A Framework for Evaluation and Design of an Integrated Public Transport System

Carl Henrik Häll
A Framework for Evaluation and Design of an Integrated Public Transport System

Carl Henrik Häll

carha@itn.liu.se
http://www.liu.se
Department of Science and Technology

ISBN 91-85523-52-6     ISSN 0280-7971

Printed by LiUTryck, Linköping 2006
Abstract

Operators of public transport always try to make their service as attractive as possible, to as many persons as possible and in a so cost effective way as possible. One way to make the service more attractive, especially to elderly and disabled, is to offer door-to-door transportation. The cost for the local authorities to provide this service is very high and increases every year.

To better serve the needs of the population and to reduce the cost for transportation of elderly and disabled, public transportation systems are evolving towards more flexible solutions. One such flexible solution is a demand responsive service integrated with a fixed route service, together giving a form of flexible public transport system. The demand responsive service can in such a system be used to carry passengers from their origin to a transfer location to the fixed route network, and/or from the fixed route network to their destination.

This thesis concerns the development of a framework for evaluation and design of such an integrated public transport service. The framework includes a geographic information system, optimization tools and simulation tools. This framework describes how these tools can be used in combination to aid the operators in the planning process of an integrated service. The thesis also presents simulations made in order to find guidelines of how an integrated service should be designed. The guidelines are intended to help operators of public transport to implement integrated services and are found by evaluating the effects on availability, travel time, cost and other service indicators for variations in the design and structure of the service.

In a planning system for an integrated public transport service, individual journeys must in some way be scheduled. For this reason the thesis also presents an exact optimization model of how journeys should be scheduled in this kind of service.
Acknowledgement

First of all I would like to thank my two supervisors Jan Lundgren and Peter Värbrand for their support, encouragement and valuable advices. Henrik Andersson also deserves special thanks for all the interest he has shown, in a number of mathematical discussions that have helped me forward in this work. Bengt Holmberg and Yngve Westerlund have introduced me to the field of public transport and the two of them together with Mats Börjesson have all shown me different aspects and perspectives of this field. Thank you all. Thanks also to Anders Peterson, who I until recently have shared my office with, and therefore also have had many valuable discussions with. Anders Wellving has taught me many valuable tips regarding GIS, for which I am very thankful. Thanks also to Mark Horn at CSIRO who made it possible for us to use the modeling tool LITRES-2 during this thesis. Finally, thanks to all! Family, friends and colleagues.

Norrköping, May 2006
Carl Henrik Häll
# Contents

Abstract I
Acknowledgement III
Table of Contents V

List of Figures VII
List of Tables IX

1 Introduction 1
1.1 Background ............................. 1
1.2 Objectives and Contributions .................. 2
1.3 Outline .................................. 3

2 Planning of Public Transport 5
2.1 The Planning Process .......................... 5
2.2 Planning of Fixed Route Services ................ 7
2.3 Demand Responsive Services .......................... 9
2.4 Route Deviation ................................ 11
2.5 Dial-a-Ride ................................ 15
2.5.1 Simulation of Dial-a-Ride Services .................. 15
2.5.2 The Dial-a-Ride Problem (DARP) .................. 17
2.5.3 Solution Methods for DARP .................. 18

3 Planning of Integrated Services 23
3.1 Integration of Area Covering and Fixed Route Services .................. 23
3.2 Modeling of Integrated Services .................. 26
3.3 A Framework for Planning of Integrated Services .................. 30
3.4 Benefits of the GIS Module .................. 34

4 The LITRES-2 Modeling System 37
4.1 Description of LITRES-2 .................. 37
4.2 LITRES-2 Architecture .................. 38
4.3 Input to a LITRES-2 Simulation .................. 40
4.4 Output from a LITRES-2 Simulation .................. 42
4.5 The Use of LITRES-2 in Planning of Integrated Services .................. 46
4.6 Comments Regarding LITRES-2 .................. 50

5 Simulations of an Integrated Service 51
5.1 The Gävle Case .................. 51
5.1.1 Road Network .................. 51
5.1.2 Market Segments .................. 51
5.1.3 Demand .................. 52
5.1.4 Zones .................. 53
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.1.5</td>
<td>Bus Network</td>
<td>55</td>
</tr>
<tr>
<td>5.1.6</td>
<td>Meeting Points</td>
<td>57</td>
</tr>
<tr>
<td>5.2</td>
<td>Performed simulations</td>
<td>58</td>
</tr>
<tr>
<td>5.2.1</td>
<td>Number of Demand Responsive Vehicles</td>
<td>59</td>
</tr>
<tr>
<td>5.2.2</td>
<td>Capacity of the Demand Responsive Vehicles</td>
<td>60</td>
</tr>
<tr>
<td>5.2.3</td>
<td>Number of Transfer Nodes</td>
<td>61</td>
</tr>
<tr>
<td>5.2.4</td>
<td>Time Windows</td>
<td>62</td>
</tr>
<tr>
<td>5.2.5</td>
<td>Travel Factor of the Demand Responsive Service</td>
<td>63</td>
</tr>
<tr>
<td>5.2.6</td>
<td>Pricing Alternatives</td>
<td>64</td>
</tr>
<tr>
<td>5.2.7</td>
<td>Door-to-door Versus the use of Meeting Points</td>
<td>65</td>
</tr>
<tr>
<td>5.2.8</td>
<td>Demand Responsive Service without any Fixed Routes</td>
<td>66</td>
</tr>
<tr>
<td>5.3</td>
<td>Some Comments about the Results of the Simulations</td>
<td>66</td>
</tr>
<tr>
<td>6</td>
<td>An Exact Model for IDARP</td>
<td>67</td>
</tr>
<tr>
<td>6.1</td>
<td>Model Formulation</td>
<td>67</td>
</tr>
<tr>
<td>6.2</td>
<td>Strengthening the Mathematical Model</td>
<td>70</td>
</tr>
<tr>
<td>6.2.1</td>
<td>Arc Elimination</td>
<td>71</td>
</tr>
<tr>
<td>6.2.2</td>
<td>Variable Substitution and Subtour Elimination</td>
<td>72</td>
</tr>
<tr>
<td>6.3</td>
<td>An Illustrated Example</td>
<td>74</td>
</tr>
<tr>
<td>6.4</td>
<td>Some Concluding Comments about IDARP</td>
<td>79</td>
</tr>
<tr>
<td>7</td>
<td>Conclusions and Future Research</td>
<td>81</td>
</tr>
</tbody>
</table>
List of Figures

1. The planning process for different types of services .................................... 7
2. Flexibility of different demand responsive services ...................................... 10
3. Cost and level of service for different demand responsive services ............ 11
4. Different forms of route deviation ............................................................... 12
5. Benefits of an integrated service ................................................................. 24
6. Description of the integrated service ......................................................... 25
7. Different ways of traveling with the integrated service ............................... 26
8. The information flow in the framework ....................................................... 31
9. Description of the framework ...................................................................... 33
10. The system architecture of LITRES-2 ....................................................... 39
11. Visualization of one vehicle’s itinerary, with planned, ongoing and finished assignments ................................................................. 43
12. Visualization of all demand responsive vehicles within an area ................. 44
13. Visualization of accepted and denied requests .......................................... 45
14. Visualization of the planning process ....................................................... 45
15. Structure of the integrated service used in the simulations ....................... 47
16. Description of centroid zones and the simulated area .............................. 47
17. The bus lines and meeting points in the simulated area ............................ 48
18. The integrated service, as handled in LITRES-2 ...................................... 50
19. The distribution of requests over the day ............................................... 53
20. The original centroid zones ....................................................................... 54
21. Description of the final zones .................................................................... 54
22. The bus network of Gävle ......................................................................... 56
23. Selection of meeting points ....................................................................... 57
24. Input to the test case .................................................................................. 75
25. Visualization of the optimal solution to the test case .............................. 78
26. The final routes of the solution .................................................................. 78
List of Tables

1. Example of how the numerical output can be presented ....................................... 42
2. Time weights for the two market segments .............................................................. 52
3. Results from simulations of the number of demand responsive vehicles ............... 60
4. Results from simulations of the number of transfer nodes (T-nodes) ................. 62
5. Results from simulations of the travel factor of the demand responsive service ....................................................... 63
6. Results from simulations of the pricing alternatives ............................................... 64
7. Results from simulations of door-to-door service, versus the use of meeting points ........................................................................... 65
8. OD-cost matrix of the test case, created in a GIS .................................................. 76
9. Time windows for pick-up nodes $i \in P$ ................................................................ 77
10. Optimal solution to the test case ........................................................................... 77
1 Introduction

1.1 Background

The road based transportation system has become the most important part of the infrastructure in almost all developed countries. It is not only important as the physical structure of the society, but also as the foundation for social and economic development. Throughout the years, most attention has been focused on improving the traffic system on behalf of private transportation. The auto traffic is though causing problems, most of all in forms of congestions and environmental impacts.

An increased demand for personal mobility also increases the problems caused by traffic. Public transportation gets a more and more important role in reducing these problems. Increased demand for public transportation opens up for lower headways and a more effective use of the vehicles. Because of this, the level of service can improve with an increased demand, which is hardly the case for private transportation. This emphasizes that when the demand of personal transport increases, so do the importance of a well functioning public transport system.

To be a well functioning public transport system, the system must give a high level of service and be available to as many as possible. Today, there are a lot of people to whom the public transport is not available. Around one million people in Sweden have problems using the regular public transport services, due to some physical or mental impairment. Out of these, about 400 000 (4.6% of Sweden’s population) have a special needs permit, allowing them to use the Transportation of the Disabled Services, a service normally operated by taxi. A relatively small number of persons (about 25000) also have the possibility to use the National Mobility Services. This service provides people with severe disabilities the possibility to make trips all over the country at normal public transport prices. Sweden is the country in Europe with the most extensive Transportation of the Disabled Services. This is of course also quite costly. The Transportation of the Disabled Services and the National Mobility Services together cost approximately SEK 2 billion per year, out of which about 25% are paid by the travelers themselves (Finnveden (2002)).

Public transportation systems are evolving towards more flexible solutions, in order to better serve the needs of the population, to capture additional travel demands from other transportation modes and of course to increase profitability. One such flexible solution is a demand responsive service integrated with a fixed route. This type of system can use the already existing fixed route service for the major part of the journeys and thereby only needs to use the more expensive demand responsive service for shorter distances. Integrated services can be used to extend the public transportation service into low-density markets (both low-density areas, as well as to new customer segments) or can be used to substitute parts of the fixed route service. By using the right transportation mode in the right situation, an operator of public transport can in this way benefit from both the cost-efficiency of fixed route services
and the flexibility of a demand responsive service. This can reduce operating costs and increase the level of service to passengers, since a door-to-door service can be provided.

The fact that the service can be operated door-to-door enables the possibility of using this integrated public transport system also for some of the people previously directed to the Transportation of the Disabled Services. Without an integrated service, the costs for the Transportation of the Disabled Services will keep increasing. Since the average lifetime of the inhabitants in most western countries is increasing, so is the number of elderly and disabled in need of transportation, and therefore the cost of this type of service also increases. An integrated service can significantly reduce the cost for these journeys.

In a number of situations, other travelers can also have interest in an integrated service. Examples include traveling in bad weather, when carrying a lot of baggage, when there is a long way to the nearest bus stop, safety reasons (at night), and in low-density areas where no other public transport is offered.

1.2 Objectives and Contributions

In this thesis, we focus on strategic and tactical planning of an integrated public transport system. The main objective is to present a framework for evaluation and design of such a system. General guidelines of how to implement and operate an integrated public transport service shall also be given.

This thesis contributes to the area of public transport planning in the following ways. The thesis:

- gives a survey of modeling of integrated public transport services, with special emphasis on the use of optimization and simulation models.
- presents a framework for evaluation and design of an integrated public transport system. This framework consists of a geographical information system, optimization tools and simulation tools.
- evaluates how a general simulation tool for public transport systems can be applied to the analyze of an integrated public transport system.
- presents a number of guidelines to how to operate an integrated service. These guidelines have been found by simulation and includes for example suitable vehicle size of demand responsive vehicles and the number of transfer nodes to use.
- presents an exact mathematical formulation for the problem of how to assign passengers to vehicles in an integrated service.
1.3 Outline

Chapter 2 describes the planning process of a public transport system, and it is shown how optimization can be a useful tool for the different problems in the process. Different forms of demand responsive services are also described. Chapter 3 presents the design of a framework for planning of integrated services. It also presents previous work done on integrated services. Chapter 4 describes, and evaluates, a modeling tool intended to model the operation and performance of urban public transport systems, including multi-modal journeys. The architecture of the software, as well as necessary inputs and possible outputs are described. Chapter 5 describes a number of simulations intended to find guidelines to help operators of public transport when designing an integrated service. Chapter 6 describes an exact model to the problem of assigning requests to vehicles in an integrated system. This chapter further explains how the mathematical model can be strengthened and gives an illustrated example with inputs and the corresponding optimal solution. Chapter 7 presents the conclusions of this thesis and gives a discussion of future research topics.
2 Planning of Public Transport

This chapter describes how planning of public transport is performed both for fixed route services and for demand responsive services. We also discuss how operations research is of help in the planning process. Operations research uses mathematical models, statistics and algorithms to aid in decision-making. It is one form of applied mathematics, most often used to analyze complex real-world systems. The goal is generally to improve or optimize the performance of the studied system.

Section 2.1 describes the process of planning a public transport service, and the optimization problems appearing in this process. Previous work on some of these problems are described in Section 2.2. Section 2.3 explains the concepts of demand responsive services, and describes different forms of such services. Route deviation services are described in Section 2.4 and dial-a-ride services are further explained in Section 2.5.

2.1 The Planning Process

When planning a public transport system, or any other public service, the planning must be made from several aspects such as efficiency, effectiveness and equity, for example described in Savas (1978). These aspects should be put together to formulate the objective of the planning. No matter what the objective is, planning of public transport always involves a number of difficult, combinatorial problems, where operations research in general and optimization in particular, is of highest importance and can be a really useful tool. To understand the role of operations research in planning of an integrated public transport system, it is necessary to first understand the different problems involved in the process of planning public transport in general. A planning process can usually be described at strategic, tactical or operational level. In this thesis, we do not distinguish between strategic and tactical planning. The planning process will in this way only be divided into strategic planning and operational planning.

The strategic and operational planning of fixed route services can be described as a systematic decision process, first presented in Ceder & Wilson (1986). The strategic planning consists of three steps, network design, frequency setting and timetabling. The network design problem is to create an overall layout of the network in such a way that the construction/implementation costs are minimized. The problem of setting frequencies is to find the optimal frequencies in a given network. This must be made in a way so the demanded transportation volume can be satisfied. The problem has two competitive objectives, to minimize the operating costs and to minimize user inconvenience. Given decided routes and frequencies, the last problem in the strategic planning process is to create a detailed timetable. Also in this problem the objective can focus on the operator or on the customer, e.g. minimize the number of vehicles or minimize the transfer times.
In the strategic planning it is essential to have good background information about the travel demand in the area. For this purpose, OD-matrices are used. Each entry in such a matrix describes the number of passengers wanting to travel between a given origin and a given destination (points or zones) in the network during a given time period. All steps in the strategic planning process are based on the OD-matrices. Because of this, it is very important that the OD-matrices contain as accurate data as possible.

In the operative planning, the timetables are the basis from which vehicle schedules and crew schedules are created. Except from these scheduling problems the operative planning also includes a number of ”what-if problems”. This is a class of problems that always can occur and that demand fast solutions. Examples of this kind of problem are, what shall be done if; a vehicle breaks down or if a driver calls in sick? A more detailed flowchart of the whole planning process, including input and output corresponding to the different steps, can be found in Ceder (2003a). For demand responsive services, further described in Chapter 2.3, there is also the additional problem of allocating passenger requests to vehicles.

In an integrated service (explained in Chapter 3) combined by different modes, one must also consider the problem that different modes and/or vehicles can perform different parts of the same journey. So when planning an integrated service it is important to consider both strategic problems as well as operational problems already during the design of the service. To be able to do this it is important that the behavior of the customers and the response by the service operator can be predicted or simulated. The planning process for fixed route services, demand responsive services and integrated services are described in Figure 1.
2.2 Planning of Fixed Route Services

To be able to start the planning process, demand data must be accessible. The problem of estimating OD-matrices for transit networks is described in Wong & Tong (1998) and Wong & Tong (2003). The data to the OD-matrices are many times gathered through on-board measurements instead of through estimation. Demand data is used in both the network design as well as in the timetabling of the service. The estimation, or gathering, of demand data is therefore very important. With the demand known, the problem of network design can be addressed.

The network design affect frequency setting as well as vehicle and crew scheduling. Because of this, the design step is very important. The problem known as the ”Transit Route Network Design Problem” often includes both the actual network design problem as well as the frequency setting. The work of Fan & Machemehl (2004a) gives a very good overview of this problem. In Shih et al. (1998), also suitable vehicle sizes at the different routes are considered.

Regarding heuristics used for transit route network design, Pattnaik et al. (1998), Bielli et al. (1998) and Chakroborty (2003) all used genetic algorithms. In Fan & Machemehl (2004b) a tabu search heuristic is compared to a genetic algorithm and shown to outperform the genetic algorithm on the example network used. Still, in Fan & Machemehl (2006) a genetic algorithm approach is used to further study the characteristics of this problem, but now with variable demand.

In Borndörfer et al. (2004), the problem of setting frequencies are addressed through the use of two multi-commodity flow models, both with the objective to minimize
Planning of Public Transport

a combination of operating costs and passenger traveling time. Guan et al. (2004) handles the configuration of a pre designed network and the passenger assignment problem at the same time, formulated as one single model.

How the timetables are planned, highly affects the customers. This is an area where a lot of work has been done. Many times the objective is to minimize the total waiting time or the sum of the waiting time and the ride time. In Domschke (1989), the objective is to find departure times at the terminal stations of all routes (busses and trains) so that the total waiting times at transfer nodes for all passengers is minimized. The work contains heuristics, regret-methods with improvement algorithms and simulated annealing, as well as a branch and bound algorithm.

Desilet & Rousseau (1992) describes a software designed for the synchronization of transfers. The software uses a model that from a set of possible starting times for different routes chooses the best one, with the objective to minimize the total penalty associated with transfers. Also Liu & Wirasinghe (2001) describes a simulation model intended to design schedules for a fixed route bus service. The model determines which stops that shall become time points, fixed in the timetable, and how much slack time every time point should have allocated. Such a simulation tool can of course be of great help when designing schedules. Near optimal solutions are found in the given examples.

In Ceder et al. (2001) the problem of creating a timetable with maximum synchronization given a network of a fixed route bus service is addressed. This is done with the object to maximize the number of buses simultaneously arriving at the transfer nodes of the network. The problem is formulated as a mixed integer linear programming problem and a heuristic is used to solve the problem in polynomial time. The problem of synchronization is also addressed in the work of Fleurent et al. (2004) and in Voß (1992) that used a quadratic assignment problem to model the minimization of passengers waiting times at transfer nodes.

In Hagani & Banihashemi (2002), an exact formulation as well as two heuristic approaches to the "multiple depot vehicle scheduling problem with route time constraints" is presented. This problem is to create schedules for individual vehicles, busses, belonging to different operators and stationed at different depots. To add the route time constraints is to say that a specific vehicle must return to the depot within a given time. This can for example be due to fuel consumption or working hours of the driver. Adding this constraint for each vehicle reduces the size of the problem significantly. In Ceder (2003b) both timetabling and vehicle scheduling are handled.

This section has shortly described what has been done on the different problems in the strategic planning of a fixed route service. As can be seen from this literature review, optimization can be used to address all the problems in the strategic planning process of a fixed route service.
2.3 Demand Responsive Services

Fixed route services are not always enough to satisfy the needs of all customers who want to use the public transport service. Demand responsive services are therefore often a necessity to satisfy all the demand. The term ”demand responsive service” is a label used for many different services. According to the definition of Kirby et al. (1974) a demand responsive transit service is a service that ”provides door-to-door service on demand to a number of travelers with different origins and destinations”. The door-to-door part of this definition is usually not so strictly followed. Many services do not pick up and drop off passengers at exact addresses, but the service still respond to a certain demand at a specific time. A better description of the service is that it offers flexible routes and schedules and that it at least partially responds to requests from passengers. Demand responsive services are meant to fill the gap between fixed route mass transit and ordinary taxi service, both in terms of flexibility as well as in terms of cost.

In this section, different forms of demand responsive services will be described. Some of the forms of demand responsive services are similar to a fixed route service, while others are more of area covering services. The main types are:

- **Hail-a-Ride** is a fixed route service in local areas. The routes are normally highly frequented and provide access to healthcare, schools, shopping centers etc. Passengers can be picked up or dropped off anywhere along the route, i.e. embarking and disembarking is allowed anywhere along the route. Since this is still a fixed route service, it is the least flexible form of demand responsive services.

- **Route deviation** is a form of demand responsive service with low flexibility. The service is normally used in low-density areas, which also implies a low frequency of the service. The passengers, who are not able to reach the ordinary bus stops, can call in a request in advance so that the bus driver becomes aware of that a passenger wishes to be picked-up on the deviation part of the route. Unless a request has been called in, the bus travels the ordinary way, without the deviation.

- **Dial-a-Ride** is a type of service in which passengers can (and in most cases should) call in requests in advance. Normally this service operates between two scheduled stops, and the most common is to let the vehicle travel freely within a corridor between the two stops, rather than along a fixed route. In some cases of this service hails enroute are also responded. Often, dial-a-ride services are also operated as a multihire taxi, but serving a predefined area.

- **Multihire taxi** operates as a normal taxi, but the vehicle can be shared with other passengers traveling between other origins and destinations. This is the most flexible form of demand responsive service.
The route flexibility and timetable flexibility for the different forms of demand responsive services are presented in Figure 2. It can be seen that the two flexibilities are correlated. The relation between cost and level of service is described in the same way, see Figure 3.

![Figure 2: Flexibility of different demand responsive services](image)

It should be noticed that depending on how the service ”Route deviation” is operated (later explained in Figure 4) the flexibilities, level of service and cost of the service can vary a lot. Because of this, ”Route Deviation” can be placed anywhere from the fixed route service up to the truly flexible services, depending on how many (and what kind of) deviations the service allows.
As can be seen in the descriptions of the demand responsive services, multihire taxi and dial-a-ride are in many ways alike. This indicate that the operational planning of a multihire taxi service is similar to that of planning a dial-a-ride service, and can regarding the planning be seen as a form of dial-a-ride. For this reason applications and solution methods for both dial-a-ride as well as multihire taxi will now on be referred to as dial-a-ride.

The main idea of an integrated service is to reach very close to customers’ origins and destinations (preferably door-to-door) and the importance of an area covering demand responsive service can not be emphasized enough. Therefore methods for dial-a-ride and route deviation, are the most interesting to use in an integrated service and modeling and solution methods for these services are therefore discussed more detailed in Section 2.5 respectively Section 2.4. Hail-a-ride is too much like a fixed route service to be really interesting to use in an integrated service, and will therefore not be further studied. Integrated services are explained in Chapter 3.

2.4 Route Deviation

The most commonly used form of route deviation can be described as having a fixed route, with fixed timetable. On this fixed route there is only a few extra stops, many times only one, that will be visited only if someone has requested to be picked up or dropped off at that location. This kind of route is illustrated in Figure 4(a). With this kind of service a request on an extra stop is always accepted.
The extra time needed for the deviation part of the route must always be included in the total time scheduled for the route. This means that if there is no demand on an extra stop, all the scheduled time will not be needed. Vehicles and drivers are thereby not in use the full cycle time.

Most of the recent interest in route deviation has been focused on services with quite sparse compulsory stops, and more than one extra stop between each pair of sequential compulsory stops, as illustrated in Figure 4(b). This brings the service closer to that of a dial-a-ride service, and at the same time the problem areas and solution methods also become more like those of interest when dealing with a dial-a-ride service. Examples of research within this area are the work of Quadrifoglio et al. (2006b), Smith et al. (2003) and Malucelli et al. (1999).

In many ways this description resembles the "Flexroute" operating in Gothenburg Sweden, described in Westerlund et al. (1999), with the difference that the Flexroute only operates between two compulsory stops, and then uses a large number of "meeting points" (checkpoints) that can be visited upon demand. The Flexroute can therefore be seen as more of a Dial-a-Ride service. It is also very much like the system analyzed in Daganzo (1984), where the feasibility of a checkpoint dial-a-ride system is studied. Cost-effectiveness is compared both to fixed route services and to a regular dial-a-ride service operating door-to-door. A model presented in Pratelli & Schoen (2001) is formulated to choose a certain number of demand points that minimizes the total disadvantage experienced by the other passengers.

In Quadrifoglio et al. (2006b), an insertion heuristic for scheduling a route deviation service called "Mobility Allowance Shuttle Transit" is presented. A set of simulations, with different demand levels, is carried out to describe the behavior of the algorithm. Different performance parameters are formulated to evaluate the efficiency. The same type of service is studied in Quadrifoglio et al. (2006c), where the relation between the width of the service area and the longitudinal velocity is
in focus, and bounds on the maximum longitudinal velocity are presented. The longitudinal velocity is of great importance for any route deviation service, since the main objective of such a service is to transport customers along a given direction. In Quadrifoglio & Dessouky (2004) the performance of a route deviation service is compared through simulation to that of a fixed route service. The results show that under certain demand distributions, the route deviation service performs better.

Two design parameters of highest interest when planning a service based on route deviation are studied in Smith et al. (2003). These parameters are service zone size and slack time distribution. The service zone is the area between two compulsory stops in which requested deviations can be serviced. A maximum distance away from the ordinary road between the two stops limits such zones. The slack time is the extra time needed to be built into the schedule to make the deviations possible.

To simplify the planning process, the effects of these two parameters have been evaluated through the use of a multi-objective binary optimization model. The two objectives used were to maximize the number of feasible deviations per hour, and to minimize the total unused slack time. By using these, both the operator’s and the customers’ perspectives are taken into account. The operator wishes to serve as many customers as possible (make as many deviations as possible), and the customers do not want any unused slack time in the schedule, since this only renders unnecessary waiting times. To determine the best distribution of the slack time two different methods have been used. These are a weighted average of the none stop travel time between the two fixed stops of the zone and a weighted average of the total number of origins and destinations of trips by customers with special needs permits. The model can be used also with more alternative methods.

In Malucelli et al. (1999), a transportation system called "Demand Adaptive System" is presented. The system consists of a set of lines, described by a set of timetabled trips, in other words as a normal fixed route service. The stops included in the original timetable are compulsory stops. The flexibility of the system consists of that the vehicles are allowed to transit by each compulsory stop \( h \) during a specified time-window \([a_h, b_h]\). Between every pair of compulsory stops there are a number of optional stops that can be activated by a user. If a user wants to be picked up or dropped off at an optional stop, the user must send a request specifying a stop where to be picked up and a stop where to be dropped off. The extra time, that the time-window admits, is used to visit a number of such optional stops, if any request for these has been made. If a user wants to travel between two compulsory stops, no request shall be sent. In this case the service can be used as an ordinary fixed route service, with the exception that the user only knows that the vehicle will depart within the specified time-windows.

All activated optional stops between two successive compulsory stops must be visited between that the compulsory stops are visited. Therefore optional stops will be visited at a time later than the earliest time the vehicle can leave the preceding com-
pulsory stop and no later than the latest time it can leave the succeeding compulsory stop. An optional stop between the compulsory stops $h$ and $h+1$, must therefore be visited between $[a_h, b_{h+1}]$.

Three variants of the Demand Adaptive System, depending on how the requests are handled, are presented.

- Users are picked up and dropped off at the stops that they have requested. If the acceptance of a request cause infeasibility or is not economically worthy, the request can be rejected.

- The user is picked up at the requested stop, and dropped off in the neighborhood of the requested drop-off stop, if the stop itself can not be part of the vehicle’s itinerary. For this, the user may travel at reduced cost.

- The user is picked up and dropped off in the neighborhood of the requested stops, if the stops can not be part of the vehicle’s itinerary. Also here a discount of the fare price is applied.

All these variants of the system are formulated as Mixed Integer Linear Programming problems, and heuristic procedures for their solutions are provided. How these variants changes under dynamic conditions, that is where requests can arrive while the vehicle is operating, are also discussed.

A system where each transit vehicle operates with route deviations in a predefined zone, but operates between zones as fixed route vehicles, are studied in Cortés & Jayakrishnan (2002). This system gives the possibility of travel between any two points with only one transit between vehicles. Systems like these, but where the fixed route part and the part where deviations are allowed are performed by different vehicles, and in this way often require two transits, are presented Chapter 3.

This literature review of route deviation systems clearly shows the interest of more and more advanced deviation strategies. The original (and still most commonly used in practice), where only one optional stop is used on an otherwise fixed route is not so interesting to study from a mathematical point of view. The new, more advanced, form of deviation services offers a more flexible service. This form of deviations indicates that the service can be useful in an integrated service. It also shows that the boundary between route deviation services and dial-a-ride services not always are clear.
2.5 Dial-a-Ride

Different kinds of dial-a-ride systems have recently gained a large interest in planning of public transportation systems, mainly because it provides suitable transportation for elderly and disabled. Two types of questions are most frequently appearing in work done on dial-a-ride systems. How shall the system be designed to be as effective as possible, and in what situations are a dial-a-ride system a good solution?

Applications with dial-a-ride services are limited to small-scale cases. Two things have contributed to this. First of all, it is not clear how the usability of a dial-a-ride system changes with an increase of the number of passengers, in comparison to how a fixed route system behaves in the same situation. Secondly, the mathematical problem of assigning passengers’ travel-requests to vehicles in an optimal way is a very difficult problem. The problem can be shown to be a variant of the vehicle routing problem (Li & Lim (2001), Luis et al. (1999)), introduced in Dantzig & Ramser (1959). This problem is a variant of the well-known traveling-salesman problem. A lot of studies of how to assign passengers to vehicles have been made, and resulted in a number of different optimization solution techniques. Simulation studies of dial-a-ride services are presented in Section 2.5.1. Section 2.5.2 gives a general description of the problem known as the Dial-a-Ride Problem (DARP). Work done on solving this problem will be presented in Section 2.5.3.

2.5.1 Simulation of Dial-a-Ride Services

When designing a dial-a-ride service it is important to know how different factors and routing policies affect the service. The first, and still most common, approach to study such effects is by simulation. Simulation studies of many-to-many dial-a-ride systems were for example studied already in, Heathington et al. (1968), Wilson et al. (1969) and Gerrard (1974). The work of Wilson & Hendrickson (1980) reviews earlier models used to predict the performance of dial-a-ride services. Both simulation models as well as deterministic and stochastic approaches are discussed.

If one wishes to make fixed route passengers more interested in using a dial-a-ride service, important knowledge to have is under what conditions a dial-a-ride service can be a better alternative than a fixed route service. The changes of usability depending on the number of passengers have only been studied in a few papers. Good examples of such studies are Bailey & Clark (1987) and Noda et al. (2003). In Bailey & Clark (1987), interaction between demand, service rate and policy alternatives for a taxi service were studied. Noda et al. (2003) the usability of dial-a-ride systems and fixed route systems are compared through a transportation simulation of a virtual town. The aims of the simulation was to compare the usability and profitability of dial-a-ride systems to that of fixed route systems, where usability is defined as the average time between that the request is sent until the request has been carried out, and profitability is defined as the number of requests occurring in a time period.
per bus. To make the comparisons as equitable as possible it is assumed that the requests are sent the same time the passenger wishes to commence the journey, because in the fixed route case the passenger simply goes to the bus stop. Regarding the vehicle routing policies, the vehicles are allowed to move freely within a given area, since this is the most general and the most important factor of a dial-a-ride service.

The results of Noda et al. (2003) states that if the number of vehicles remains unchanged, the usability of dial-a-ride systems degrades very quickly as the number of requests increases, since many requests then are denied. If the number of buses increases while keeping the ratio of requests and vehicles fixed, that is that the number of vehicles increase with fixed profitability, usability of the dial-a-ride system improves faster than for the fixed route system. This actually means that for a given number of requests per vehicle, when the total number of requests increases and the number of vehicles do as well, not surprisingly so does the usability. This is due to the advantage of having more possible combinations of vehicle itineraries.

Simulations to study the effects of a dial-a-ride service are also done in Quadri-foglio et al. (2006a). How time window settings and zoning vs. no-zoning strategies affect the total trip time, deadhead miles and fleet size are studied. The fleet size is also studied in Diana et al. (2006). A continuous approximation model is used instead of simulation to determine the number of vehicles needed to give a predefined quality of the service.

In Deflorio et al. (2002) a simulation system is proposed that is able to evaluate quality and efficiency parameters of a dial-a-ride service. The system can simulate a number of uncertainties caused both by passengers and drivers. An other simulation system is described in Fu (2002b). The purpose of this system is to evaluate what effects new technologies such as automatic vehicle location can have on a dial-a-ride service. The work of Jayakrishnan et al. (2003), gives a more general discussion about the needs of a simulation system intended to simulate different commercial fleets and different types of vehicles and services, such as dial-a-ride.

Despite that all of these aspects are very important to consider when planning a dial-a-ride service (deciding if dial-a-ride is the proper service form for the intended area), most work have been done in order to find methods, or algorithms, for routing of vehicles within such services. To be able to perform simulations of the kind described above, there is of course a need of an algorithm describing the process of how requested journeys are being assigned to the different vehicles. This problem, to plan how the requests shall be scheduled to the vehicles, is known as the dial-a-ride problem (DARP) and will be explained in the next section.
2.5.2 The Dial-a-Ride Problem (DARP)

The dial-a-ride problem (DARP) is a specific case of the pick-up and delivery problem. Travel requests belonging to individual passengers or groups of passengers are to be executed. To each request there is a specific origin and destination defined. The most common application of this problem type is within transportation of elderly and disabled. In such applications each request is to be carried out from one address to another, i.e. that the transportation service is of a door-to-door type. What really characterize the DARP from any other pick-up and delivery problem is the control of user inconvenience. User inconvenience can be stated as waiting time, travel time or deviations from desired departure and arrival times. This is to reflect the necessity of balancing user inconvenience against minimizing the operating costs, when transporting passengers.

In practice, dial-a-ride services can be operated according to one of two modes, static or dynamic. The static mode is when all requests are known in advance, which also allows vehicle itineraries to be planned in advance. Static versions of the DARP are for instance described in Feuerstein & Stougie (2001) and Melachrinoudis et al. (2006). Opposite to this is the dynamic mode, for example described in Teodorovich & Radivojevic (2000), Colorni & Righini (2001) and Coslovich et al. (2006). In the dynamic mode the number of requests gradually increases as the customers call in requests and the planning starts before all requests are known. Most studies on the DARP assume the static case. Often assumed is also a homogenous vehicle fleet based at one single depot. Important to remember is however that this is not always the case in practice. Several depots, as well as different vehicle types, for example some equipped to handle wheelchairs, are of course common in practice. Also when working with the dynamic case, the static case is often solved on a known set of initial requests, since some requests usually are known prior to scheduling and therefore can be used to find a starting solution.

There are normally two objectives from the operators’ point of view as well as two from the customers’ point of view that can be part of the overall objective of a DARP. Out of the operators’ perspective, the goal is to minimize the total number of vehicles needed as well as the total travel time of those vehicles. From the customers’ perspective, the goal is to minimize service time deviations and minimize the ride times, Fu & Teply (1999). More detailed explanations of the DARP can be found in Cordeau & Laporte (2003a) and in Cordeau et al. (2004).
2.5.3 Solution Methods for DARP

Some early work on DARP are those of Wilson et al. (1971), Stein (1978), Psarfatis (1980), Psaraftis (1983a) and Psaraftis (1983b). Wilson et al. (1971) investigate the dynamic DARP. Stein (1978) and Psarfatis (1980), handles both the static case where all requests are received in advance, as well as in the dynamic case where requests can occur at any time. In Psarfatis (1980) the single-vehicle, many-to-many problem is investigated. Many-to-many implies that the customers all can have different origins and destinations. The objective function is to minimize a weighted combination of the total time needed to service all customers, and the total inconvenience of those who have to wait for service. This is done with respect to constraints regarding vehicle capacity and priority rules. If only the first part of the objective would be considered, the objective would be the same as that of the traveling salesman problem. The single vehicle dial-a-ride problem is studied also in Psaraftis (1983a) and Psaraftis (1983b). The work of Psaraftis (1983b) also extended the problem to include time windows.

Most of these papers are focused from the operators’ perspective, trying to minimize the total distance of the vehicles. This is the objective also in Desrosiers et al. (1986) where a forward dynamic programming algorithm is used for the single vehicle DARP. In Psarfatis (1986), two different algorithms for the static version of the multi-vehicle DARP are compared. One of these algorithms is based on a clustering technique. Clustering is a technique also used by many others, among which Jaw et al. (1986), Ioachim et al. (1995) and Borndörfer et al. (1997) are good examples.

Jaw et al. (1986) developed a method based on clustering for the static version of the multi-vehicle DARP with service quality constraints. A large number of clusters are constructed through column generation. Experiments have been made on instances of 50 to 250 requests, as well as on a real-life problem consisting of 2545 requests. In the experiments, the clustering approach is also compared to a parallel insertion heuristic. The experiments show that the clustering algorithm improves both the quality of the solutions as well as the computation times. What the algorithm does is to sequentially process each travel request in the list, assigning each request to a vehicle until the list is completely traversed. The processing of a request \( i \), can be described as follows. For each vehicle \( j \) (\( j = 1, \ldots, n \)): find all feasible ways of inserting request \( i \) into the schedule of vehicle \( j \). If no feasible insertion can be found, continue with the next vehicle, otherwise, find the insertion of request \( i \) in the schedule of vehicle \( j \) that gives the least additional cost, and call this cost for \( COST_j \). If no feasible insertion of request \( i \) can be found into the schedule of any vehicle, then the request is declared as a “rejected request”, otherwise, request \( i \) is assigned to the vehicle \( j^* \), which has the lowest of all the \( COST_j \), that is for which the additional cost of request \( i \) is lower then for any other feasible insertion of request \( i \) into any other vehicle \( j \). In this way the insertion that minimizes the additional disutility experienced by other travelers is identified. The algorithm developed by Jaw et al. (1986) was then adapted by Alfa (1986). The main adaptation is the use of variable capacities on the vehicles. In the studied scenario, some of the vehicles
have convertible seats that can be transformed to accommodate wheelchair passengers. Additional constraints are formulated to handle this.

The technique in Ioachim et al. (1995) is based on a mini-clustering method, which involves solving a multi vehicle pick-up and delivery problem with time windows by column generation. What this technique does is to group the trips into clusters, and then the algorithm uses a shortest path technique to re-optimize the distribution of the mini-clusters to the different vehicles. The shortest path technique used generates a new column consisting of a new vehicle trip. The goals of re-assigning the clusters to the vehicles are to minimize the number of vehicles, the number of vehicle trips and the total travel time.

Borndörfer et al. (1997) use a set partitioning approach consisting of a clustering step and a chaining step. The clustering step generates possible clusters (by complete enumeration) and solves the clustering set partitioning problem to select the best set of orders such that each request is part of exactly one order. The biggest reason for this step is to reduce the size of the problem. The chaining step generates a set of feasible tours and solves the chaining set partitioning problem, in this way choosing a best set of tours. Although the chaining problem is much larger then the clustering problem, both these set partitioning problems are solved with the same branch-and-cut algorithm.

During the last decade, the interest in heuristics, and especially metaheuristics, have increased dramatically. This is something that has been quite noticeable in the work regarding DARP. Since the DARP is a computational demanding problem, the use of heuristics has dominated during the last years. Tabu search is the most commonly used metaheuristic for solving the DARP. Cordeau & Laporte (2003b) uses a tabu search algorithm on several different data sets. To model the DARP in a more realistic way, the authors use the time windows for pickup and drop off in a certain way. For outbound trips they let users define time windows on the arrival times and on inbound trips on the departure time. In addition to this there is also an upper limit of the ride time of any user as well as constraints regarding vehicle capacity and route duration. During the search, relaxations of vehicle capacity and time window constraints are allowed. In this way the authors have the possibility of exploring infeasible solutions during the search. Other authors have used the same data as in Cordeau & Laporte (2003b), for example Bergvinsdottir et al. (2004) and Attanasio et al. (2004).

Chan (2004) uses a cluster-first route-second approach, where the clustering has been made both with tabu search as well as with scatter search. For both clustering methods two different techniques for routing have been used. The first one generates a route that is always feasible, and where requests that cannot be assigned in a feasible way, will be left unassigned. The second one generates a tour that might be infeasible, but where all requests are assigned.
Ho & Haugland (2004) study the probabilistic DARP, where each user requires service with a certain probability. This is an instance of the DARP especially useful when creating vehicle tours to be used for a given time period (more than just at one particular time). Reoptimizations are not considered on daily basis, only removals of customers not requiring service are allowed. Regarding the heuristics, both a tabu search heuristic as well as a heuristic that is a hybrid of tabu search and GRASP (Greedy Randomized Adaptive Search Procedure) is used. The conclusions are though that the tabu search performs better then the hybrid GRASP-tabu search.

Attanasio et al. (2004) test different parallel heuristics, based on tabu search, for the dynamic DARP. The authors uses the static tabu search of Cordeau & Laporte (2003b) to find a solution to the static problem of the requests known at the start of the planning horizon. The experiments made indicate that parallel computing can be beneficial when solving real-time DARP.

Except for tabu search, also other metaheuristics have been tried on the DARP. Baugh et al. (1998) use simulated annealing to solve the DARP. A cluster-first, route-second approach is used. Simulated annealing is used for the clustering, and a greedy algorithm is used for the routing after that the clusters are made. The clustering starts with randomly assigning customers to clusters. Two types of operations are then used to alter the clusters. The first one is to simply lift one customer out of the cluster, and insert into an other cluster. This operation can change the number of clusters. The second operation possible is to let two costumers change clusters, always giving the same number of clusters as before the operation. In each iteration of the simulated annealing, the routing algorithm is run on those clusters that have been changed, giving a new objective value. The objective is evaluated on the total distance (of all vehicles), the number of vehicles used as well as on the total disutility observed by the customers.

Uchimura et al. (2002) use genetic algorithms to solve the DARP. Pick-up and drop-off nodes are stringed out by random and in this way creating the individuals of the first population. After this, the algorithm iterates for 1000 generations, where for each iteration the individuals with the best fitness value are kept as the next population. The experiments made focuses on comparison of the genetic algorithm, a standard edge exchange algorithm (2-opt algorithm) and a combination of the genetic algorithm and the 2-opt.

Bergvinsdottir et al. (2004) present a genetic algorithm based on a cluster-first, route-second approach. The algorithm is tested on the same data used by Cordeau & Laporte (2003b), and the results are also comparable to these. One interesting feature of this work is the possibility of altering several factors of both cost of operation as well as service level. This enables the possibility of evaluating the consequences of different scenarios.
Many authors have also used different forms of insertion heuristics to solve the DARP. Madsen et al. (1995) use an algorithm based on an insertion heuristic to solve the DARP with multiple capacities and multiple objectives. The algorithm was developed to solve a real-life problem of scheduling transportation for elderly and disabled in Copenhagen, Denmark. Diana & Dessouky (2004) presented a parallel regret insertion heuristic to solve large instances of the DARP. Data sets of 500 and 1000 requests have been tested. Toth & Vigo (1997) developed a parallel insertion heuristic to be able to find good solutions also to large instances within quite small computational times.

More special instances of the dial-a-ride problem have also been studied. In Daganzo (1978) an analytic model is presented, that forecast average waiting times and ride times for a dial-a-ride service. In Fu (2002a), the problem of scheduling dial-a-ride under time-varying, stochastic congestion is studied. The work of Dessouky et al. (2003), presents a methodology for optimizing cost, service and environmental consequences of a dial-a-ride system. Results of simulations show that it is possible to reduce environmental impacts to a large extent at the same time as operating costs and service delays only are increased slightly. In Xiang et al. (2006) a quite realistic instance of the DARP is studied. This study includes a heterogeneous vehicle fleet and drivers with different qualifications. In practice, this is often the real situation.

An other quite well studied application of the DARP is that of scheduling elevators. This has been studied by simulation in Grötschel et al. (1999). In Hauptmeier et al. (2001), a cargo elevator system is used to illustrate a dial-a-ride system with precedence-constraints for requests starting at the same vertex. In this example, conveyor belts deliver goods to the elevator, where all goods arriving to the elevator at the same floor must be handled in a first-in-first-out manner. The work of Coja-Oghlan et al. (2005) also uses the elevator scheduling task to show that a dial-a-ride problem with a caterpillar network structure, and a server only handling one request at a time, is NP-hard in worst case but in most cases can be solved in an efficient way.

This illustrates the fact that many different kind of problems in which something (or someone) is to be transported from one position to an other by some sort of server can be formulated as a DARP. Even if the real world applications of the DARP can differ quite some, the model formulation is still more or less the same. Because of this, experiences from one application can many times be useful for an other.

This literature review shows that the DARP is a well studied problem. Both static and dynamic versions of the problem have been studied as well as the probabilistic DARP. Since the DARP is hard to solve, focus has been on heuristics.
3 Planning of Integrated Services

As described in Chapter 2, a lot of work is done on how to plan both fixed route services as well as demand responsive services separately. In this chapter, focus is on how to plan an integrated service. This service is intended to be used in urban traffic systems and the service should suit a wide range of customers from different market segments. Both the category of elderly and disabled as well as any other public transport customer shall be able to use the service.

Section 3.1 describes how the integrated service is intended to operate. Section 3.2 presents previous work on integrated services. How an integrated service can be planned, and a framework for the tools that are needed for this, is described in Section 3.3. Section 3.4 describes the benefits of having a planning system based on a geographic information system (GIS) and how different planning tools can be included in such a system.

3.1 Integration of Area Covering and Fixed Route Services

An integrated service built up by a demand responsive service and a fixed route service can be designed in a number of ways. The differences between these designs primarily depend on what type of demand responsive service that is used. If elderly and disabled shall be able to use the integrated service, it is essential that the service can be provided very close to the desired points of origin and destination (or even at the exact addresses). The demand responsive service must therefore be an area covering service and there are two main services of interest. These two are dial-a-ride and multi hire taxi.

The demand responsive service can be used to carry passengers from their origin to an appropriate transfer location to the fixed route network, and/or from the fixed route network to their destination. It may be of great advantage to the provider of public transportation, due to cost effectiveness if a demand responsive service could be combined with the fixed route service. Also the passengers can benefit from this due to increased availability to the public transport service, and an increased level of service. These benefits of an integrated service are described in Figure 5, and are also the reason that many transit agencies have considered this possibility.
Most applications and modeling methods involving dial-a-ride systems are made from the operator’s perspective. The objective is then to minimize the total travel distance, or a generalized cost, of the demand responsive vehicles subject to constraints regarding time-windows for requests and vehicle capacities. From the passengers’ perspective the total travel time is more relevant, but usually not included in the objective function. An ordinary taxi service operates from this perspective, since when no ride sharing occurs this is the most economical way to handle the requests. This way to operate the vehicle fleet is a quite expensive one. In the case of taxi customers, these customers themselves have to pay for the high expense of this service. But in the case of integrated traffic the operator is more interested in minimizing the first objective, since this minimizes several variable costs such as the number of vehicles and drivers needed, fuel consumption etc. Techniques like the two mentioned in Section 2.5.3 by Jaw et al. (1986) and Ioachim et al. (1995), are well suited to use for planning the demand responsive part of an integrated journey, in the case of a static planning situation.

The integrated service is intended to be used in such a way that a user can travel with the demand responsive service to a transfer point connecting the demand responsive service to the bus network. Transfer points are the bus stops at which it is possible to change between a demand responsive vehicle and a fixed route bus, and vice versa. If necessary, the passenger can then transfer again from another transfer point in the bus network to a second demand responsive vehicle, operating in another area of the city, and with this vehicle travel to the destination. The journey can of course also include transfers between bus lines. A typical route, including two demand responsive vehicles, is described in Figure 6.

Figure 5: Benefits of an integrated service
Alternative use of the integrated service include only one demand responsive vehicle in addition to the fixed route bus service for travel from an origin to a destination, or include one single demand responsive vehicle taking the passenger all the way from origin to destination. The relative use of these alternatives depends on the demand pattern, the cost structure and on the service levels offered to the customer. These factors also affect the overall performance of the integrated service.

The above description shows how the service is intended to be used when passengers are picked up and dropped off at the exact addresses of their origin and destination, i.e. when door-to-door service is provided. An other way of operating the service is to use a large number of meeting points (bus stops for the demand responsive service) scattered over the area where the service shall be available. The only difference is that the customers have to walk to, and from, these meeting points. By this reason it is very important that a large number of meeting points are used. In this way, journeys can be built up in the different ways presented in Figure 7.
The fixed route service could be of any kind suitable for urban traffic, for example bus, tram or light rail. In a substantial part of all Swedish towns however, busses are the only fixed route public transport available. The fixed route service should have highly frequented routes. If routes with low departure frequencies should be used, coordination at the transfer locations must be made between the vehicles, and in this way complicating the construction of integrated journeys. The use of low frequented routes without any coordination to the demand responsive vehicles increases the transfer times. In case of coordination between a demand responsive service and a fixed route service there are the additional problem of different travel times for the demand responsive service. This type of problem is very complex since the travel time depends on what requests that have been assigned to the specific demand responsive vehicle. Timetable planning for integrated services is an area where more work must be made.

### 3.2 Modeling of Integrated Services

Integration between different services with fixed routes is not anything new. For example, busses arriving at train stations, even with coordinated timetables between the two transportation modes have been around for quite some time. The problem of scheduling an integrated service consisting of two fixed route services, train and bus, operated by two different operators are for example studied in Li & Lam (2004). Also in Martins & Pato (1998) a combination of train and bus services is studied. The problem is to design a feeder bus network given a rail network, with the objective to minimize a cost function considering both the operator’s and the customers’ interests.
This type of integrated systems as well as the form of flexible systems presented in Section 2.4 that combines features of both fixed route and dial-a-ride services are of course interesting. However, since a fixed route service always will be necessary to provide a service effective enough for those who demand a fast public transportation service, a combination of a fixed route service and a demand responsive service seems more useful. The kind of integrated services that combines a fixed route service and a demand responsive service has not been studied in the same way, especially not for local area (city-based) services.

Nevertheless, some work is done on this kind of integrated transport services, mainly focusing on reducing the costs for transit agencies to transport elderly and disabled. Also some work has been done on flexible and integrated public transportation systems intended for a general public and not only for paratransit customers. The main difficulty of operating an integrated service is to schedule transit trips as a combination of demand responsive and fixed route transit service. Both passenger trips and vehicle trips must be scheduled, and it therefore makes the planning of an integrated service more complex than that of a single mode service. To solve the problem of scheduling trips to an integrated public transport service, a number of inputs are important to be known, and can be regarded as essential.

- the location of the passengers’ origins and destinations
- the passengers’ requested times, and associated time windows, in which pickups and drop-offs must occur
- the location of fixed route stops
- the schedules of all fixed route vehicles
- the accessibility level of all fixed route vehicles and transfer points
- the time windows in which demand responsive vehicles are permitted to meet fixed route vehicles at transfer points
- vehicle capacities
- passenger loading and unloading times
- the distance between stops
- minimum passenger level of service standards

All these inputs are necessary to be known to be able to plan the integrated trips. As for the DARP, both the operator’s and the passenger’s perspective must be considered when planning integrated trips. Either both perspectives are part of the objective, or one perspective is part of the objective while the other is controlled by constraints.
Integrated public transport systems were studied already in Potter (1976) and Wilson et al. (1976). Potter (1976) describes an integrated service where 45 dial-a-ride vehicles and 36 express buses are operated in Ann Arbor, Michigan. The bus routes cover all places with a high number of requests going to or from. The dial-a-ride vehicles are assigned to different zones, and acts as feeders to the fixed route service as well as taking care of intrazonal journeys. In time periods with low demand, connections between different dial-a-ride vehicles are also made.

The work of Wilson et al. (1976) is more focused on algorithms for planning the journeys. The problem has a passenger utility function as its objective, and this function is maximized subject to a series of level of service constraints. In this way, the costs of the operator are not included explicitly in the model. A trip insertion heuristic is used to schedule both passenger and vehicle trips. Opposite to this, the work of Liaw et al. (1996) has a model for the integrated problem with the operating costs as its objective.

Hickman & Blume (2001) take both passengers and operators objectives into account, by explicitly inserting the transit agency cost as well as the passenger level of service in the model. The goal in scheduling vehicle trips is from the operator’s perspective to minimize the total cost of the service, while it from the passengers’ perspective is to maximize the level of service; i.e., minimize travel time, transfer time and the number of transfers. The way this is implemented, so that the objectives of the operator and the passengers are balanced, is a heuristic that schedules the integrated trips in a way so the operators costs minimizes, subject to passenger level-of service constraints.

The method proposed divides the problem into two parts. The first part is to find feasible itineraries, for the requests suitable to integrated service, that connect the passenger’s origin and destination in a way that maximizes the traveler’s level of service. If the itinerary meets all of the level of service constraints, the trip is scheduled. The second part is that the demand responsive legs of the passenger’s itinerary must be added to a specific vehicle, so that the legs are included in a vehicle’s schedule. This is done in a way that minimizes the costs of the operator. The thoughts behind this decomposition was to make the technique improved over that of Liaw et al. (1996), by having the passenger’s level of service considered explicitly, and over that of Wilson et al. (1976) by explicitly including operating costs into the decision making process of the vehicle scheduling.

In Horn (2002) the way a dial-a-ride system interacts with long-distance transportation systems have been studied. Procedures for planning journeys combining fixed route and demand responsive modes are further described in Horn (2004). The tests made with simulated demand shows that the procedures are well suited for a real-time traveler information system.
In Aldaihani & Dessouky (2003), the objective function contains two measures of performance, minimize total travel distance of demand responsive vehicles and minimize total travel time of passengers. The proposed heuristic is tested on real-life data obtained from Antelope Valley Transit Authority. Regarding the methods proposed to solve this problem it is assumed that the requests are known before the scheduling process takes place, i.e. a static problem is considered. In the model it is further assumed that all demand responsive vehicles dispatches from, and at the end of the day must return to, their depot. The capacity of the fixed route buses is assumed enough not to be considered as a constraint, whereas the demand responsive vehicle have known capacity. The method is also based on the fact that a maximum of two transfers is allowed.

An other type of integrated public transport service is presented in Crainic et al. (2001). This service is based on flexible routes of the form presented in Malucelli et al. (1999), and described in Section 2.4, combined with a conventional fixed route service. It is assumed that the fixed route service is a fast service of some kind, light rail or an express bus for example, and that the compulsory stops of the flexible service are typically such that the users here can transfer to other lines (fixed and flexible). In an integrated network, passengers can travel from an optional stop to any other, with or without transfers between lines (both flexible and fixed route). The itinerary for each user is not determined uniquely, but depends on the schedule and availability of the system. This means that two trips, from a given origin to a given destination, occurring at the same time two different days doesn’t necessary follow the same itinerary. The itineraries depend on how the vehicles have been routed to respond to other requests. This solution, that a passenger can travel from an optional stop at one side of the network to an other optional stop anywhere in the network, gives an almost personalized transit system. A personalized transit system, such as a dial-a-ride service operating door-to door, but with costs almost as a traditional fixed route system.

An other important problem regarding integration between a demand responsive service and a fixed route service is how to design the zones in which the service should operate, and to decide what geographical areas that are of interest for such a service. The purpose of Durvasula et al. (1998) was to demonstrate the technical feasibility of operating a route deviation bus service, and in that way show that spatial analysis capacities provided by a GIS, can be used to support the operations of a route deviation service. To achieve this, design alternatives for a route deviation service where investigated. As a part of this work, the effects on the system depending on the service zone size where analyzed.

Analysis somewhat like these, but for an integrated service is presented in Aldaihani et al. (2004). An analytical model meant to be an aid in designing the network of an integrated service is developed. The problem has an area divided into zones, by a fixed route grid, where each zone is served by a number of demand responsive vehicles. If the destination of a request is in another zone, then the demand respon-
sive vehicles transfer the passenger to a fixed route. Since a grid structure of the fixed routes is assumed, it is possible to travel from any zone to all the others. If the destination of a request is in the same zone, the demand responsive vehicle will transport the passenger from the origin to the destination. The model developed in Aldaihani et al. (2004) determines the number of zones, demand responsive vehicles, fixed routes, and buses in each route. This is done in the way that minimizes the total cost, as a combination of both the operator’s and the passengers’ costs. It should be noted that in this model only one passenger (or group of passengers with the same origin and destination) is served at a time by a demand responsive vehicle. That is, no ridesharing is allowed.

The work of Ceder & Yim (2003) also concerns how a service should be designed. A demand responsive service is to be designed to feed a train station in a so effective way as possible. This study is done by simulation, and ten different routing strategies are simulated.

As can be seen by this literature review, most work in the field of integrated public transport systems have been made on heuristics to solve the problem of assigning passengers to vehicles. The complexity of this problem indicates that heuristics are necessary for operational planning in real time. For strategic planning the algorithms does not have to work in real time, and therefore exact models can in some situations be an alternative.

3.3 A Framework for Planning of Integrated Services

Many of the problems involved in planning of a public transport system are complex, and therefore it is difficult to find good solutions to them (as previously described in Chapter 2). In addition, there is also the problem of how the solution of one problem affects the solutions of other problems. Even if optimal solutions are not required to all problems (or even to any of the problems) it is many times hard to know how the different problems affect each other.

The fact that the problems are so difficult and complex indicates the necessity of a computerized tool as an aid in the planning process. When planning any public transport service today, several different tools are usually used. Some tools for analyzing demand, some for strategic planning and some for operational planning. Analyzing demand and the effects on availability to the service depending on different scenarios are many times studied in a GIS, while the actual planning is done using other tools.

If operators of public transport shall be able to implement an integrated service in large scale, it is necessary that they have tools for both strategic and operational planning. Several different tools exist today for the different problems. With a graphical interface and a simple way to handle all the necessary data, there are good
possibilities to create a framework that can be of great help in planning of integrated public transport. The problems involved in both strategic and operational planning highly depend on each other, indicating the advantage of having one framework for studying all these different aspects. It is preferable that such a framework can handle as many problems as possible of the ones involved in the planning process.

Three different components must be included. The framework must have the possibility to handle and analyze different types of data. Most of this data has one common feature, information about geographical positions. For this reason, a GIS is an essential part of the framework. The complex structure of the included problems emphasizes the importance of different optimization tools. Optimization tools can be useful in both the strategic planning as well as in the operational planning. To analyze the effects of a chosen strategic design of the service, a simulation tool must also be part of the framework. The GIS can be seen as the central part of the framework. It is used to prepare input to, and visualize output from, the two other components of the framework, the optimization and simulation tools. A sketch of the framework is described in Figure 8.

![Figure 8: The information flow in the framework](image)

How the different modules in the framework interact with each other can be seen from two different perspectives. The most common way, in which optimization and simulation tools are used in combination, is to first find an optimal solution to a specific case, and then simulate the effects of this solution. This is the first perspective. The other is to find a good overall design by the use of simulation, and then use optimization to find the best solution to a specific instance of the given design. In both cases, one can iterate between the optimization and the simulation.

The first perspective can be described as using an "operative" tool to describe the effects already in the strategic planning. The second perspective is normally used when going from strategic planning to operational planning.
In this framework, a model of the real world is represented in the GIS. As in any model, the model in the GIS is of course a simplification. This model can still include much more information than needed by a specific optimization or simulation tool. The information needed for a specific optimization tool is taken from the GIS module to the optimization module. The optimizations are performed and the results are sent back to the GIS module and/or to the simulation module. In the GIS, the results are visualized, and in the simulation module it can be studied what effects the new results have on the overall performance of the service. In the simulation module, most of the input is taken from the GIS. If results are obtained from any optimization tool, this data complement the data from the GIS. The output of the simulations is then also sent back to the GIS for visualization.

The different steps in the planning process use the separate modules in different ways. In the design steps of the strategic planning (network design and design of the demand responsive service), the GIS can many times be used as the actual planning tool. For the frequency setting, timetabling and the steps in the operational planning, optimization is the tool to use. For all steps, the effects of the solutions can be studied through simulation. For most steps, simulations are done to see what the effects are for the passengers. But for vehicle scheduling and crew scheduling simulations are done to find out how sensitive the solution is to different disturbances. Nevertheless simulations can be useful in all steps of the planning process. For frequency setting, timetabling and assigning passengers to vehicles, the GIS is also a good help in visualizing the results.

To create a framework of this kind is in many ways to put different tools for optimization and simulation together, in a user-friendly environment. In the suggested framework, new intelligence in the form of planning tools, are added to the GIS. These new tools together with the original possibilities of the GIS create the components of the framework in the way described in Figure 9.
For most of the optimization tools, the needed mathematical models have already been studied extensively, as described in Chapter 2. The most important step in planning of an integrated service is how to schedule customers to vehicles. This includes determining which vehicle that shall pick up the customer, at what transfer location the transfer to the fixed route shall take place, which vehicle the passenger shall transfer to from the fixed route and at what transfer location this shall be done. This also includes creating itineraries for all demand responsive vehicles. Work done on this problem is presented in Section 3.2, and an exact model to this problem is presented in Chapter 6.

Assigning passengers to vehicles is part of the operational planning process. But a mathematical model used for assigning passengers to vehicles can still be used also for strategic planning. In the strategic planning of an integrated public transport service, the operational decisions should be simulated, according to the description in Chapter 2, and Figure 1. It is though not likely that a model developed to find optimal solutions to the operational problem, as the one described in Chapter 6,
could be used in actual operational planning. Such models are most likely too com-
putational demanding to be used in real operational planning situations. How the
described planning tools can be added to a GIS, and what the benefits are of using
a GIS, are explained in the next section.

3.4 Benefits of the GIS Module

The use of GIS in transport planning has increased during the last years. The rea-
son why GIS should be used in transport planning can be seen from two different
perspectives, as very well explained in Berglund (2001).

From the perspective of those using a GIS for different spatial analysis, a quite
common opinion is that these systems often lack sufficient tools for studying mobi-
licity. From this perspective, the aim is therefore to include such modeling tools in an
existing GIS. The other perspective is that of transport modelers wanting to use a
GIS as an aid to visualize modeling results or to prepare input data to such a tool.
From this perspective, the aim is often to simply transform data from planning tools
to a GIS, and vice versa.

Tools for transport planning are usually restricted to only handle information needed
for the specific planning situation the tool is intended for, i.e. the information needed
by the algorithm. If a tool used for such planning also should be possible to use for
analyzing how different scenarios affect customers, the tool must be able to also
include much more information. One example of this is the representation of the
most important piece of information in any transport problem, the network. In most
transport planning tools the network is represented by links and nodes where each
link connects a from-node and a to-node, but without describing the actual geometry
of the link. In this way, each link is actually described as a straight line between the
two connecting nodes. In a GIS however, also the geometry of the links are easily
represented.

The way in which the geometry of objects is described is very important in a number
of situations. To use a GIS as the central part of the framework gives the possibil-
ity to use different level of details for different analysis and planning tools. For an
optimization tool in which a high level of detail is not needed, the geometry can be
simplified before it is sent from the GIS to the optimization module. At the same
time, all details are still available if needed for analyzing the results. In this way,
the information can be taken from the GIS and transformed in any way suitable for
the different tools. Even though the data is used in many levels of detail in different
tools, every type of object only needs to be represented in one single database. This
is something that really simplifies the management of such a database and reduces
the risk of inconsistency. It can also contribute to reducing redundant information
(in the form of duplicates), even if there is a risk that unnecessary information (that
is not used by any tool) also is kept in the database.
When creating a modeling tool, there are more advantages of using an existing GIS to include the tool in, instead of creating a stand-alone system. The most important advantage is perhaps the flexibility of a GIS. Even though the modeling tool only uses a limited number of predefined data sets, the original tools for analyzing data in the GIS can be used on any data set, in this way enabling follow up analyzes on the results given from the modeling tool. The fact that any data with geographical information at any time can be added to a GIS gives a great flexibility of what kind of such analyzes that can be made.

The benefits of using the existing interface of a GIS are also not to be neglected. A well functioning interface is of course very important to create a user-friendly environment. The possibility of good visualization of the data used in the operations, as well as of the results, are also essential for this kind of analyzes. A GIS is not intended for a specific planning or analyzing situation, and is therefore easy to adapt and to include new tools in.

To include an own designed tool, or to connect the GIS to an other existing program, can be done in different ways. There are basically three different ways to add new functionality to a GIS, as described in Berglund (2001). First of all, data can be exported from the GIS to the modeling tool and vice versa. In this way all computation is done in either the GIS or in the modeling tool. As an example, input data can be prepared in the GIS, exported to the external modeling tool and computed there, and thereafter exported back to GIS for visualization or further analyzes of the results. This can either be done manually, or called from the GIS interface. If done manually, the new functionality of the modeling tool is actually never added to the GIS. The second way is to add macros written inside the GIS, using languages provided by the GIS. Such macros are often quite simple to create, but often requires relatively large amount of computation in relation to what the code performs. By this reason macros are generally only used for short and simple programs. The third and final way to add new functionality is to fully integrate the modeling tool inside the GIS. Normally this requires the source code of one of the systems to be combined. Nowadays providers of GIS also provide developing tools to simplify integration of own developed tools or other existing software. Still, in research projects the simplicity of the first way of including new functionality to a GIS is most often used. Fully integrated systems are more common in the development of commercial products.
4 The LITRES-2 Modeling System

When designing and evaluating an integrated public transport service one need a simulation tool that is able to describe effects on both fixed route services as well as on demand responsive services. One such tool is the LITRES-2 that will be described and evaluated in this chapter. In simulations performed to evaluate this software, input data from the town of Gävle has been used. Section 4.1 gives an overview of how the system works and what it can be used for. In Section 4.2 we described how different parts of the system interacts. Section 4.3 describes the input needed to build a simulation model, and Section 4.4 describes what output that could be obtained from a simulation. In Section 4.5 we identify what specific help LITRES-2 can offer in planning of an integrated service. Finally, in Section 4.6 we give some summarizing comments about LITRES-2.

4.1 Description of LITRES-2

LITRES-2 is a simulation software developed by the Commonwealth Scientific and Industrial Research Organisation, Mathematical and Information Sciences. The software is designed to model the operation and performance of urban public transport systems, including multi-modal (integrated) journeys, in order to estimate their likely performance.

The system uses simulation to describe how travel requests are being executed, given a set of possible transportation modes. LITRES-2 takes the passengers’ perspective and simulates the route choices and use of the service based on experienced generalized costs, consisting of monetary costs and travel times. The planning of the vehicle fleets is made as a consequence of the passengers’ choices. So, basically the main objective is to simulate travelers’ behavior as realistic as possible. To obtain this, the objective is to minimize the generalized cost for each request. The generalized cost, $C$, for a journey with fares $F_j$ and time-components $t_i$, for travelers from market segment $c$, with time-weights $w^c_i$, is defined as follows:

$$C = \sum_{j \in J} F_j + \sum_{i \in I} w^c_i \cdot t_i$$

where $j \in \{\text{in-transit legs}\}$ and $i \in \{\text{walk, in-transit legs, origin-wait, transfer-wait}\}$, Horn et al. (1998).

A number of different forms of public transportation can be simulated with LITRES-2. Static public transport (bus, train, tram etc.), demand responsive services (taxi) and hybrid services (roving bus and smart shuttle) and also journeys built up by combinations of these modes can be simulated.

37
• The static public transport services have fixed routes with designated stops and fixed timetables.

• The taxi services can be in form of a regular single hire taxi, or a multi hire taxi. Both these forms can be modeled as opportunistic or planned (book-ahead) services.

• The hybrid services available are Roving bus and Smart shuttle. A Roving bus operates between bus stops, but has no fixed route or timetable, and only visits stops where there is a request. This makes the Roving bus quite similar to a multiple-hire taxi, but only operated between predefined stops. Smart shuttle is timetabled but only runs on demand.

The hybrid services are normally regarded as demand responsive services, but in LITRES-2 they are referred to as just hybrid services. The hybrid services are called this because of their mix between static public transport and taxi service. Different modules of the software handle static public transport and taxi (see Section 4.2), but the hybrid services are handled by a combination of both these modules.

Since the software actually simulates the way people choose to travel given a number of alternatives, it is suitable for planning and analyzing scenarios of the following kind:

• Changes to existing services (public transportation network, timetables etc.)

• Changes to fare structures

• Changes in the distribution of transport demand

• Inter-modal coordination

• Introduction of new type of services

4.2 LITRES-2 Architecture

LITRES-2 consists of two main components, a simulator and a control system, see Figure 10. The simulator models the demand and vehicle movements, and passes the events to the control system. Then, the control system plans journeys on the basis of the events received from the simulator, and also manages the itineraries of the demand responsive vehicles. One reason for this distinction between the two parts is that modules within the control system are intended for future use in real-time contexts, especially algorithms in the request-broker and journey-planner modules, Horn (2003b).
The LITRES-2 Modeling System

The simulator module consists of several sub-modules. Two major such sub-modules are the demand simulator and the network simulator. The demand simulator transforms an aggregated demand into a stream of travel-requests. These travel-requests are thereafter sent to the control modules. The aggregated demand model consists of information regarding market segments, party sizes and time periods, and can be represented either by OD-models (Origin-Destination) or by SS-models (Source-Sink). The SS-models are more suitable when a large number of requests is to be carried out to, or from, the exact same location, since the OD-models only uses predefined regions and not exact addresses.

The network module tracks movements of demand responsive vehicles, and therefore interacts closely with the fleet-scheduling module. Apart from these two simulator-sub-modules, there are some more handling vehicle breakdowns, delays and no-shows.

A schematic sketch of the different parts of the LITRES-2 architecture is shown in Figure 10.

The fleet-scheduler and the journey-planner are the central components of the control module. The fleet-scheduler manages the deployment of demand responsive vehicles, and plans each vehicles itinerary on the basis of trip-requests received from the request-broker (also part of the control module). The trip-requests contain information regarding transport mode and time-windows for departure and arrival. When
choosing implementation for a specific trip, the objective for the fleet-scheduler is to minimize the costs for the operator. A sub-problem to this is to find the shortest-time path between successive pair of stops. In other words, the fleet schedulers task is to assign a set of passenger trips to a set of vehicles, and to construct routes for all included vehicles. This is the problem known as the “pickup and delivery problem with time windows”, Horn (2002). The journey-planner plans multi-leg journeys, with the objective to minimize the generalized cost for the traveler, as described in Section 4.1. The journey-planner uses a branch-and-bound algorithm, where each branch of the tree represents a possible way of travel between the origin and the destination. Since it is most likely to find journey alternatives with low generalized costs among those with a few numbers of legs, the tree of possible journeys is searched with a breath-first approach, Horn (2002). By using this approach a fairly low generalized cost will be found quickly, and therefore also give an effective value for the upper bound of the generalized cost.

4.3 Input to a LITRES-2 Simulation

To perform a simulation with LITRES-2, input data of several kinds must be defined. The road network consists of a set of nodes and a set of directed links connecting the nodes. The nodes and the links are defined in separate files. The nodes file contains only information about the coordinates of the nodes. The links file contains for each link information about which nodes the link is connecting, the length of the link and the link speed. The link speed can vary over time, but the speed for each such time period must be defined in the links file. The model of the road network is used both for estimating travel times and for routing vehicles, in the fleet-scheduling and network simulation modules. To find shortest-time paths, variants of Dijkstra’s algorithm are used Horn (2002), with the addition of off-network allowance for travel between the starting point and a nearby network node, and the same for the end point.

The region to be studied is divided into several polygonal bounded area units, zones. These zones are used as spatial reference to the demand models and for defining fare structures, Horn (2003b). For each zone, the number of vertices on the polygonal boundary of the zone, and all geographical coordinates of the vertices must be defined, Horn (2003a), and stored in a region file.

Different market segments can be defined to describe the behavior of different types of travelers. Each market segment defines parameters for a specific set of travelers. The most important to define for each market segment is the set of modes acceptable to members of the segment. Another important parameter that must be set is the maximum walking distance. Furthermore the cost of waiting, walking and transiting must be defined. All this information regarding the market segments is stored in a separate market segments file.
Aggregated demand models can be represented in the form of OD-models or SS-models, or both, as mentioned earlier. In an OD matrix file, a number of matrices can be defined. For each such matrix, the start and end times must be stated. The values inside the matrices are not averages of journeys per hour, but the actual number of journeys that shall commence during the time period for which the matrix is valid. A Source-Sink file consists of sub models called DemSource and DemSink. The DemSource sub model specifies journeys from a fixed origin point to each zone, while the DemSink sub model specifies journeys to a specific destination point from each zone, Horn (2003a).

The aggregated demand models must be disaggregated into specific requests. This is done in a randomized fashion. The start and end points for a journey to be made between zone $i$ and $j$, during the period $(t_{lo}, t_{hi})$ for which the OD-model is valid, is randomized in the following way. To find a starting-point, the smallest rectangle enclosing the entire zone $i$ is used. The width and height of this rectangle is referred to as $w_i$ and $h_i$. The starting point is then $(x, y)$, where

$$x = x^1_i + w_i \cdot p$$
$$y = y^1_i + h_i \cdot q$$

and where $x^1_i$ and $y^1_i$ are the left-hand edge, and the bottom-most edge respectively of the rectangle enclosing zone $i$, and $p$ and $q$ are random numbers in the range $[0, 1]$. The same method is used to find a destination point in $j$. The start time for the requested journey, $t_{start}$, is set in a similar way, but here $r$ is a random number in the range $[0, 1]$ Horn (2003b).

$$t_{start} = t_{lo} + r(t_{hi} - t_{lo})$$

To include timetabled services in the simulation model, four different types of files are needed. First of all a public transport points file must be defined. Such a file contains the coordinates for all points at which public transport service is provided. This is similar to how the network nodes were defined in the nodes file. The points used for transfers are marked in this file and each public transport point is also given a identification number, that is cross-referenced to the second file needed, the timetable file. The timetable file contain all information regarding a specific public transport route, including route-name, what mode that is used in the route, all the stops in the route, and departure times. Before a LITRES-2 simulation including timetabled modes can be accomplished, a timetables list must be created. This file should contain a list of the timetables specified in the timetable files, i.e. the route-names. The last type of file needed to describe timetabled services in the simulation model is a fares file. This file contains a matrix specifying the fares between all zones. These zones can be the same as those defined in the region file used for the OD-models and/or SS-models.

To use demand responsive modes in the model the vehicles belonging to these modes needs to be specified in a demand responsive fleet, this is done in a fleet file. In each fleet file information about all the vehicles belonging to that specific fleet is stored.
The information stored is passenger capacity, the mode or modes the vehicle can offer, the start and end times of the operating shifts of the vehicles and terminal nodes.

With all the required files given in the way described above, simulations can be performed. What output that can be withdrawn from such simulations are described in the next section.

4.4 Output from a LITRES-2 Simulation

The output from LITRES-2 is grouped in three different categories, travel statistics, scheduling statistics and route-usage statistics.

The travel statistics contain the most interesting information when evaluating an integrated service. Included here are, for example, the total number of passengers carried and journey-legs completed, with average travel times, distances and costs. From the scheduler, a lot of statistics regarding the performance of the demand responsive fleets are provided. As an example, the total number of passengers carried by each demand responsive fleet, by each mode and by each vehicle, can be extracted from the scheduling statistics. They also include measures of efficiency and vehicle occupancy. The route-usage statistics contain information such as the number of passengers on each timetabled mode and for each route, with average trip times and occupancy ratios.

The travel statistics, scheduling statistics and the route-usage statistics can all be stored either for the whole simulation period, for pre-nominated time intervals, or for both. The travel statistics can also be broken down by leg-sequence, by market segment and by transport mode, Horn (2003b).

The output from the simulations can for example be presented in a table. Table 1 gives an example of how the effects of the number of demand responsive vehicles can be presented.

<table>
<thead>
<tr>
<th>Number of vehicles</th>
<th>Accepted requests</th>
<th>Generalized cost</th>
<th>Monetary cost</th>
<th>Average total time</th>
<th>Average travel time</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>69,05</td>
<td>26,6</td>
<td>9,5</td>
<td>1277</td>
<td>434</td>
</tr>
<tr>
<td>1</td>
<td>73,91</td>
<td>28,6</td>
<td>10,4</td>
<td>1356</td>
<td>465</td>
</tr>
<tr>
<td>2</td>
<td>78,91</td>
<td>30,0</td>
<td>11,1</td>
<td>1398</td>
<td>477</td>
</tr>
<tr>
<td>3</td>
<td>83,29</td>
<td>31,4</td>
<td>11,7</td>
<td>1445</td>
<td>499</td>
</tr>
</tbody>
</table>
Except for the numerical output, graphical information can be obtained while carrying out simulations in LITRES-2. The network and zones described in Section 4.3 can of course be viewed. More interesting is the information regarding requests, planning and operation of vehicles that is also available. For example, for one specific demand responsive vehicle all planned, ongoing and finished journeys can be studied, see Figure 11. Figure 12 gives an illustration of how all demand responsive vehicles within an area can be viewed and kept track of.

Figure 11: Visualization of one vehicles itinerary, with planned, ongoing and finished assignments
The LITRES-2 Modeling System

Figure 12: Visualization of all demand responsive vehicles within an area

Information about accepted and denied requests is presented as in Figure 13. This kind of information can be very useful in the way that it early in a simulation can be seen in what areas the requests are hard to carry out. If noticed, this can be used in the way that the simulation is terminated, and changes are made directly instead of letting the simulation run to the end. The planning process of one single request can also be viewed, as shown in Figure 14. This is not information suitable to be studied during simulation, but if studied initially it can be used to understand how the planning process in LITRES-2 works.
Figure 13: Visualization of accepted and denied requests

Figure 14: Visualization of the planning process
The Use of LITRES-2 in Planning of Integrated Services

When introducing an integrated public transport service consisting of a demand responsive service and a fixed route service, at least three out of the five original scenarios suitable for LITRES-2 (described in Section 4.1), are represented. Of course, it is an introduction of a new service, and since more than one transport mode is used for some journeys, it is also somewhat a question of inter-modal coordination. In order to find out how to operate the service, simulations of changes to the service can reveal what parameters that affect the operation and in what way.

We have defined and simulated a number of test scenarios in order to find out in what way the software can be of use when planning an integrated public transport system. These test simulations have led to some conclusions about which type of analysis that are possible to do in LITRES-2 and which are of interest when planning an integrated transport service. The numerical outputs of the conducted simulations have not been studied in detail. The simulations have been made with input from Gävle town, provided by the local authority in Gävle.

A relatively small area (approximately 15 km$^2$ of which some parts are sparsely inhabited) has been used in the simulations. The road network within this area is in LITRES-2 represented by a graph consisting of 683 nodes (intersections and endpoints) and 1643 arcs connecting the nodes. The network of bus lines is represented by a similar graph of which the elements also can be elements in the road network. This information as well as timetables and travel requests have all been kept fixed throughout the simulations.

All passengers request transportation from some point of origin inside the area to the city centre, as described in Figure 15. In this way only journeys to the city centre are simulated and no concern is taken to where the travelers actually are going. This is of course a simplification of the real situation, but sufficient for these tests. In Figure 16, the simulated area of Gävle, including the road network and the centroid zones is visualized, and Figure 17 shows the locations of bus stops within that area. The requests have been stated for two different market segments (categories of customers). One segment for all of those normally traveling by bus, and one for those normally traveling with the Transportation of the Disabled Services.
Figure 15: Structure of the integrated service used in the simulations

- Bus (or other fixed route service)
- Flexible route for the demand responsive service
- Transfer point between a fixed route and the demand responsive service
- Meeting point or other passengers' origins/destinations
- Origin

Figure 16: Description of centroid zones and the simulated area
When introducing a new service the main question is how it should be planned and operated. Two categories of parameters that can affect the service, and therefore can be of interest to simulate, have been identified. These categories are planning parameters and service parameters.

**Planning parameters**

- Design parameters of the fixed route service (networks, timetables etc.)
- Number of, and placement of, meeting points
- Number of, and placement of, transfer nodes
- Number of demand responsive vehicles
- Passenger loading and unloading times
- Pricing alternatives (fare rates and whether to have a combined price for the whole journey or different prices for different journey-legs)
- Vehicle capacity (for fixed route vehicles and demand responsive vehicles)
- Whether or not to have requirements on minimum distance for journey-legs executed by the demand responsive service
- Whether to use meeting points or door-to-door service
Customer service parameters

- Acceptable notice time required by the operator
- Acceptable prolonged travel time (only relevant for demand responsive journey-legs)
- Acceptable size of time windows
- Acceptable total travel time
- Acceptable walking distances for different market segments

The only one of these parameters that cannot be tested in LITRES-2 is the parameter regarding distance-based requirements on journey-legs executed by the demand responsive service. There are also a number of these parameters that can be tested, only with some limitations. Regarding the pricing alternatives, fare rates can be tested both for fixed route services and for demand responsive services. A fixed cost from the origin to the destination, regardless of what mode or modes that are used, can however not be defined. The vehicle capacity can be set for all demand responsive vehicles, but for a fixed route service the capacity is always assumed to be infinite. The acceptable prolonged travel time, size of time windows and total travel time can only be given as percentage of the minimum travel time, and cannot be stated in absolute time. All of the other parameters can be tested in LITRES-2.

In an integrated public transport system the problem is to schedule transit trips that may be performed as a combination of demand responsive and fixed route transit service, and where both passenger trips and vehicle trips must be scheduled. It should be remembered that in LITRES-2 a trip must not be scheduled as an integrated trip just because it is feasible. The simulation works mainly from the passengers’ point of view and simulate the traveler’s behavior and choice.

The fact that travelers choose to build up their journeys with the journey-legs that are most suitable for them, all requests cannot be coordinated in an optimal way from the operators point of view. If the traveler wishes to use more than one mode, the transfer between modes is made at the transfer point most suitable for the customer. This means that the requests handled by a demand responsive vehicle when planned from the operators point of view (as in Figure 6 in Chapter 2) actually would be handled by two vehicles (or routes) in LITRES-2, as described in Figure 18.
4.6 Comments Regarding LITRES-2

LITRES-2 is a useful tool for simulating several different scenarios, and especially useful for studies of different integrated transport services. For simulations of integrated public transport systems, LITRES-2 offers a unique opportunity to study markets response to changes in the range of transport services. The main thing missing in LITRES-2 is the possibility to plan an integrated system fully from the operator’s perspective. In such a perspective, the operator must have the possibility to choose which of the requests that should be performed as integrated journeys, and which that should be served by the demand responsive service the entire way from origin to destination.

During this evaluation we have used a UNIX-based version of LITRES-2. The software may be a bit too detailed and complicated for regional authorities and smaller companies in the public transportation market. For research projects within the area of public transportation, LITRES-2 can though be a useful tool. Some of the models working in the LITRES-2 software would certainly be interesting to see in a GIS-based environment. Since a lot of the input data needed for simulating public transport services many times (almost all the time) involve spatial references (also in LITRES-2) they are quite easy to handle in a GIS. The simplified environment that a GIS-based software offers would probably increase the number of potential users.
5 Simulations of an Integrated Service

As described in Section 4.5, there are several parameters that affect the attractiveness and efficiency of an integrated service. In this Chapter, the effects of customer service parameters and planning parameters are studied through simulation. The aim is to find guidelines to help operators of public transport to design a service that combines a fixed route bus service with a demand responsive service, resulting in an integrated public transport service. This is done by evaluating the effects on availability, travel time, cost and other service indicators for variations in the design and structure of the service. The goal is to design a transportation service that gives the regular public transport riders an increased service level, attracts some of those persons in hold of special needs permits, and contributes to an overall cost-efficient public transport. The simulations are done with the software LITRES-2, described in Chapter 4. When performing such simulations, there is a number of assumptions that must be made, such as travel times for different vehicle types, loading and unloading times for different market segments etc. These assumptions and other input to the simulations are explained in Section 5.1. Section 5.2 describes the performed simulations and their results. In Section 5.3 we make some concluding remarks about the outcome of the simulations.

5.1 The Gävle Case

For the simulations we have used data from the town of Gävle. Gävle is a medium sized Swedish town with about 90 000 inhabitants, situated on the east coast of Sweden 170 kilometers north of Stockholm.

5.1.1 Road Network

The network representing the town of Gävle consists of 4087 nodes and 9384 directed links. The area covered is approximately 130 km$^2$. The original network was extracted from the Swedish national road database. The description of the network was then modified in the way needed to represent it in LITRES-2.

5.1.2 Market Segments

Two different market segments are defined. The main differences between these segments are the maximum walking distance acceptable by the customers, and the walking speed. In the first segment a maximum walking distance of 1.0 km are used, while in the second segment 0.2 km is used. The values are assumed average of all people belonging to the market segment. In reality, of course, the maximum walking distance is distributed continuously since some customers accept even more than 1.0
Simulations of an Integrated Service

km and some less than 0.2 km. For the majority however, the maximum acceptable walking distance is somewhere in between. Walking speeds are set to 5.0 kilometers per hour for market segment 1, and 2.5 kilometers per hour for market segment 2. Just as in the case with maximum walking distances, assumptions are made for the average of all people in each market segment. Except for the maximum acceptable walking distances, the cost weights for the time-dependent parts of the generalized cost presented in section 3, also differ between the segments.

It is difficult to find suitable time weights for different parts of a journey. In Persson (2003) a comparison of two different studies by Sjöstrand (2001) and Norheim, & Stangeby (1993) show that different values for the time weights can be obtained also from cities equal in size. In Vägverket (1992) further different results are presented. To find reasonable time weights for the Gävle case, all these studies have been taken under consideration. The finally assumed costs for the two market segments are presented in Table 2.

<table>
<thead>
<tr>
<th>Journey part</th>
<th>Market segment 1</th>
<th>Market segment 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cost of waiting at origin</td>
<td>40</td>
<td>30</td>
</tr>
<tr>
<td>Cost of waiting at transfer</td>
<td>70</td>
<td>70</td>
</tr>
<tr>
<td>Cost of walking</td>
<td>60</td>
<td>100</td>
</tr>
<tr>
<td>Cost of transiting</td>
<td>50</td>
<td>30</td>
</tr>
</tbody>
</table>

5.1.3 Demand

The total demand of public transport services within the studied area comprises approximately 17 000 journeys per day. This demand is defined with respect to 104 zones. 90 percent of all these journeys are assumed to belong to the first market segment, and the remaining 10 percent to the second. The original demand was given by four OD-matrices representing different time intervals. To get a better accuracy, the demand has been broken down to 19 time intervals, each representing approximately one hour. For this we have used information from Holmberg, & Hydén (1996) that describes how the demand of public transportation normally is distributed over a day in Swedish towns of different sizes.

How the demand finally used, varies over the day can be viewed in Figure 19. All journey requests are sent exactly 30 minutes prior to the requested departure time.
5.1.4 Zones

In the simulations of test site Gävle, 104 centroid zones are used. These zones are the traffic prognosis zones used by the local authorities in Gävle. These original zones cover the entire town, and not only areas where people actually live and work, as can be seen in Figure 20. This is a problem since LITRES-2 divides the number of requests (origins and destinations) equally over an entire zone. This means that journeys will be requested also from and to uninhabited areas (such as parks, sport fields etc). To get around this problem, the zones have been changed to only cover the areas around the road network. For this reason, buffer zones of all areas within 30 meters distance from the road network have been created. The 104 zones finally used in the simulations covers only the area of these buffer zones. The zones have then been discretized, and the coordinates of the vertexes have been identified. All these steps have been made in a geographical information system (GIS). Thereafter a VBA code has been used to restructure the information in the way that it must be presented in LITRES-2. A part of the total area finally covered by the zones can be viewed in Figure 21, where the black fields are areas not part of any zones.

Figure 19: The distribution of requests over the day
Figure 20: The original centroid zones

Figure 21: Description of the final zones
5.1.5 Bus Network

In the town of Gävle there are three major (highly frequented) fixed route bus lines, as well as a number of less frequented. In the simulations, the highly frequented have been kept, while the less frequented have been neglected. So the fixed route network used in the simulations consists of three lines and 128 stops, see Figure 22. On tow-way streets, bus stops represents the bus stops in both directions, even if these are somewhat apart in reality. On one-way streets the bus stops are of course only used by buses in the stipulated direction. All buses are in LITRES-2 assumed to have infinite capacity. The real timetables for the bus lines have been used. All these lines are in service approximately from 5 am to 11.30 pm. The cost for using the bus is set to 15. If two different fixed routes are used in one single journey, the price is doubled, i.e. the cost is 15 for each fixed route leg of a journey.
Figure 22: The bus network of Gävle
5.1.6 Meeting Points

Normally, there are many practical details to take into account when deciding the number and location of stops (or as in this case meeting points). In this study however, the exact location and number of meeting points have not been considered.

The decision of where the meeting points should be located has been made in the following way. First of all, a grid of points, separated by 200 meters, has been put over the entire map. Thereafter, all points within a distance of 70 meters from any part of the road network have been selected, see Figure 23. In the Gävle case, this makes about 1000 selected meeting points.

Figure 23: Selection of meeting points
5.2 Performed simulations

The simulations have been made to estimate what effect different parameters have on the attractiveness and efficiency of the integrated service. The parameters that are assessed to be the most interesting when planning an integrated service, and that also are possible to test in LITRES-2, are the following:

- The number of demand responsive vehicles
- The capacity of demand responsive vehicles
- The number of transfer nodes
- Acceptable size of time windows for customers
- Acceptable travel factor of the demand responsive vehicles
- Pricing alternatives

In addition to these parameters, the following two more fundamental changes of the service have also been tested.

- Door-to-door service versus the use of meeting points
- The use of demand responsive vehicles without any fixed route service during time periods with low demand

To perform the simulations, a standard case has been defined in which all parameters have been set to specified values. Thereafter, one parameter at the time has been changed. In this way the effects of each parameter are made visible as clearly as possible. The standard case was based on the use of meeting points and the following parameter values.

- Number of demand responsive vehicles: 14
- Capacity of the demand responsive vehicles: 7
- Number of transfer nodes: 15 (3 kilometers in between)
- Acceptable time window for pick up: 100%
- Acceptable travel factor of the demand responsive vehicles: 200
- Fixed cost of the demand responsive service: 5
- Distance based cost of the demand responsive service: 10
The results have been evaluated from the following criteria:

- Number of requests accepted
- Generalized cost
- Monetary cost
- Average time consumption for a journey
- Number of journeys with at least one demand responsive leg (DR-leg)

For each criterion (except the number of demand responsive journeys) the results have been studied both in total and for each market segment.

### 5.2.1 Number of Demand Responsive Vehicles

How the number of demand responsive vehicles affect the service is of course interesting to know. Not the least to be able to make approximate assumptions of how many demand responsive vehicles that are needed in different situations. Tests on the number of demand responsive vehicles have been made in the interval from 1 to 30 vehicles.

The number of requests that can be accepted seems to depend linearly on the number of demand responsive vehicles used. This is not quite so. The number of accepted requests of course always increases with the number of demand responsive vehicles used, but with a slightly decreasing marginal effect. Within ± a few vehicles from the 14 used in the standard case, the relationship can be assumed linear. In Table 3, the results from the standard case as well as ± 5 vehicles are presented.
Table 3: Results from simulations of the number of demand responsive vehicles

<table>
<thead>
<tr>
<th>Number of vehicles</th>
<th>Market segment 1</th>
<th>Market segment 2</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accepted requests</td>
<td>14722</td>
<td>622</td>
<td>15344</td>
</tr>
<tr>
<td>Acceptance percentage</td>
<td>94,1</td>
<td>40,1</td>
<td>89,2</td>
</tr>
<tr>
<td>Generalized cost</td>
<td>40,8</td>
<td>58,2</td>
<td>41,5</td>
</tr>
<tr>
<td>Monetary cost</td>
<td>16,5</td>
<td>34,8</td>
<td>17,3</td>
</tr>
<tr>
<td>Time (s)</td>
<td>1740</td>
<td>1979</td>
<td>1750</td>
</tr>
<tr>
<td>Journeys including a DR-leg</td>
<td>-</td>
<td>-</td>
<td>916</td>
</tr>
</tbody>
</table>

Number of vehicles: 14

| Accepted requests | 14900 | 766 | 15666 |
| Acceptance percentage | 95,2 | 49,4 | 91,1 |
| Generalized cost   | 41,1  | 57,7 | 41,9 |
| Monetary cost      | 17,0  | 35,5 | 17,9 |
| Time (s)           | 1728  | 1846 | 1734  |
| Journeys including a DR-leg | - | - | 1359 |

Number of vehicles: 19

| Accepted requests | 15058 | 909 | 15967 |
| Acceptance percentage | 96,2 | 58,6 | 92,8 |
| Generalized cost   | 41,3  | 57,9 | 42,2 |
| Monetary cost      | 17,4  | 36,6 | 18,5 |
| Time (s)           | 1709  | 1735 | 1710  |
| Journeys including a DR-leg | - | - | 1823 |

5.2.2 Capacity of the Demand Responsive Vehicles

When planning any service it is important to know the effects of different vehicle types, especially the effects of different capacities. This is not least important when trying to estimate the economical benefits possible to be drawn from each vehicle type. In the simulations, the capacities of the demand responsive vehicles were 3, 7 and 14 (car, mini bus and midi bus).

The tests made on the capacity of the demand responsive vehicles indicate that vehicles with capacity 7 are enough. Tests with different vehicle capacities have also been made with higher travel factor than in the standard case. These tests show that when the acceptable time for a DR-leg of a journey is increased, there are some positive effects when using vehicles with capacity 14. But, most of the time these vehicles have a low level of usage.
5.2.3 Number of Transfer Nodes

The number and placement of transfer nodes is a parameter most interesting for the operator. It is often preferable for the operator if the number of transfer nodes is as low as possible. The transfer nodes should have a higher standard than the ordinary bus stops, and therefore there are economical reasons to keep the number of these nodes down.

The first transfer node to be established at each fixed route is the most central bus stop (the same for all routes). From this node, the rest of the transfer nodes are then placed more or less equally distributed along the routes. The distance between the transfer nodes have been set to (approximately) 2, 3 and 4 kilometers. The distances are approximate since only original bus stops are used and no new are added. When altering the distance between the transfer nodes in this way, the number of transfer nodes is altered between 10, 15 and 27. In addition to this, a test where transfers were allowed at all (128) bus stops has also been performed.

As mentioned, a relatively small number of transfer nodes would be preferable for the operator. The simulations show that a smaller number of transfer nodes work just as well, regarding the number of journey requests that can be accepted by the operator. The costs for the customers using the demand responsive service nevertheless increase somewhat, since the fare partially was based on distance during these tests. A lower number of transfer nodes gives the effect that a lower number of journeys involving the demand responsive service are requested from market segment one. The effect of this is that a higher number of requests from market segment two can be carried out, as can be seen in Table 4. An assumption can be made that the placement of the transfer nodes are maybe more important than the number of such nodes.
Table 4: Results from simulations of the number of transfer nodes (T-nodes)

<table>
<thead>
<tr>
<th>Number of T-nodes</th>
<th>Market segment 1</th>
<th>Market segment 2</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accepted requests</td>
<td>14859</td>
<td>811</td>
<td>15671</td>
</tr>
<tr>
<td>Acceptance percentage</td>
<td>95,0</td>
<td>52,3</td>
<td>91,1</td>
</tr>
<tr>
<td>Generalized cost</td>
<td>41,0</td>
<td>61,2</td>
<td>42,0</td>
</tr>
<tr>
<td>Monetary cost</td>
<td>16,9</td>
<td>38,6</td>
<td>18,0</td>
</tr>
<tr>
<td>Time (s)</td>
<td>1723</td>
<td>1877</td>
<td>1731</td>
</tr>
<tr>
<td>Journeys including a DR-leg</td>
<td>-</td>
<td>-</td>
<td>1390</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Number of T-nodes</th>
<th>Market segment 1</th>
<th>Market segment 2</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accepted requests</td>
<td>14900</td>
<td>766</td>
<td>15666</td>
</tr>
<tr>
<td>Acceptance percentage</td>
<td>95,2</td>
<td>49,4</td>
<td>91,1</td>
</tr>
<tr>
<td>Generalized cost</td>
<td>41,1</td>
<td>57,7</td>
<td>41,9</td>
</tr>
<tr>
<td>Monetary cost</td>
<td>17,0</td>
<td>35,5</td>
<td>17,9</td>
</tr>
<tr>
<td>Time (s)</td>
<td>1728</td>
<td>1846</td>
<td>1734</td>
</tr>
<tr>
<td>Journeys including a DR-leg</td>
<td>-</td>
<td>-</td>
<td>1359</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Number of T-nodes</th>
<th>Market segment 1</th>
<th>Market segment 2</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accepted requests</td>
<td>14912</td>
<td>769</td>
<td>15681</td>
</tr>
<tr>
<td>Acceptance percentage</td>
<td>95,3</td>
<td>49,6</td>
<td>91,2</td>
</tr>
<tr>
<td>Generalized cost</td>
<td>41,0</td>
<td>55,0</td>
<td>41,7</td>
</tr>
<tr>
<td>Monetary cost</td>
<td>17,0</td>
<td>33,0</td>
<td>17,8</td>
</tr>
<tr>
<td>Time (s)</td>
<td>1724</td>
<td>1812</td>
<td>1729</td>
</tr>
<tr>
<td>Journeys including a DR-leg</td>
<td>-</td>
<td>-</td>
<td>1409</td>
</tr>
</tbody>
</table>

5.2.4 Time Windows

In many situations the size of the time windows for pick up (or drop off) can be a very important factor, both when planning the overall service as well as the vehicle itineraries. In LITRES-2, the size of time windows for picking up customers is given as a percentage of the shortest possible travel time. The size of acceptable time windows at pick up, have therefore been tested in the interval from 0 percent to 100 percent.

The effects of changes to the acceptable size of time windows are not so easily studied, especially not when the journey-legs with the demand responsive service are so short as they are in an integrated service like this. In LITRES-2, 100 percent of shortest possible travel time is the maximum time window. For short journey-legs, that time is to small to see any effect of the time window size. With the possibility to study larger time windows it is quite likely that effects would be quite large. If the operator is given the possibility to decide what transfer nodes the customers should use (not possible in the LITRES-2 model) the effects of the time window size would be even larger.
5.2.5 Travel Factor of the Demand Responsive Service

The travel factor is the maximum allowed time for a DR-leg as percentage of the shortest time possible. This parameter is of great importance when planning any demand responsive service (or integrated service). How much DR-legs are allowed to be prolonged of course influence how the itineraries can be put together. An increase in acceptance from the customers regarding the time a DR-leg of a journey is allowed to take, result in much better planning possibilities for the operator. Tests on the travel factor have been carried out with values of 200, 300 and 400.

The performed simulations have shown that the travel factor affects the efficiency of the service. When increasing the travel factor by 50 percent, i.e. from 200 to 300, the number of journeys including the demand responsive service, is increased by 5 percent, see Table 5. At the same time the total average travel time for all those using the demand responsive service, increased less than one minute (55 seconds). To increase the travel factor further, to 400, does not give the same effect.

Table 5: Results from simulations of the travel factor of the demand responsive service

<table>
<thead>
<tr>
<th>Travel factor: 200</th>
<th>Market segment 1</th>
<th>Market segment 2</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accepted requests</td>
<td>14900</td>
<td>766</td>
<td>15666</td>
</tr>
<tr>
<td>Acceptance percentage</td>
<td>95.2</td>
<td>49.4</td>
<td>91.1</td>
</tr>
<tr>
<td>Generalized cost</td>
<td>41.1</td>
<td>57.7</td>
<td>41.9</td>
</tr>
<tr>
<td>Monetary cost</td>
<td>17.0</td>
<td>35.5</td>
<td>17.9</td>
</tr>
<tr>
<td>Time (s)</td>
<td>1728</td>
<td>1846</td>
<td>1734</td>
</tr>
<tr>
<td>Journeys including a DR-leg</td>
<td>-</td>
<td>-</td>
<td>1359</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Travel factor: 300</th>
<th>Market segment 1</th>
<th>Market segment 2</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accepted requests</td>
<td>14927</td>
<td>836</td>
<td>15764</td>
</tr>
<tr>
<td>Acceptance percentage</td>
<td>95.4</td>
<td>53.9</td>
<td>91.7</td>
</tr>
<tr>
<td>Generalized cost</td>
<td>41.2</td>
<td>59.1</td>
<td>42.1</td>
</tr>
<tr>
<td>Monetary cost</td>
<td>17.0</td>
<td>35.9</td>
<td>18.0</td>
</tr>
<tr>
<td>Time (s)</td>
<td>1732</td>
<td>1923</td>
<td>1742</td>
</tr>
<tr>
<td>Journeys including a DR-leg</td>
<td>-</td>
<td>-</td>
<td>1422</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Travel factor: 400</th>
<th>Market segment 1</th>
<th>Market segment 2</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accepted requests</td>
<td>14925</td>
<td>868</td>
<td>15793</td>
</tr>
<tr>
<td>Acceptance percentage</td>
<td>95.4</td>
<td>56.0</td>
<td>91.8</td>
</tr>
<tr>
<td>Generalized cost</td>
<td>41.2</td>
<td>58.3</td>
<td>42.1</td>
</tr>
<tr>
<td>Monetary cost</td>
<td>16.9</td>
<td>34.7</td>
<td>17.9</td>
</tr>
<tr>
<td>Time (s)</td>
<td>1739</td>
<td>1948</td>
<td>1751</td>
</tr>
<tr>
<td>Journeys including a DR-leg</td>
<td>-</td>
<td>-</td>
<td>1422</td>
</tr>
</tbody>
</table>
5.2.6 Pricing Alternatives

The way the pricing is set is of course very important to how popular the service becomes. This is a very important factor to consider, particularly for being able to interpret other results in the right way. Tests of pricing alternatives have only been made on the demand responsive service. The fixed route service has the same pricing throughout all tests. Regarding the demand responsive service, the price consists of two parts, one fixed cost and one distance based cost. The tests concern both these costs.

The main problem regarding the pricing is that if the cost is set too low (especially concerning the distance based part of the cost) then the increase in popularity bring the effect that quite a lot chooses to use the demand responsive service even though they do not need to. Meaning that those in need of the service, primarily customers from the second market segment do not get access to it, see Table 6.

<table>
<thead>
<tr>
<th>Fare base: 0, Fare rate: 15</th>
<th>Market segment 1</th>
<th>Market segment 2</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accepted requests</td>
<td>14849</td>
<td>827</td>
<td>15676</td>
</tr>
<tr>
<td>Acceptance percentage</td>
<td>94.9</td>
<td>53.3</td>
<td>91.2</td>
</tr>
<tr>
<td>Generalized cost</td>
<td>41.1</td>
<td>62.1</td>
<td>42.2</td>
</tr>
<tr>
<td>Monetary cost</td>
<td>17.0</td>
<td>40.1</td>
<td>18.2</td>
</tr>
<tr>
<td>Time (s)</td>
<td>1729</td>
<td>1822</td>
<td>1734</td>
</tr>
<tr>
<td>Journeys including a DR-leg</td>
<td>-</td>
<td>-</td>
<td>1311</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Fare base: 5, Fare rate: 10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accepted requests</td>
</tr>
<tr>
<td>Acceptance percentage</td>
</tr>
<tr>
<td>Generalized cost</td>
</tr>
<tr>
<td>Monetary cost</td>
</tr>
<tr>
<td>Time (s)</td>
</tr>
<tr>
<td>Journeys including a DR-leg</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Fare base: 25, Fare rate: 0</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accepted requests</td>
</tr>
<tr>
<td>Acceptance percentage</td>
</tr>
<tr>
<td>Generalized cost</td>
</tr>
<tr>
<td>Monetary cost</td>
</tr>
<tr>
<td>Time (s)</td>
</tr>
<tr>
<td>Journeys including a DR-leg</td>
</tr>
</tbody>
</table>
5.2.7 Door-to-door Versus the use of Meeting Points

If the demand responsive service picks up and drops off customers at the exact location of their origins and destinations instead of at meeting points close to these locations, this of course gives a better service for the customers. So when planning a demand responsive service these two systems should be compared to see how much this increase in service costs.

Whether the demand responsive service uses meeting points or pick up and drop off passengers at their exact addresses does not seem to give any major differences on the results. This indicate that the higher service level provided by the door-to-door service can be offered without any noticeable loss in efficiency, see Table 7. If the operator is given the possibility to decide which transfer nodes the customers should use (not possible in the LITRES-2 model) the use of meeting points could be more beneficial.

Table 7: Results from simulations of door-to-door service, versus the use of meeting points

<table>
<thead>
<tr>
<th>Use of meeting points: Yes</th>
<th>Market segment 1</th>
<th>Market segment 2</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accepted requests</td>
<td>14900</td>
<td>766</td>
<td>15666</td>
</tr>
<tr>
<td>Acceptance percentage</td>
<td>95,2</td>
<td>49,4</td>
<td>91,1</td>
</tr>
<tr>
<td>Generalized cost</td>
<td>41,1</td>
<td>57,7</td>
<td>41,9</td>
</tr>
<tr>
<td>Monetary cost</td>
<td>17,0</td>
<td>35,5</td>
<td>17,9</td>
</tr>
<tr>
<td>Time (s)</td>
<td>1822</td>
<td>1846</td>
<td>1734</td>
</tr>
<tr>
<td>Journeys including a DR-leg</td>
<td>-</td>
<td>-</td>
<td>1359</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Use of meeting points: No</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accepted requests</td>
</tr>
<tr>
<td>Acceptance percentage</td>
</tr>
<tr>
<td>Generalized cost</td>
</tr>
<tr>
<td>Monetary cost</td>
</tr>
<tr>
<td>Time (s)</td>
</tr>
<tr>
<td>Journeys including a DR-leg</td>
</tr>
</tbody>
</table>
5.2.8 Demand Responsive Service without any Fixed Routes

The simulations of how the demand responsive service would cope without any fixed route service during time periods with low demand (early morning and late evening) do point out some difficulties with this. The main problem is that the demand is spread over the entire town, but is quite sparse. This makes it difficult to coordinate the journeys in a way acceptable for the customers. The journey requests with origin or destination in the outskirts of the town involve longer journeys and an even more sparse demand. As a result of this the operator of the demand responsive service prioritizes the demand in the central areas, and thereby giving the customers in the outskirts of the town an even lower level of service.

Forcing the operator to accept journey requests from the outskirts, would probably lead to a lower level of coordination between journeys and therefore a lower level of usage of the vehicles. The total demand is still so large, and sparse, that it would require a large number of demand responsive vehicles. The use of a demand responsive service without a fixed route service would yet be possible to operate effectively in smaller towns. But in towns of the size of Gävle, it does not seem suitable.

5.3 Some Comments about the Results of the Simulations

The overall small differences in results obtained from the different simulations is probably mainly due to two things. First of all that the customers are allowed to decide what transfer nodes they wish to use, described by the difference of Figure 6 and Figure 18. This reduces the planning possibilities for the operator. If the operator of the integrated service is given the right to decide at which transfer nodes the customers should be transferred between the demand responsive service and the fixed route, and vice versa, as described in Section 2, substantial effects could be assumed. Secondly, the time windows for pick up of customers have not been possible to set to large enough sizes. Small time windows of course also reduce the possibilities for the operator to plan good itineraries.

To let the demand responsive vehicles operate freely, without being allotted to any specific part of the town doesn’t seem so successful. During time periods with high demand, the demand is denser in the central part of the town. By this reason the operator of the demand responsive vehicles chooses to satisfy as much as possible of the demand in that area. This means that the customers in the outskirts of the town are given a reduced access to the service. This is especially clear when operating the demand responsive service without any fixed route service available.

The method of simulating the effects of one single parameter at the time was quite important. The effects of changing the parameters were overall quite small and could therefore have been very hard to detect if the simulations had been performed in an other way.
6 An Exact Model for IDARP

In planning a framework for integrated public transport, as the one described in Chapter 3, an optimization model performing the actual planning of the journeys is an essential tool. This chapter presents an exact model formulation, describing the planning of journeys in an integrated public transport system. The service combines dial-a-ride and fixed route services, and the problem of scheduling journeys in such a system will therefore from here on be referred to as the integrated dial-a-ride problem (IDARP). In the described planning process of Chapter 2, this is the problem of assigning passengers to vehicles. For large problem instances, this exact formulation take very long time to solve, but for strategic planning of small instances it can be useful.

In Section 6.1, the mathematical model for planning of the IDARP is presented. Section 6.2 describes methods to strengthen this formulation in order to make the problem easier to solve. Section 6.3 illustrates a numerical example and describes how input and output data can be handled in a GIS. Finally, in Section 6.4 some conclusions regarding this exact formulation of the IDARP are presented.

6.1 Model Formulation

We are given a number of requested journeys that are to be scheduled. Each request has a specified pick-up location (node) and a drop-off location (node), as well as a time window specified for one of these nodes, or both. Further more, each request also has a maximum ride time. The requests can be carried out by a dial-a-ride service, but some part of each journey can be carried out by a fixed route service. This means that the passenger can either be carried by the dial-a-ride service from the pick-up node directly to the drop-off node, or from the pick-up node to a transfer node, where the passenger can change to a fixed route service. If the passenger is carried to a transfer node, another vehicle in the fleet of dial-a-ride vehicles must pick up the passenger at another transfer node and carry the passenger to its drop-off node. Each request contains one or several passengers and requires a certain capacity in a vehicle, for the persons and any wheelchairs, walkers, luggage etc. This needed capacity generates a load at each node describing the capacity required if the request is picked-up at the node, or made available if the request is dropped-off, at that node. The fleet of dial-a-ride vehicles is assumed to be homogeneous. All vehicles have the same capacity, the same maximum duration time (time from when the vehicle leaves the depot until it returns to the depot) and they originate from the same depot.

The formulation has an underlying network structure. This problem is based on the dial-a-ride problem (DARP), described in Section 2.5.2, and therefore we have based our IDARP formulation on the directed graph formulation of the DARP, described in Cordeau (2006). The main expansion of the original DARP-formulation is that in the IDARP both vehicle itineraries and customer itineraries are scheduled.
In the directed graph $G = (N, A)$, $N$ is the set of all nodes, including pick-up nodes, drop-off nodes, depot nodes and transfer nodes, and $A$ is the set of arcs connecting the nodes. Each arc $(i, j) \in A$ has an associated cost $c_{ij}$ and a travel time $t_{ij}$. Among all the nodes in $N$, the subset $P = \{1, \ldots, n\}$ contains the pick-up nodes for the $n$ requests, and $D = \{n + 1, \ldots, 2n\}$ is the corresponding set of drop-off nodes. In this way the node $n + i$ is the drop-off node associated with pick-up node $i$. For each geographical transfer location there is a separate transfer node for each request. The total number of transfer nodes is therefore equal to the number of geographical locations, $g$, times the number of requests. These nodes form the subset $C = \{2n+2, \ldots, 2n+1+ng\}$ of $N$. The set $C^r$, is the set of nodes in set $C$ associated with request $r$. The origin depot is node 0 and the destination depot is node $2n + 1$. Finally, to each node there is a time window $[e_i, l_i]$ associated.

The set $C$ makes the main difference between the formulation of the DARP and the formulation of IDARP. In the DARP all nodes $i \in N$ are visited, but for each request $r$ in the IDARP at most two transfer nodes in the set $C^r$ are visited. In the DARP, the same vehicle must visit both the pick-up node and the delivery node belonging to a request, which is not the case in the IDARP.

We introduce the following notation:

**Sets:**
- $R$: set of requests
- $P$: set of pick-up nodes
- $D$: set of drop-off nodes
- $C$: set of transfer nodes
- $C^r$: set of transfer nodes associated with request $r$
- $N$: set of all nodes, pick-up, drop-off, depot and transfer nodes
- $K$: set of demand responsive vehicles

**Variables:**

$x_{ij}^k = \begin{cases} 1, & \text{if vehicle } k \text{ travels from node } i \text{ to node } j \\ 0, & \text{otherwise} \end{cases}$

$y_{ij}^r = \begin{cases} 1, & \text{if request } r \text{ travels by dial-a-ride from node } i \text{ to node } j \\ 0, & \text{otherwise} \end{cases}$

$z_{ij} = \begin{cases} 1, & \text{if the fixed route between transfer nodes } i \text{ and } j \text{ is used} \\ 0, & \text{otherwise} \end{cases}$

$B_k^k = \text{the time at which vehicle } k \text{ leaves the depot}$

$B_i = \text{the time node } i \text{ is reached}$
$D_i = \text{time from which a vehicle leaves the depot, until it arrives at node } i$

$w_i = \text{waiting time at node } i \text{ of the vehicle visiting node } i$

**Parameters:**

$Q = \text{capacity of a dial-a-ride vehicle}$

$T = \text{maximum duration time of a dial-a-ride vehicle}$

$c_{ij} = \text{cost for any vehicle to travel from node } i \text{ to node } j$

$t_{ij} = \text{travel time from node } i \text{ to node } j$

$q^r = \text{load of request } r$

$d_i = \text{service duration at node } i$

$e_i = \text{earliest time at which service may begin at node } i$

$l_i = \text{latest time at which service may begin at node } i$

$f^r_i = \begin{cases} 
1, & \text{if node } i \text{ is the pick-up node of request } r \\
-1, & \text{if node } i \text{ is the drop-off node of request } r \\
0, & \text{otherwise} 
\end{cases}
$

$\bar{L}_i = \text{maximum ride time of request } i$

$M = \text{large positive number}$

The mathematical formulation becomes:

$$\min \sum_{k \in K} \sum_{i \in N} \sum_{j \in N} c_{ij} x_{ij}^k \quad (1)$$

Subject to:

$$\sum_{k \in K} \sum_{j \in N} x_{ij}^k = 1 \quad i \in P \cup D \quad (2)$$

$$\sum_{j \in N} x_{0j}^k = 1 \quad k \in K \quad (3)$$

$$\sum_{j \in N} x_{ji}^k - \sum_{j \in N} x_{ij}^k = 0 \quad i \in P \cup D \cup C, k \in K \quad (4)$$

$$\sum_{i \in N} x_{i,2n+1}^k = 1 \quad k \in K \quad (5)$$

$$\sum_{j \in N} y_{ij}^r + \sum_{j \in C^r} z_{ij} - \sum_{j \in N} y_{ji}^r - \sum_{j \in C^r} z_{ji} = f_i^r \quad r \in R, i \in C^r \quad (6)$$

$$\sum_{j \in N} y_{ij}^r - \sum_{j \in C^r} y_{ji}^r = f_i^r \quad i \in N \setminus C^r, r \in R \quad (7)$$

$$\sum_{r \in R} q^r y_{ij}^r \leq Q \sum_{k \in K} x_{ij}^k \quad i \in N, j \in N, (i, j) \notin (C \times C) \quad (8)$$

$$B_j \geq B_j^k + t_{0j} - M(1 - x_{0j}^k) \quad j \in N, k \in K \quad (9)$$

$$B_j \geq B_i + d_i + w_i + t_{ij} - M(1 - \sum_{k \in K} x_{ij}^k) \quad i \in N, j \in N \quad (10)$$

$$B_j \geq B_i + w_i + t_{ij} - M(1 - z_{ij}) \quad (i, j) \in (C^r \times C^r), r \in R \quad (11)$$
The objective is to minimize the total cost of the dial-a-ride vehicles. Constraint (2) says that all pick-up and drop-off nodes must be visited. Constraints (3)–(5) ensure that each route starts at the depot, that a vehicle leaves each node it arrives at and that each route ends at the depot. Constraints (6)–(7) describe the node balance of the requests. In constraint (6), describing the transfer nodes, passengers can arrive and depart both with dial-a-ride vehicles and by the fixed route. In constraint (7) only the dial-a-ride vehicles have to be considered. Constraint (8) ensures that the sum of the passengers in all requests traveling between node \( i \) and node \( j \) may not be greater than the capacity of the vehicles. This also implies that if no vehicle travels between the nodes, no passengers can travel between them. Constraints (9)–(11) describe when each node can be reached. Constraint (12) makes sure that all nodes are serviced within their time window. Constraint (13) guarantees that the ride time is less than the maximum ride time allowed for each request. Constraints (14)–(15) guarantee that the duration times of the vehicles are lower than the maximum duration time.

### 6.2 Strengthening the Mathematical Model

The model presented in Sect. 6.1 grows very rapidly in size with the number of requested journeys. Each request adds \( 2 + g \) nodes, where \( g \) is the number of transfer nodes, to the network and also adds all the associated arcs. In order to solve larger instances of the model, it is very important to be able to reduce the size of the network. In Section 6.2.1, we present some arc elimination rules that reduce the number of arcs, i.e. the number of variables. We also show, in Section 6.2.2, how the number of binary variables can be reduced by introducing a new set of variables and how subtour elimination constraints can be added to the model.
6.2.1 Arc Elimination

Many of the arcs in the graph $G$ represent infeasible transportation opportunities. The requirements on time windows and maximum ride times make these arcs impossible to use in a feasible schedule. Therefore, we can eliminate these arcs, which in our formulation directly lead to a reduction of the number of binary variables. We first apply some arc elimination rules previously defined for the DARP, described for example in Cordeau (2006). We will then present some more elimination rules, specific for the IDARP.

Some arcs connected to depot nodes can always be eliminated. These are the following:

- No arc can go to the node representing the origin depot, i.e. all arcs $(i, 0)$ are infeasible for $i \in N$.
- No arc can go from the node representing the destination depot, i.e. all arcs $(2n + 1, i)$ are infeasible for $i \in N$.
- No arc can go from the origin depot to a drop-off node $i$, i.e. all arcs $(0, n + i)$ are infeasible for $i \in P$.
- No arc can go from a pick-up node $i$ to the destination depot, i.e. all arcs $(i, 2n + 1)$ are infeasible for $i \in P$.

There are some further elimination rules valid both in DARP and IDARP. These rules refer to arcs between nodes in the same request and arcs that can be eliminated due to time constraints.

- No arc can end at the same node as it starts, i.e. all arcs $(i, i)$ are infeasible for $i \in N$.
- No arc can go from a drop-off node to the pick-up node of the same request, i.e. all arcs $(n + i, i)$ are infeasible for $i \in P$.
- If $e_i + d_i + t_{ij} > l_j$ for $i, j \in N$, the arc $(i, j)$ is infeasible.
- If $t_{ij} + d_j + t_{j,n+i} > \bar{L}_i$ for $i \in P, j \in N$, the arcs $(i, j)$ and $(j, n + i)$ are infeasible.
- If the path $\hat{P} = \{ j, i, n + j, n + i \}$ is infeasible, so is the arc $(i, n + j)$, with $i, j \in P$.
- If the path $\hat{P} = \{ i, n + i, j, n + j \}$ is infeasible, so is the arc $(n + i, j)$, with $i, j \in P$. 

71
An Exact Model for IDARP

In an IDARP there are some more arcs that can be eliminated, all of these are in some way connected to at least one transfer node. For some of these elimination rules to be valid, we have assumed that the fixed route service never travels faster then the dial-a-ride service. In the rules presented below, \( S_i \) is any transfer node belonging to request \( i \).

- In the same way as a pick-up node \( i \) cannot be visited after the drop-off node of the same request, a transfer node for the request, cannot be visited after the drop-off node. Nor can a pick-up node be visited after a transfer node. All arcs \((n + i, S_i)\) and \((S_i, i)\) are therefore infeasible for \( i \in P \).

- If \((i, n + j)\) is infeasible, all arcs \((i, S_j)\) are also infeasible, for \( i \in N, j \in P \).

- If \(i, j \in P\), all arcs \((S_i, S_j)\), \((S_i, n + j)\) and \((S_i, j)\) are infeasible if \((i, n + j)\) is infeasible.

- There is never any reason to why a dial-a-ride vehicle should travel directly between two transit nodes associated with the same request, i.e. two nodes in \( C^r \). By this reason all arcs \((S_i, S_i)\) are infeasible.

There are also some commonly used elimination rules valid for the DARP that are not valid for the IDARP, simply because it is not certain that the same vehicle that picks up a request at the pick-up node also drops the request off at the drop-off node.

### 6.2.2 Variable Substitution and Subtour Elimination

Since the fleet of vehicles is homogeneous and uses the same depot, and each node can only be visited once, it is not necessary to explicitly know which vehicle that visits each node. As long as the information about what arcs that are used is available, it is possible to build up all the vehicle itineraries anyway. This makes it possible to reduce the number of binary variables substantially. In order to do this, we introduce the following new variable:

\[
v_{ij} = \begin{cases} 
1, & \text{if a vehicle travels from node } i \text{ to node } j \\
0, & \text{otherwise} 
\end{cases}
\]

Then the variable \( x_{ij}^k \) does not have to be binary. It is now enough if \( x_{ij}^k \) is non-negative. In this way we reduce the number of binary variables with a factor equal to the number of vehicles. By adding constraint (19) and (20), and replacing constraint (2) and (16) with (21) and (22) respectively, the model is ready to be solved with a significantly lower number of binary variables.
\begin{align}
\sum_{k \in K} x_{ij}^k &= v_{ij} \quad i, \in N, j \in N \\
v_{ij} &\in \{0, 1\} \quad i, \in N, j \in N \\
\sum_{j \in N} v_{ij} &= 1 \quad i \in P \cup D \\
x_{ij}^k &\geq 0 \quad i, \in N, j \in N, k \in K
\end{align}

The final way, in which we try to make the model easier to solve, is by adding constraints that eliminates subtours in the LP-relaxation of the problem. The solution to the LP-relaxation of the problem is used to compute a lower bound on the optimal objective function value of the original problem. We want the objective value of the LP-solution to be as high as possible. In order to increase the objective value of the LP-solution, and thereby the lower bound of the original problem, constraints eliminating subtours can be added.

In a feasible solution, only one vehicle can travel between any pair of nodes, and only in one direction. This implies that between any two nodes in \(N\), we simply have that the sum of the flow \((i, j)\) and \((j, i)\) must be less than or equal to one, as described in constraint (23).

\begin{align}
v_{ij} + v_{ji} &\leq 1 \quad (i, j) \in A
\end{align}

We also identify clusters with a total inflow of vehicles less than one. Since at least one vehicle must visit each such cluster in a feasible solution, one can add a constraint saying that each such cluster must have an inflow (or outflow) greater than or equal to one. If \(S\) is the set of all identified subtours (clusters) and \(S\) is one such subtour, constraint (24) describes how these subtours can be eliminated.

\begin{align}
\sum_{i \in S} \sum_{j \in N \setminus S} v_{ij} &\geq 1 \quad S \in S
\end{align}

Constraint (24) says that for each identified cluster, the outflow of vehicles from the nodes in the cluster to other nodes must be greater than or equal to one.
6.3 An Illustrated Example

We have solved the model described in Sect. 6.1 for several small instances. The modeling language has been AMPL, and CPLEX 7.0 has been used to solve it to optimality. So far, problem instances up to 10 requests have been solved. All tests were performed on a Sun Enterprise 450 computer with 3074 Mb of memory and four 400 MHz processors. No parallel computing is performed, so only one of the processors has been used. This section will describe one of the test cases, including four requests, three transfer locations and two dial-a-ride vehicles. This section also explains how the input and results can be visualized in a GIS.

A GIS can be of great help for creating input files to any kind of routing problem. It can also be a useful tool for displaying the results when the problem is solved. The simplicity in visualizing network based problems and the handling of such data are two of the advantages of using a GIS for this kind of problem. But the most important advantage is perhaps that most of the input data already is available in this format. Road networks, the location of fixed route services and customer registers are available (at least for all areas in Sweden).

The GIS used is ArcGIS 9.1 with the network analyst extension. All the nodes $i \in N$ have been pin placed, with the small exception that for the transfer nodes it is enough to place one node at each geographic location, there is no need of creating all the nodes in $C$. Each kind of node (pick-up, drop-off, depot and transfer) has been created in its own layer. All interesting and relevant information regarding the nodes, for example time windows, have been stored as attributes of each feature in the layers, in this way simplifying the handling of the information.

The input to our test case can be seen in Figure 24. The test case includes four requests. The pick-up nodes are therefore represented with the numbers 1 – 4 and the drop-off nodes with numbers 5 – 8. The origin depot has number 0 and the destination depot has number 9. The three transfer locations are represented with number 10 – 12. According to how the set $C$ is defined, there are one node for each request at each geographical transfer location. So for each number representing a transfer location in the figures, there are actually four nodes on top of each other, with the exact same geographical position.
A road network has been used to create an OD-cost matrix (available by the use of the network analyst extension) with all the relations between all nodes. The calculations of the cost of the different relations have been made directly in the GIS. The travel time of each link in the network has been used as input for these calculations. Consequently, the values in the matrix represent travel times, and they are presented in Table 8. Other measures of costs can of course be used, and calculated in the same way.
Table 8: OD-cost matrix of the test case, created in a GIS

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>23</td>
<td>173</td>
<td>131</td>
<td>195</td>
<td>170</td>
<td>129</td>
<td>160</td>
<td>153</td>
<td>0</td>
<td>120</td>
<td>21</td>
<td>163</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>166</td>
<td>123</td>
<td>173</td>
<td>179</td>
<td>125</td>
<td>138</td>
<td>168</td>
<td>23</td>
<td>135</td>
<td>26</td>
<td>159</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>173</td>
<td>0</td>
<td>57</td>
<td>289</td>
<td>41</td>
<td>268</td>
<td>254</td>
<td>79</td>
<td>173</td>
<td>57</td>
<td>191</td>
<td>301</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>131</td>
<td>123</td>
<td>57</td>
<td>0</td>
<td>258</td>
<td>72</td>
<td>236</td>
<td>222</td>
<td>60</td>
<td>131</td>
<td>29</td>
<td>148</td>
<td>270</td>
</tr>
<tr>
<td>4</td>
<td>195</td>
<td>173</td>
<td>289</td>
<td>258</td>
<td>0</td>
<td>318</td>
<td>73</td>
<td>35</td>
<td>306</td>
<td>195</td>
<td>273</td>
<td>174</td>
<td>39</td>
</tr>
<tr>
<td>5</td>
<td>170</td>
<td>179</td>
<td>41</td>
<td>72</td>
<td>318</td>
<td>0</td>
<td>296</td>
<td>282</td>
<td>53</td>
<td>170</td>
<td>51</td>
<td>191</td>
<td>329</td>
</tr>
<tr>
<td>6</td>
<td>129</td>
<td>125</td>
<td>268</td>
<td>236</td>
<td>73</td>
<td>296</td>
<td>0</td>
<td>48</td>
<td>282</td>
<td>129</td>
<td>248</td>
<td>108</td>
<td>34</td>
</tr>
<tr>
<td>7</td>
<td>160</td>
<td>138</td>
<td>254</td>
<td>222</td>
<td>35</td>
<td>282</td>
<td>48</td>
<td>0</td>
<td>271</td>
<td>160</td>
<td>237</td>
<td>139</td>
<td>57</td>
</tr>
<tr>
<td>8</td>
<td>153</td>
<td>168</td>
<td>79</td>
<td>60</td>
<td>306</td>
<td>53</td>
<td>282</td>
<td>271</td>
<td>0</td>
<td>153</td>
<td>35</td>
<td>174</td>
<td>316</td>
</tr>
<tr>
<td>9</td>
<td>0</td>
<td>23</td>
<td>173</td>
<td>131</td>
<td>195</td>
<td>170</td>
<td>129</td>
<td>160</td>
<td>153</td>
<td>0</td>
<td>120</td>
<td>21</td>
<td>163</td>
</tr>
<tr>
<td>10</td>
<td>120</td>
<td>135</td>
<td>57</td>
<td>29</td>
<td>273</td>
<td>51</td>
<td>248</td>
<td>237</td>
<td>35</td>
<td>120</td>
<td>0</td>
<td>141</td>
<td>283</td>
</tr>
<tr>
<td>11</td>
<td>21</td>
<td>26</td>
<td>191</td>
<td>148</td>
<td>174</td>
<td>191</td>
<td>108</td>
<td>139</td>
<td>174</td>
<td>21</td>
<td>141</td>
<td>0</td>
<td>142</td>
</tr>
<tr>
<td>12</td>
<td>163</td>
<td>159</td>
<td>301</td>
<td>270</td>
<td>39</td>
<td>329</td>
<td>34</td>
<td>57</td>
<td>316</td>
<td>163</td>
<td>283</td>
<td>142</td>
<td>0</td>
</tr>
</tbody>
</table>

Since we have four requests and three transfer locations (leading to 12 nodes in the set \( C \)) we have a total of 22 nodes in set \( N \). These 22 nodes give a total of 484 original arcs. By the arc elimination rules described in Sect. 6.2.1, we have reduced this number to 359 arcs. How many arcs that can be eliminated of course depends very much on how the time windows are defined. More narrow time windows and more scattered requests, both in time and in space, gives a larger number of arcs that can be eliminated. In our test case, the time windows have been defined as in Table 9, where each request has a time window defined for their pick-up node, but none for the drop-off node. For the drop-off nodes, the time windows have simply been calculated as:

\[
e_{n+i} = e_i + d_i + t_{i,n+i} \quad (25)
\]

\[
l_{n+i} = l_i + \bar{L}_i \quad (26)
\]

The earliest time a drop-off node can be serviced is described in equation (25), saying that this is equal to the earliest time the pick-up node of that request could be serviced, plus the service time at that node and the lowest possible travel time to the drop-off node. The latest time a drop-off node can be serviced, as described in equation (26), is the sum of the latest possible time for servicing the associated pick-up node and the maximum ride time of the associated request.
Table 9: Time windows for pick-up nodes $i \in P$

<table>
<thead>
<tr>
<th>$i$</th>
<th>$e_i$</th>
<th>$l_i$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>50</td>
<td>950</td>
</tr>
<tr>
<td>2</td>
<td>350</td>
<td>1250</td>
</tr>
<tr>
<td>3</td>
<td>650</td>
<td>1550</td>
</tr>
<tr>
<td>4</td>
<td>950</td>
<td>1850</td>
</tr>
</tbody>
</table>

The solution to this test case is presented in Table 10. The table also shows which customers that embark or disembark the vehicles at each node. An illustration of the solution is presented in Figure 25.

Table 10: Optimal solution to the test case

<table>
<thead>
<tr>
<th>Vehicle 1</th>
<th>Vehicle 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Node</td>
<td>Node type</td>
</tr>
<tr>
<td>0</td>
<td>depot</td>
</tr>
<tr>
<td>1</td>
<td>pick-up</td>
</tr>
<tr>
<td>3</td>
<td>pick-up</td>
</tr>
<tr>
<td>2</td>
<td>pick-up</td>
</tr>
<tr>
<td>5</td>
<td>drop-off</td>
</tr>
<tr>
<td>10</td>
<td>transfer</td>
</tr>
<tr>
<td>8</td>
<td>drop-off</td>
</tr>
<tr>
<td>9</td>
<td>depot</td>
</tr>
</tbody>
</table>

As can be seen in Table 10, vehicle one departs from the depot to pick-up request 1, 3 and 2. Thereafter request 1 is dropped off at its destination, and the vehicle proceeds to transfer location 1 (node 10) where request 2 and 3 are transferred to a fixed route service. At the same transfer location, the vehicle then waits for request 4 to arrive with the fixed route service. During this time vehicle two picks up request 2 and 3 at transfer location 2 (node 11) and takes them to their destinations. On this route, request 4 is picked up and taken to transfer location 2, and vehicle two returns to the depot. Vehicle one picks up request 4 at transfer location 1 and proceed to the requested destination. Thereafter also vehicle one can return to the depot. We can notice that one vehicle gets a quite long waiting time at transfer location 1, caused by the fact that only the arcs have costs in this formulation.
When the optimal routes have been computed they are fed back into the GIS. In the GIS, the exact routes, with the exact links in the road network are visualized. This is simply done by finding the cheapest paths between the nodes in the route, and easily done in the network analyst extension. The final results of what these routes look like in the GIS can be seen in Figure 26.

Figure 25: Visualization of the optimal solution to the test case

Figure 26: The final routes of the solution
6.4 Some Concluding Comments about IDARP

In this chapter an exact formulation of the integrated dial-a-ride problem (IDARP) has been introduced. The model has been tested on several cases and proven to work as intended. Small-scale cases can be solved to optimality. In order to solve medium to large sized problem instances, more efficient solution methods need to be developed. The next step in our work of the IDARP is to evaluate how different characteristics of input data, such as time window sizes and dispersion of requests in both time and space affect the complexity of the problem. We will also study how different branching techniques affect the computation time. We believe that with further research on suitable branching techniques for the (IDARP), medium sized problems can be solved.

We have also shown the benefits of using a geographic information system when creating the input files to the problem and when visualizing the results.
7 Conclusions and Future Research

We have described how a framework for planning of an integrated service should be designed. The importance of a connection between a GIS, simulation tools and optimization tools has been emphasized.

The evaluated modeling tool LITRES-2 has proven to be a useful tool for simulating multi modal public transport journeys. It offers a unique opportunity to study markets response to changes in the range of transport services. What is missing is the possibility to plan an integrated public transport system fully from the operator’s perspective. Performed simulations have given some guidelines of how to operate an integrated service. They have shown that small vehicles can efficiently be used in an integrated public transport service operated door-to-door.

The exact model of how to assign passengers to vehicles in an integrated public transport system has been tested on several cases and proven to work as intended. It has also shown that it is possible to solve small instances of the integrated dial-a-ride problem to optimality. The literature review shows that more work has to be done on effective planning of journeys in integrated services. Also research on systems for planning and design of such services are needed, as well as to study the effects of an integrated service, to passengers, operators and society.

One focus on future research within this area should be on how to develop more efficient methods to plan integrated journeys. The first we will do is to test our exact formulation on a small real-world case. The purpose of this is to see how effective our methods to strengthen the formulation are with real-world data. Further development of these strengthening methods can also be necessary. We will also study how different branching techniques can improve the efficiency of the solution method. Finally, a heuristic method will be developed to find good solutions to larger instances.

An other important issue to focus future research on is how to combine solution methods with a GIS. We believe that computational tools must be integrated with a GIS to simplify the use of such tools, and especially give simple ways to prepare input and visualize results. By this reason, we will combine our heuristic method with a GIS.

More research must also be made on how the cost for the operator is affected by the introduction of an integrated public transport service. We will evaluate what kind of economical savings that can be made regarding customers that today uses the Transportation of the Disabled Services. This can be an important factor to convince local authorities and operators of public transport of the benefits of an integrated service.
References


References


References


