	0.25
Std. Dev.	0.51	24.26	254	0.52	20.98	254	0.08	0.18

The performance of our insertion method is fair because this algorithm simply tries to find the cheapest feasible insertions and has not been applied to any local search method to improve the solution. However, even with the improvement procedure, it is known that the classical insertion method will encounter a 2% to 10% gap between it and the optimal solution, on average. In this research, we focus on trying to develop an algorithm which can effectively solve our proposed model in practical sizes, rather than finding the best known solutions. Furthermore, since we use the same algorithm to construct solutions for different service designs and demand configurations, the effects of gaps generated by the algorithm equivalently apply to each scenario.

5.6 Computational Experiment

In order to demonstrate the productivity and level of service provided by the proposed zoning with transfer paratransit system, we compare the results of zoning without transfer and no-zoning with the same sequential insertion algorithm proposed in the previous section. Below, we present the real demand data provided by the Metropolitan Transit Authority of Harris County, which was used to generate the random samples. Then, we describe the configurations of three organizational structures. Finally, an analysis of the simulation results is provided including the sensitivity analysis of the maximum ride time factor R.

5.6.1 Demand data description

METROLift is a paratransit service in Harris County, Texas, currently in compliance with the Americans with Disabilities Act (ADA). On average, over 5,000 trips are made daily from 3:45 a.m. to 1:30 a.m. the following day. The fare for a single ticket is $1.15 per ride. All trips must be scheduled one day in advance. Once customers make a reservation, the schedule operator gives them an estimated scheduled pick-up time. The time can change, plus or minus 20 minutes, resulting in a 40 minute time window (other US cities have adopted 20 or 30 minute windows). Comparisons to other systems have been provided in Table 8.

Table 8 Operating characteristics and populations served by different systems

	Service Area (Square Miles)	Service Area Population (Million)	Number of ADA customers	Service Hours	Boarding (Minutes)		Disembarking (Minutes)	
					Lift-Required	None-Lift Required	Lift-Required	None-Lift Required
Houston	751	3.2	17695	3:45AM-1:30AM	6	1	4	1
Chicago	3750	8	42516	24hours	7	3	6	2
Boston	729	2.5	67329	6AM-1AM	5	2	3	2
Washington, DC	1500	3.4	25575	5AM-12AM	7	2	6	2

Test samples were generated according to the locations (pick-up and drop-off) and time distributions. Using GIS software, we count the number of pick-up and drop-off locations for every square mile area (see Figure 5). The actual pick-up time distribution is shown in Figure 6. Because the pick-up and drop-off locations are independently generated, the pick-up and drop-off points are occasionally unrealistically generated within the same square mile area. In these rare cases, new drop-off locations are generated.

5.6.2 Zoning configurations

The configuration of a zoning structure is defined by its boundaries; transfer locations are often located at a zone boundary. We use the following four rules to build the sub-zones, as shown in Figure 24:

1. It is better not to situate a popular destination or high demand density area in one exclusive zone.
2. Each zone should accommodate a certain volume of trips originating from it.
3. The percentage and number of inter-zonal trips attached to each zone should be close.

4. Zones should be mutually adjacent, so that more than one transfer can be avoided.

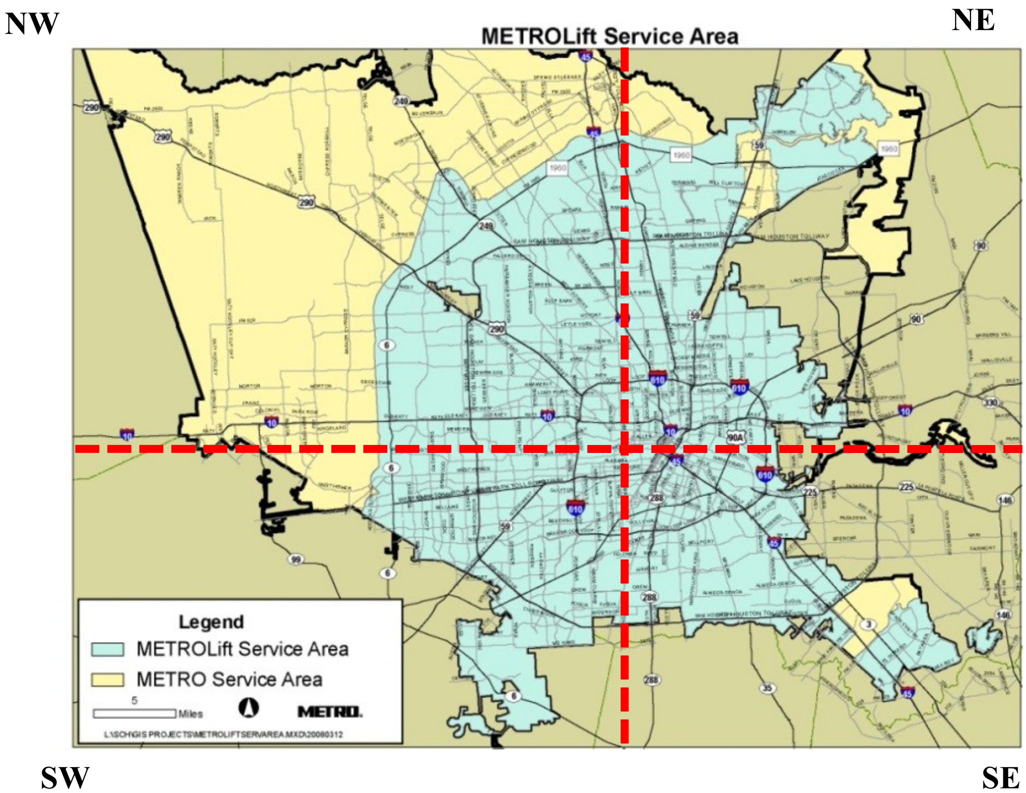


Figure 24 Zones built in Houston region

By checking the pick-up/drop-off location distribution, we locate a one square mile area with an extremely high demand density (250 pick-ups per day). This spot sits roughly in the lower center section of the service area. The origins of the trips leading to this spot and the destinations desired from this spot are both uniformly scattered throughout the area. Therefore, this area makes an idea center from which to form zones. If we include this spot in one specific zone, other zones would have had to make more inter-zonal trips, which in turn would decrease the overall service quality. Based on the selection of this spot, the service area is administratively divided into four geographical quadrants of unequal size: the northwest (NW), the northeast (NE), the southeast (SE), and the southwest (SW). Trips in each zone are observed to be large enough to maintain a minimum level of operational scale, although individual trips from each zone are not equal in length. In practice, passengers do not usually require a transfer if their destinations are just one or two blocks beyond a particular zone boundary. Therefore, we set a one mile-wide buffer area along each zone boundary.

For the zones that we generate, five locations provide for transfer needs. The center of the four quadrants is selected to be the transfer location for all inter-zonal trips traveling between NW and SE, or NE and SW. Our research found that transfer locations are best located at the edges of the zones nearest the major inter-zonal corridor.

The default parameters used in the simulation are as follows:

Vehicle speed: 25 miles per hour.

Average boarding or disembarkation time: 1 minute.

Maximum ride time factor *R*: customers have different numbers according to their direct travel distance. R =1.5 if $DRT > 72$ minutes; R =2 if 48 minutes $< DRT \leq 72$ minutes; R =2.5 if $DRT \leq 48$ minutes.

Pick-up time windows: 40 minutes from the desired pick-up time.

The three scenarios listed below were tested on the randomly generated instances, and 10 replications were run to deal with the randomness of the simulation.

1. Zoning without transfer: The region is divided into four service zones, and each zone has its own carrier. Customers are zoned by their pick-up locations and served by their designated service carrier. Vehicles in each zone cross boundaries only to drop off inter-zonal customers.

2. Zoning with transfer: This scenario respects the same geographical zones and carrier design as in scenario 1. Vehicles in this system, however, are always within a single zone. Customers need to transfer at the zone boundary.

3. No-zoning: The region is served by a single carrier. Current Houston paratransit service adheres to this scenario.

The statistics reported are the averages taken of 10 replications. The heuristic is implemented via computer program C and runs on a 2.33 GHz Core2 Duo Computer with 2 GB of memory.

5.7 Performance and Results Analysis

We investigate the performance of scenarios from the aspects of system efficiency and service quality. For system efficiency, the number of vehicles used is the most straightforward indicator for a comparison of alternative scenarios. Deadhead miles are miles a vehicle travels from its home depot to its first pick-up node, and from its last drop-off node to its home depot. Vehicle revenue miles are defined as the summation of traveled miles from first pick-up location to last drop-off location, for all vehicles.

Vehicle revenue miles with no passengers on board are defined as empty miles. Total miles include revenue miles and deadhead miles.

Passenger trips per vehicle revenue hour is an important performance measurement for capturing the productivity of a particular demand responsive system. Higher numbers of passenger trips per vehicle hour usually mean that more trips can be scheduled within a given time period.

Passenger miles traveled is calculated as the summation of traveled miles multiplied by the number of customers on board for each travel segment. Passenger miles per vehicle revenue mile is another performance measurement used to calculate the productivity of the demand responsive system. It captures the differences between the systems with longer or shorter trips, on average. Vehicle idle time is the time gap between the vehicle arrival time and the earliest pick-up time at the pick-up location.

Except in terms of efficiency, we thoroughly analyze the service quality of various different strategies. From the service quality point of view, deviation from the desired pick-up time and passenger ride time are the major passenger concerns (other than fare amount). Passenger wait time is calculated as the time difference between the requested pick-up time and the scheduled pick-up time. Passenger ride time is the actual drop-off time minus the actual pick-up time. Again, the passenger ride time cannot exceed the maximum ride time factor for both intra-zonal and inter-zonal requests.

Table 9 shows the results generated by the three test scenarios. First, we observe that the no-zone system has the smallest number of vehicles, while the zoning with transfer and zoning without transfer policies have larger numbers. This may be attributed to two reasons. The no-zoning system has no restrictions regarding a choice of the next unassigned trip; thus, the probability of finding a better insertion is higher. In addition, in favor of the sequential insertion method, the number of trips in earlier build routes is higher than in the latter build route. Therefore, if the latter build route has only one or two inter-zonal trips, it might possibly be served by one vehicle in a no-zoning system or in a zoning without transfer system; it must be served by two vehicles in a zoning with transfer case.

Table 9 Comparison for three zoning scenarios in various R

Scenario	# of vehicles	Total miles	Deadhead miles	Empty miles	Passenger miles	Passenger miles/total mile	Passenger trips/revenue hour	Vehicle Idle time	Average Passenger waiting time*	Average Passenger Ride Time*
Zoning with transfer	254	49,170	4,205	9,236	80,377	1.63	2.32	35,656	22.9	41.8
No-zoning	208	46,124	5,473	7,427	71,518	1.55	2.13	21,128	24.1	34.9
Zoning without transfer	266	48,907	5,839	16,149	71,251	1.46	1.74	25,586	22.8	34.7
* Time in minute										

By allowing a transfer for zoning policy, deadhead miles and empty miles decrease, as compared to the zoning without transfer policy. For the operator, the smaller number of empty miles provides a better result because the passenger miles per vehicle revenue mile increases as the empty miles decrease. Zoning with transfer shows a significant improvement in passenger miles over both the no-zoning and zoning without transfer policies. The higher number of passenger miles contribute to the longer trip's travel length and the higher rideshare rate. Since we use the same data set to run the simulation, we conclude that the zoning with transfer system has a higher rideshared rate. Although the zoning constrains the likelihood of finding a better insertion, we can see from the results that the transfer policy not only recovers the deficit from the no transfer case, but also significantly increases the number of rideshare rate. Due to the lowest number of total miles among the three cases, the no-zoning policy shows the highest number of passenger miles per total mile.

Zoning with transfer significantly improves the passenger trips per revenue hour. With the schedule and route coordination of inter-zonal customers at a particular transfer point, this strategy demonstrates that a zonal service that acts as a feeder and distributor increases productivity. Such a transfer policy increases the vehicle idle time, partially due to the vehicle's time spent idling at a transfer point to pick up inter-zonal customers for latter build routes. As for the level of service, the coordination at the transfer location slightly increases the passenger wait time, as compared to the wait time in a zoning without transfer system. However, the passenger wait time is still significantly lower as compared to a no-zoning case. A zoning with transfer policy shows the highest passenger ride time among the three scenarios. This is due to the inter-zonal trips requiring a change of vehicles at various transfer locations; thus, the system requires some extra travel distance and additional wait time. Passengers usually can endure longer travel times than wait times.

We further investigate the performance of the three strategies when we use a constant maximum ride time factor where R is 2.5 for all trips (as has been done in other cities such as Los Angeles). Table 10 shows the results generated by the three test scenarios.

Table 10 Comparison for three zoning scenarios in fixed R

Scenario	# of vehicles	Total miles	Deadhead miles	Empty miles	Passenger miles	Passenger miles/total mile	Passenger trips/revenue hour	Vehicle Idle time*	Average Passenger waiting time*	Average Passenger Ride Time*
Zoning with transfer	238	55,531	3,899	7,770	85,980	1.55	2.47	30,626	23.5	44.9
No-zoning	191	50,085	4,952	5,255	85,698	1.71	2.30	17,990	24.4	42.1
Zoning without transfer	240	59,909	5,219	11,443	84,875	1.42	1.91	21,190	23.0	41.6
* Time in minute										

First, in this case all scenarios have a lower number of vehicles and a higher number of passenger miles. By allowing larger maximum ride time factor for longer trips, the longer trips have larger drop-off time windows, increasing the possibility of being inserted into an existing route. The differences in the total miles between the zoning without transfer and the zoning with transfer policies are nearly equal to the sum of the differences between the deadhead miles and the empty miles. Due to its having lowest number of deadhead miles and the lowest number of empty miles among the three cases, the no-zoning policy shows the highest number of passenger miles per total mile. With constant maximum ride time factor ($R = 2.5$), it inherently favors the no-zoning and zoning without transfer policies as shown by the increase of passenger miles; however, the zoning with transfer system still has the highest number of passenger trips per revenue hour. As for the level of service, the coordination at the transfer location also increases the passenger wait time, as compared to the wait time of a no-zoning system. The cost of such higher efficiency is a decrease in the service level. All three scenarios increase passenger wait times and ride times, especially in the zoning without transfer and no-zoning policies.

The number of vehicles arriving at a specific transfer location at the same time determines the number of parking spaces required. Table 11 shows the maximum number of vehicle arrivals at the same time for each transfer location. This is the minimum number of parking spaces, then, that must be provided. However, it is desirable to add extra parking spaces to cope with unexpected situations. Furthermore, for dispatching reasons, drivers are allowed to switch to other shifts at transfer locations. Transferring passengers also enjoy a more comfortable experience if the basic facilities are provided.

Table 11 Maximum vehicles at each transfer location

	NortheWest	NorthEast	SouthWest	SouthEast
NortheWest	-	14	9	8
NorthEast	14	-	10	12
SouthWest	9	10	-	9
SouthEast	8	12	9	-

In general, our results show that the zoning with transfer design is suitable for a large service area where the majority of trips are short and there is a determinable number of long trips. Hence, productivity is retained by focusing on shorter trips within a denser area. For a small community, it is unlikely that there will be longer trip lengths to contribute to an increase in empty trip miles. As a result, transfer designs will not be able to increase productivity, and instead will only downgrade the service level. However, to what extend a transfer design might benefit by increasing the service area should be further investigated.

5.8 Applications in Different Demand Configurations

In order to understand to what extent the conclusions can be extrapolated to other cities with different configurations, we further test the proposed transfer design under the distribution of multiple different communities. To our knowledge, no test instances are described in the literature that illustrate the coordinated zoning systems discussed in this research. To analyze the benefits and costs of using coordinated strategies in zonal systems that cover several neighboring cities, we generate forty instances. In these generated instances, we use four squares to represent four separate communities within the 40 mile $\times$ 40 mile area. Two types of service area characteristics for each square are proposed: compact cases and disperse cases. The side length of each square in the compact service area is 10 miles $\times$ 10 miles; the side length of each square in the disperse service area is 15 miles $\times$ 15 miles. Origins and destinations are generated according to a uniform distribution within each square. The distributions from which the pick-up times of the samples are drawn is based on an empirical distribution derived from Houston METROLift. For both types of service areas, we generate two levels of inter-zonal request percentages: 15% and 30%. There are five instances for each scenario, and each instance has 5,000 requests.

The boarding and disembarking times for passengers with wheelchairs are set at 6 minutes and 4 minutes, respectively. The service times for other passengers are set at 1 minute. The travel times are determined on the basis of the travel distance, assuming a constant speed of travel (20 mph). The maximum idle time at transfer nodes is set to ten minutes and the maximum wait time at the pick-up nodes is set to 40 minutes.

In order to determine in what situations our methodology performs better than other possible organizational designs, we implement three different variants. In scenario 1 the vehicles are allowed to pick up and deliver any customers within the entire service region. This also can be seen as a no-zoning case. Scenario 2 applies the zonal structure, and each cluster can be seen as one zone. Vehicles can pick up requests whose pick-up nodes are within the designated pick-up zone no matter what the location of the drop-off nodes might be. Scenario 3 represents our proposed service design, as it introduces transfers for inter-zonal requests; thus, the vehicles stay within their designated zones. The algorithm is coded in C language and was executed on a personal computer.

The computational results for the sets of 15 and 30 inter-zonal percentages of a compact type are shown in Tables 11 to 14. The numbers reported in those Tables correspond to the mean values obtained from runs of five random samples, as described above. The number in parenthesis is the percentage change over scenario 1. The Tables also report additional information, as follows. The first column indicates the scenario, and the second column is the number of vehicles used. The third and fourth columns depict the total miles and the miles traveled with no passengers on board. In order to eliminate the effect of the depot location, the above two measurements exclude the miles vehicles travel to and from the depots. In practice, the scheduler does not actually consider the depot location for scheduling purposes. The reason for this is that they want the vehicles to begin making pick-ups anywhere in the service area at the beginning of the revenue hour for that route. Relatively, the vehicles can make the last dropoff anywhere before the end of the revenue service.

In the following two columns, the passenger miles per total vehicle miles and unlinked passenger trips per revenue hour are shown. In the following four columns, we report the total vehicle idle time, total vehicle wait time, the customer's average deviation from the desired pick-up time, and the customer's average ride time.

Table 12 Computational results for the compact type, 15 percent inter-zonal requests

Scenario	vehicles	Total miles	Empty miles	Passenger miles/total mile	Unlinked Passenger trips/revenue hour	Vehicle Idle time*	Vehicle wait time*	Average Passenger deviation time*	Average Passenger Ride Time*
1	262	53,842	18,639	0.94	1.53	0	19,532	15.4	31.8
2	275	50,136	15,231	1.02	1.59	0	23,128	14.3	32.3
	(4.83%)	-(6.88%)	-(18.29%)	(8.80%)	(4.00%)	-	(18.41%)	-(7.09%)	(1.61%)
3	272	47,782	11,642	1.21	1.78	2,346	28,751	16.3	38.3
	(3.68%)	-(11.26%)	-(37.54%)	(28.82%)	(16.31%)	-	(47.20%)	(5.43%)	(20.50%)
* Time in minute									

Table 13 Computational results for the compact type, 30 percent inter-zonal requests

Scenario	vehicles	Total miles	Empty miles	Passenger miles/total mile	Unlinked Passenger trips/revenue hour	Vehicle Idle time*	Vehicle wait time*	Average Passenger deviation time*	Average Passenger Ride Time*
1	303	63,053	18,864	1.09	1.33	0	20,767	14.6	42.9
2	338	64,261	21,020	1.08	1.32	0	19,956	13.9	43.3
	(11.56%)	(1.92%)	(11.43%)	-(1.17%)	-(1.23%)	-	-(3.91%)	-(5.35%)	(0.81%)
3	344	58,779	13,478	1.41	1.68	4,607	32,863	16.9	56.6
	(13.55%)	-(6.78%)	-(28.55%)	(29.19%)	(25.65%)	-	(58.25%)	(15.72%)	(31.81%)
* Time in minute									

Table 14 Computational results for the disperse type, 15 percent inter-zonal requests

Scenario	vehicles	Total miles	Empty miles	Passenger miles/total mile	Unlinked Passenger trips/revenue hour	Vehicle Idle time*	Vehicle wait time*	Average Passenger deviation time*	Average Passenger Ride Time*
1	321	66,476	21,259	0.988	1.263	0	23,076	14.91	41.06
2	345	63,894	18,620	1.040	1.259	0	31,526	14.17	41.64
	(7.69%)	-(3.88%)	-(12.41%)	(5.21%)	-(0.29%)	-	(36.61%)	-(4.98%)	(1.41%)
3	346	60,837	14,375	1.205	1.438	2,722	38,374	15.80	47.94
	(7.80%)	-(8.48%)	-(32.38%)	(21.92%)	(13.88%)	-	(66.29%)	(5.95%)	(16.76%)
* Time in minute									

Table 15 Computational results for the disperse type, 30 percent inter-zonal requests

Scenario	vehicles	Total miles	Empty miles	Passenger miles/total mile	Unlinked Passenger trips/revenue hour	Vehicle Idle time*	Vehicle wait time*	Average Passenger deviation time*	Average Passenger Ride Time*
1	353	72,996	21,327	1.084	1.161	0	24,520	14.48	49.25
2	384	73,700	22,931	1.078	1.149	0	25,026	13.62	49.62
	(8.88%)	(0.96%)	(7.52%)	-(0.52%)	-(1.01%)	-	(2.06%)	-(5.98%)	(0.76%)
3	386	68,994	15,891	1.339	1.445	4,479	38,792	16.55	62.05
	(9.55%)	-(5.48%)	-(25.49%)	(23.57%)	(24.45%)	-	(58.20%)	(14.24%)	(26.00%)
* Time in minute									

It is possible to draw several different conclusions from these results. Scenario 1 has a better performance, considering the number of vehicles used in all situations. This is especially true when the percentage of inter-zonal customers increase. Scenarios 2 and 3 have almost the same number of vehicles used. Without the restrictions on pick-up and delivery in scenario 1, the savings are shown in the number of vehicles needed. Most savings are seen under the compact type with 30 percent inter-zonal requests.

The improvement in total miles over scenario 1 mainly seems to be caused by the dramatic decrease in empty miles (miles traveled without passengers on board). The decreasing numbers of empty miles and total miles are almost equivalent. By checking the improvement rate of passenger miles per total miles against the decreasing percentage of total miles, it becomes obvious that scenario 3 improves the passenger miles significantly. The increase in passenger miles indirectly implies the increments in average vehicle occupancy (the rideshare rate) and, thus, the level of efficiency.

One of the most important productivity indexes is not linked to passenger trips per revenue hour. This index measures the service outputs that can be generated based on service consumption. It is clear that the transfer design improves productivity in all cases.

The vehicle idle time is defined as the time gaps during which the earlier arrival vehicles must wait while vehicles arriving later switch customers at transfer locations. Only scenario 3 has this value. Although vehicle wait times increase in scenario 3 between 47% and 66%, as compared to scenario 1, we find that most of the vehicle wait times occur with no passengers on board. Since we impose a maximum ride time for all requests, this level of service is warranted. The average passenger deviation time increases in scenario 3 in all situations, while the same value decreases in scenario 2 (as compared to scenario 1). Not surprisingly, scenario 3 increases the average passenger ride time. This result is mainly due to the increase in boarding and disembarking time of inter-zonal customers (especially for disabled passengers), and the increase in direct travel distance (the extra travel distance required for detours to different transfer locations). However, this level of service is warranted by the allowable maximum ride time.

We also investigate the effects of a number of transfer points on the coordinated zoning system. In one transfer location scenario, a transfer location is at the intersection of the lines connecting cluster centers and zone boundaries. Therefore, there are five transfer locations in the test instances. In two transfer location scenarios, the two locations are set at the boundaries between the zones. Table 15 depicts the computational results over the five samples in the defined scenarios.

Table 16 Computational results for the number of transfer locations

	Small type				Large type			
	15 % inter-zonal requests		30 % inter-zonal requests		15 % inter-zonal requests		30 % inter-zonal requests	
	one transfer location	two transfer locations	one transfer location	two transfer locations	one transfer location	two transfer locations	one transfer location	two transfer locations
vehicles	272	270	344	341	346	338	386	386
		-(0.86%)		-(0.68%)		-(2.31%)		-(0.09%)
Total miles	47,782	47,817	58,779	59,000	60,837	61,196	68,994	69,214
		(0.07%)		(0.37%)		(0.59%)		(0.32%)
Empty miles	11,642	11,574	13,478	13,279	14,375	14,544	15,891	15,304
		-(0.58%)		-(1.47%)		(1.17%)		-(3.69%)
Passenger miles/total mile	1.209	1.223	1.41	1.43	1.205	1.216	1.339	1.372
		(1.22%)		(1.86%)		(0.93%)		(2.47%)
Unlinked Passenger trips/revenue hour	1.780	1.781	1.68	1.69	1.438	1.441	1.445	1.444
		(0.04%)		(0.77%)		(0.18%)		-(0.04%)
Vehicle Idle time*	2,346	2,303	4,607	4,622	2,722	2,504	4,479	4,553
		-(1.85%)		(0.33%)		-(8.01%)		(1.67%)
Vehicle wait time*	28,751	28,572	32,863	30,417	38,374	37,008	38,792	38,053
		-(0.62%)		-(7.44%)		-(3.56%)		-(1.90%)
Average Passenger deviation time*	16.26	16.12	16.94	16.51	15.80	15.65	16.55	16.27
		-(0.90%)		-(2.57%)		-(0.93%)		-(1.65%)
Average Passenger Ride Time*	38.28	38.78	56.6	57.76	47.94	48.59	62.05	63.74
		(1.32%)		(2.10%)		(1.37%)		(2.73%)

Overall, the increments in the number of transfer locations do not dramatically change the service performance, despite the fact that intuitively it seems likely that the increments in the number of transfer locations would have a significant impact. On average, an increase in the number of transfer locations slightly decreases the number of vehicles needed; the increments are also shown to have an effect on the total miles, the passenger miles per total mile, and the average passenger ride time. The above results indicate that the rideshare rate increases with incremental increases in the number of transfer locations. The vehicle wait time and average passenger deviation time both decrease, when compared to the one transfer location scenario. A possible explanation for this result is that adding transfer points leads to increased opportunities to insert inter-zonal requests into the schedule, which in turn increases the total miles and decreases the total number of vehicles used. Because of the increased number of opportunities for the insertion of inter-zonal requests, the vehicle wait time and idle time both decrease.

CHAPTER VI

CONCLUSIONS

The effects of including transfers between service zones are examined in depth in the ADA paratransit system design. First, we construct a mixed-integer linear formulation to prove the potential benefit of a transfer design in a strict method. Second, we propose an insertion-based heuristic that is computationally practical to solve realistically-sized problems, thus helping to highlight the apparent benefits and associated costs of this transfer design.

The results indicate that zoning with transfer can provide significant benefits to paratransit operations that are managed according to a zoning structure. Our results use the demand data of the paratransit system operating in Houston, Texas (a relatively low-density region) for modeling and simulation purposes, and conclude that the zoning with transfer method proves to be a more productive organizational structure than a zoning without transfer method. It is worth noting that the no-zoning case adopted by Houston METRO still performs better than zoning cases, on average, in terms of efficiency. The transfer design in this research enables the system to increase the passenger trips per revenue hour significantly without excessively increasing in-vehicle ride times for passengers. Furthermore, we consider the simulations of the three zoning scenarios indicative of their relative performances, in general. Although the exact level of benefits will vary according to the different demand types and different operational standards, this simulation methodology is easily and quickly adaptable to any large-scale or rural paratransit system.

The heuristic is further tested on a series of instances built within the context of a large geographic area that covers several neighboring cities. The results of these experiments show that the introduction of transfers for inter-zonal customers decreases both the vehicle's total miles and its empty miles, thus increasing the vehicle's productivity as compared to no-zoning and zoning without transfer scenarios. However, these advantages need to be balanced by increments in vehicle wait time and average passenger ride time. When we increase the transfer locations between zones from one to two, the number of vehicles needed and the overall vehicle wait time decreases, but the total miles traveled and the average passenger ride time increase.

A natural extension of this study is to improve the solution by combining local search methods or tabu search methods. It is also be reasonable to consider the dynamic case of inserting real-time requests into the schedules based on our proposed algorithm. Future work should include combining the search of optimal transfer locations or the number of transfer locations to improve the performance of our proposed transfer system.

REFERENCES

2010 Public Transportation Fact Book, (2010) American Public Transportation Association, Washington, DC.

Aldaihani, M. and Dessouky, M.M. (2003) Hybrid scheduling methods for paratransit operations. Computers and Industrial Engineering, 45, 75-96.

Balog, J.N., Morrison, J.B., and Hood, M.M. (1997) Integration of paratransit and transit services: Importance of vehicle transfer requirements to consumers. Transportation Research Record: Journal of the Transportation Research Board, 1571, 97-105.

Berbeglia, G., Cordeau, J.F., Gribkovskaia, I., and Laporte, G. (2007) Static pickup and delivery problems: a classification scheme and survey. TOP, 15, 1–31.

Burkhardt, J.E., Hamby, B., and McGavock, A.T. (1995) TCRP Report 6: Users' Manual for Assessing Service-Delivery Systems for Rural Passenger Transportation. Transportation Research Board of the National Academies, Washington, D.C.

Chainey, S., Reid, S., and Stuart, N. (2002) When is a Hotspot a Hotspot? A Procedure for Creating Statistically Robust Hotspot Maps of Crime. Taylor & Francis, London, England.

Chira-Chavala, T. and Venter, C. (1997) Cost and productivity impacts of a “smart” paratransit system. Transportation Research Record: Journal of the Transportation Research Board, 1571, 81-87.

Cook, T.J., Lawrie, J.J., and Henry, A.J. (2003) From rural single-county to multicounty regional transit systems: Benefits of consolidation. Transportation Research Record: Journal of the Transportation Research Board, 1841, 54-61.

Cordeau, J.F. (2006) A branch-and-cut algorithm for the dial-a-ride problem. Operations Research, 54, 573-586.

Cordeau, J.F. and Laporte, G. (2003) A tabu search heuristic for the static multi-vehicle dial-a-ride problem. Transportation Research Part B: Methodological, 37, 579-594.

Cordeau, J.F. and Laporte, G. (2007) The dial-a-ride problem: models and algorithms. Annals of Operations Research, 153, 29-46.

Cortes, C.E., Matamala, M., and Contardo, C. (2010) The pickup and delivery problem with transfers: Formulation and a branch-and-cut solution method. European Journal of Operational Research, 200, 711-724.

Daganzo, C.F. (1978) An approximate analytic model of many-to-many demand responsive transportation systems. Transportation Research, 12(5), 325-333.

Diana, M. and Dessouky, M.M. (2004) A new regret insertion heuristic for solving large-scale dial-a-ride problems with time windows. Transportation Research Part B: Methodological, 38, 539-557.

Diana, M., Dessouky, M.M., and Xia, N. (2006) A model for the fleet sizing of demand responsive transportation services with time windows. Transportation Research Part B: Methodological, 40, 651-666.

Ellis, E. and McCollom, B. (2009) TCRP Report 136: Guidebook for Rural Demand-Response Transportation: Measuring, Assessing, and Improving Performance. Transportation Research Board of the National Academies, Washington, D.C.

Fu, L. (2002) A simulation model for evaluating advanced dial-a-ride paratransit systems. Transportation Research Part A: Policy and Practice, 36, 291-307.

Fu, L. (2003) Analytical model for paratransit capacity and quality-of-service analysis. Transportation Research Record: Journal of the Transportation Research Board, 1841, 81-89.

Häll, C.H., Andersson, H., Lundgren, J.T., and Värbrand, P. (2009) The Integrated Dial-a-Ride Problem. Public Transport, 1, 39-54.

Jorgensen, R., Larsen, J., and Bergvinsdottir, K.B. (2006) Solving the dial-a-ride problem using genetic algorithms. Journal of the Operational Research Society, 58, 1321-1331.

KFH Group, National Research Council (US). (2008) TCRP Report 124: Guidebook for Measuring, Assessing, and Improving Performance of Demand-Response Transportation. Transportation Research Board of the National Academies, Washington, D.C.

Lave, R. and Mathias, R. (2000) State of the Art of Paratransit. Transportation in the New Millennium, Transportation Research Board of the National Academies, Washington, D.C.

Li, H. and Lim, A. (2001) A metaheuristic for the pickup and delivery problem with time windows. Tools with Artificial Intelligence, Proceedings of the 13th International Conference on, pp. 160-167.

Li, X. and Quadrifoglio, L. (2009) Optimal zone design for feeder transit services. Transportation Research Record: Journal of the Transportation Research Board, 2111, 100-108.

Lin, C.K.Y. (2008) A cooperative strategy for a vehicle routing problem with pickup and delivery time windows. Computers and Industrial Engineering, 55, 766-782.

Lu, Q. and Dessouky, M.M. (2006) A new insertion-based construction heuristic for solving the pickup and delivery problem with time windows. European Journal of Operational Research, 175, 672-687.

Malucelli, F., Nonato, M., and Pallottino, S. (1999) Demand Adaptive Systems: Some proposals on flexible transit. Operational Research in Industry eds T.A. Ciriani, S. Gliozzi, E.L. Johnson, and R. Tadei, pp. 157-182. Purdue University Press, London, UK.

McKnight, C.E. and Pagano, A.M. (1984) Effect of size and type of organization on quality of special transportation services. Transportation Research Record: Journal of the Transportation Research Board, 973, 39-44.

Mitrovic-Minic, S. and Laporte, G. (2006) The pickup and delivery problem with time windows and transshipment. Infor-Information Systems and Operational Research, 44, 217-228.

Nanry, W.P. and Wesley Barnes, J. (2000) Solving the pickup and delivery problem with time windows using reactive tabu search. Transportation Research Part B: Methodological, 34, 107-121.

Paquette, J., Cordeau, J., and Laporte, G. (2009) Quality of service in dial-a-ride operations. Computers and Industrial Engineering, 56, 1721-1734.

Quadrifoglio, L., Dessouky, M.M., and Ordonez, F. (2008) A simulation study of demand responsive transit system design. Transportation Research Part A: Policy and Practice, 42, 718-737.

Shang, J.S. and Cuff, C.K. (1996) Multicriteria pickup and delivery problem with transfer opportunity. Computers and Industrial Engineering, 30, 631-645.

Shinoda, K., Noda, I., Ohta, M., Kumada, Y., and Nakashima, H. (2004) Is dial-a-ride bus reasonable in large scale towns? Evaluation of usability of dial-a-ride systems by simulation. International Workshop, MAMUS 2003: Multi-Agent for Mass User Support, August 10, 2003 - August 10, pp. 105-119.

Silverman, B.W. (1986) Density Estimation for Statistics and Data Analysis. Chapman & Hall/CRC.

Solomon, M.M. (1987) Algorithms for the vehicle routing and scheduling problems with time window constraints. Operations Research, 35(2), 254-265.

Wilson, N.H.M. and Hendrickson, C. (1980) Performance models of flexibly routed transportation services. Transportation Research, Part B: Methodological, 14B, 67-78.

Wilson, N.H.M. and Sussman, J. (1971) Implementation of computer algorithms for the dial-a-bus system. 39th National Meeting of ORSA/TIMS, Dallas, TX.

Wilson, N.H.M., Sussman, J.M., Goodman, L.A., and Higonnet, B.T. (1970) Simulation of a computer aided routing system (CARS). Third Conference on Applications of Simulation, pp. 171-83.

Xiang, Z., Chu, C., and Chen, H. (2008) The study of a dynamic dial-a-ride problem under time-dependent and stochastic environments. European Journal of Operational Research, 185, 534-551.

Zachariadis, E.E., Tarantilis, C.D., and Kiranoudis, C.T. (2009) A hybrid metaheuristic algorithm for the vehicle routing problem with simultaneous delivery and pick-up service. Expert Systems with Applications, 36, 1070-1081.

VITA

Name:	Chung-Wei Shen
Address:	Department of Civil Engineering c/o Luca Quadrifoglio Texas A&M University College Station, TX 77843-3136
Email Address:	tzungwei0610@yahoo.com.tw

Education:	B.B.A., Transportation and Communication Management Science, National Cheng Kung University, 2001 M.S., Civil Engineering, National Taiwan University, 2003 Ph.D., Civil Engineering, Texas A&M University, 2012