Design of Integrated Flexible Transit Service with Given Fixed-Route Services

Khaled Saeed · Fumitaka Kurauchi

Abstract The increased use of private cars has caused many negative environmental, social and economic effects. Local authorities always do efforts to encourage the travellers to use the existing public transport services to reduce the negative effects of the increased personal mobility. Offering flexible transit services can compete with the flexibility of private cars and attract the travellers to use the public transport. However, these type of services has high operating costs which require subsidies to maintain running the service. In this paper, we present an integrated service between the flexible transit services and the existing fixed-route services with the aim of reducing the total costs and maintaining a suitable level of service for the travellers. Two exact formulations for the proposed integrated service are presented and analysed. Computational results on a set of instances provide a clear understanding of the capability of integrated service in reducing the total costs while maintaining a suitable level of service for the travellers.

Keywords: Dial-a-ride · Integrated services · Low Demand · Operational planning · Public transport · Scheduled lines
1 Introduction

The road-based transportation is a key issue for the social and economic development of any society. The increased use of private cars has caused many environmental, social and economic problems. Environmental side effects include the emissions of air pollution and greenhouse gases which contribute to the global warming. Social problems are related to the decreased accessibility and availability of public transport for some user groups, e.g. non-drivers, low income, elderly and people with disabilities. On the economic level, public transit services may suffer from low passenger loads which make it economically inefficient and may require government subsidies to maintain their services. Encouraging the travellers to use the public transport can decrease the use of the private cars, and thereby the problems associated with the increased personal mobility.

Local authorities and transit operators always try to offer more attractive public transit services and serve as many persons as possible. However, several reasons make people give up using the public transportation and prefer using their private cars. Examples for these reasons are: the last mile problem, long waiting times at the public transport stops especially in rural areas, the fixed-route services are not frequent or sometimes not exist, and sometimes a private car is needed to give other people lifts (e.g. children to school, hospital, and shopping). To attract people to use the public transportation and overcome the previous issues, public transportation systems are evolving towards more flexible solutions.

Flexible transit services can attract travellers to use the public transportation due to its flexibility compared with the traditional fixed-route services. Dial-a-Ride (DAR) service is one of those services which can provide shared-ride service with flexible routes and schedules. The problem of designing vehicles’ itineraries and schedules is known as the Dial-a-Ride Problem (DARP). The DARP consists of designing vehicle routes and schedules for a number of users who specify pickup and delivery requests between origins and destinations (Cordeau et al., 2007). However, the operating costs for those services are higher compared with the fixed-route services. From this point of view, in this paper we seek to combine the advantages of both services, i.e. DAR and fixed-route, to offer more competitive public transport services.

The integrated service between DAR and fixed-route services can serve travellers with suitable service quality at reasonable costs. This type of services can use the existing fixed-route service for a major part of the trip and the DAR services for the shorter distances. It can reduce the total costs of the public transportation by using the cheaper fixed-route service for major part of the journey. In addition, extending the public transportation to the low-density areas where it is not advisable to use the fixed-route services. The design of this integrated service is known as the Integrated Dial-a-Ride Problem (IDARP). The goal is to design vehicles’ itineraries and schedules for the DAR service where some part of the passengers’ trips can be performed by the fixed-route service.
The contributions of this paper are the following: (1) we introduce a formulation for an integrated service includes both operators’ and passengers’ perspectives, (2) we model and control the passengers’ transfers between different integrated services, (3) considering the passenger inconvenience caused by the congestion in peak-hours for the fixed-route service by including the available on-board capacity in the model, (4) give the possibility for any request to start/end the trip using the DAR service or the fixed-route service depending on the total costs minimization, (5) track accurate passengers’ movements and transfers through the entire network using the different passenger time, passenger flow and passenger transfer variables, and (6) the capability of solving complex networks contain more than one line for the fixed-route service.

The remainder of the paper is organized as follows. Section 2 provides a brief review of the related works. In section 3 the description and mechanism of the integrated service are discussed. Section 4 introduces the model formulation, defines the key variables, provides mathematical formulations for two models and shows a comparison between them. In Section 5, we present our numerical results and describe the advantages conferred by the integrated service. Finally, we conclude our work in Section 6 and describe our plans for future work.

2 Literature Review

This section reviews the literature addressing the related problems of the integration between flexible and fixed-route services. One of the earliest attempts to integrate the flexible service (DAR) with a fixed-route service was done by Liaw et al. (1996). The authors formulated a model to maximize the number of served requests while satisfying the service time constraints for the travellers. Their results showed that improvements of up to 9% in terms of the number of served requests can be obtained using static planning. Hickman and Blume (2001) used a two-stage heuristic to solve the integrated service scheduling problem to minimize the operating costs while satisfying a set of constraints on the quality of the service. The authors concluded that the potential cost savings and passenger level of service are sensitive to the parameters of the minimum and maximum passenger trip lengths for paratransit trip "legs" and the assumed penalty for passenger transfers. Aldaihani and Dessouky (2003) used a two-stage heuristic to get an approximate solution for the integrated problem. The results showed that shifting some of the demand for a hybrid service route (18.6% of the requests) reduces the on-demand vehicle distance by 16.6%. Hall et al. (2009) introduced a mixed integer formulation to solve an Integrated Dial-a-Ride Problem (IDARP). In their model, the minimization of the operating costs was only considered. Also, the authors did not include the timetables for the fixed-route service or the passengers’ transfers in their model. In addition, they assume an infinite capacity for the fixed-route service and high frequency so the waiting time at transfer location was not considered.
Posada et al. (2017) updated the work of Hall et al. (2009) to include time synchronization between the two integrated services. The authors did not include the passengers’ transfers or waiting times in their model. Due to the complexity of the model, the authors presented a second model in an attempt to reduce the number of the variables. However, the second model didn’t make significant improvements as only 5 requests were solved optimally within 7 hrs. Lee and Savelsbergh (2017) presented a model to minimize the running costs for the demand responsive connector vehicles, which transports passengers to transit hubs and after they continue their journey via a fixed-route service. The authors assumed that a sufficient number of vehicles are available to serve all passengers which in a sense might not be a practical assumption.

One of the similar systems to the integrated service is the Pickup and Delivery Problem with Transfers (PDPT) and the Dial-a-Ride Problem with Transfers (DARPT). In these systems, the transfers may happen between different vehicles. Also, it is not required that the requests should be picked up and delivered using the same vehicle. Cortes et al. (2010) proposed a solution method based on Benders decomposition to solve the PDPT. The size of the solved problem was small (up to 6 requests, 2 vehicles, and 1 transhipment node). They reported up to 90% saving in the CPU time when compared with the traditional Branch and Bound algorithm. The PDP with transshipment was addressed by Rais et al. (2014). The results show that cost savings up to 7% can be achieved due to transfers. Masson et al. (2014) propose an Adaptive Large Neighborhood Search (ALNS) to solve the DARPT. The authors show that using the transfer points can lead to operating costs savings up to 8.25%.

Different example for the integration concerned with the passenger and freight transportation have been introduced. Trentini et al. (2012) introduced a shared passengers and goods city logistics system where the spare capacity of a bus line is used to transport goods. Ghilas et al. (2013) proposed a formulation for integrated service where pickup and delivery (PD) vehicles can carry both passengers and freights. The PD vehicles transport the requests to a station-hub and from there the requests continue the journey using scheduled lines. Their results show that using the integrated service cause a reduction up to 27% in the operation costs and a decrease up to 70% for the CO\textsubscript{2} emissions.

The contributions in solving the DARP can be useful in solving the IDARP since the DARP is considered as the basis for the IDARP. Cordeau (2006) proposed a three-indexed formulation for the DARP. Using their numerical approach, they were able to find optimal solutions for up to 36 requests. Saeed and Kurauchi (2015) used a Branch and Cut algorithm to solve the DARP for up to 65 requests within 2 hours. Muelas et al. (2013) proposed a new variable neighborhood search and test it on a set of 24 different scenarios in a large-scale DARP. Recent reviews for the problem variants and the solution methodologies can be found in Molenbruch et al. (2017) and Ho et al. (2018).

In summary, previous researchers assume that any feasible solution is acceptable to the travellers as long as it satisfies the travel time constraints. However, it is not
always true, since they did not model the transfers, hence passengers might experience a high number of transfers during their trips. Also, the on-board capacity for the fixed-route service was not considered which is not accurate especially during peak-hours as it will increase the user inconvenience and passengers might give up using the service. From the modelling point of view, the previous models cannot solve networks include more than one line for the fixed-route service since there are no variables (e.g. time and transfers) to track the passengers’ movements at each node. In this research, we present a mixed integer formulation to design an integrated service between flexible services and the existing fixed-route services. Our aim is to minimize the total costs including the operating, users’ travel, and transfer penalty costs. In addition, we seek to fulfill the drawbacks of the previous models.

3 Description of the Integrated Service

This type of services can use the existing fixed-route service for a major part of the journey and the flexible services for the shorter distances. The total costs can be reduced by using the cheaper fixed-route service for the longer part of the trips. In addition, extending the public transport coverage area to serve the low-density areas where it is not advisable to use the fixed-route service. Figure 1 shows a scheme for an integrated service between Dial-a-Ride (DAR) service and the existing bus service to serve four requests (travellers). Nodes \((P, D)\) represents the pick-up and drop-off points for the requests, respectively. For example, node \(P1\) represents the pick-up node for request 1 and node \(D3\) represents the drop-off node for request 3. While nodes \(BS\) represents the various location of the bus stops (transfer nodes).

![Fig. 1 Scheme for Satisfying a Demand of Four Requests](image-url)
Some travellers can be offered a direct trip using the DAR service between their origins and destinations (e.g. traveller 3 using path P3-D3). While other travellers might experience a combination between the DAR service and bus service (e.g. traveller 1 using path P1-P2-BS1-BS2-D1 or traveller 2 using path P2-BS1-BS3-BS4-D2). The traveller’s trip does not have to start/end with the DAR service in case that the origin/destination is near to the bus stop (e.g. node D1 destination of traveller 1). Also, it is not necessary that the same DAR vehicle which used to pick up the traveller in the beginning of his trip should drop him at the end of his trip (e.g. traveller 2 which uses DAR1 in the beginning of the trip and DAR2 in the end of the trip).

4 Model Formulation for the Integrated Service

The objective of the proposed formulation is to find the optimal itineraries and schedules for DAR vehicles and passengers, where part of the trip can be performed by the existing fixed-route service, such that the total costs are minimized. The minimized total costs include operating costs for DAR vehicles, passengers' in-vehicle travel costs, and passengers' transfer penalties costs. The passengers' in-vehicle travel costs include the costs for traveling using both the DAR service and fixed-route service. A transfer penalty cost is calculated if a passenger transfer between the DAR and fixed-route services or between different lines on the fixed-route service. A transfer will occur after satisfying a set of constraints related to the available on-board capacity and the time synchronization between the DAR and fixed-route services or between vehicles on different lines for the fixed-route service.

After studying the structure of the problem explained earlier and to solve it to optimality it is required to determine a huge number of variables. The main sources of problem complexity are the number of requests and transfer locations. Previous researchers have attempted to solve these issues to reduce the size of the problem and increase the number of solved requests. However, they only solved small instances and the computational time is still large. Also, in the models presented in this paper, we attempt to enhance the quality of the service and consider the passengers’ interests. These modifications add new variables and increase the size of the problem hence the required computational time. Two models are presented in this paper, the first one follows the same network structure used by previous researchers. In addition to the modifications and new variables we have added to the problem. In the second model, we attempt to change the problem modelling and network structure to reduce the number of variables.

In practical situations, the pick-up or drop-off points of some requests might be close enough to the transfer locations. Which allows the travellers to start/end their trips using the fixed-route service and there is no need to use the DAR service as they can walk to/from the fixed-route stop directly. For these situations, in both models, we have created a special subset $N_j$ of the pick-up and drop-off nodes.
which is close enough to a transfer location. These requests have the ability to start/end the trips using the DAR service or walk to the transfer locations to use the fixed-route service. The allowable walking distance can be set as an input in the both models.

In the next subsections, we will present the variable definitions used in both models, providing the details of the two models and the network’s descriptions as well as a comparison between the two models. Also, numerical experiments will be used to show the difference in computational times for both models.

4.1 Variable Definitions

In the mathematical formulations, we aim to find the optimal itineraries and schedules for DAR vehicles and passengers, where part of the trip can be performed by the existing fixed-route service, such that total costs are minimized. The minimized total costs include: operating costs for DAR vehicles, passengers' in-vehicle travel costs, and passengers' transfer penalties. To address this problem, we determine variables which are relevant to the design of DAR such as: (1) DAR arc flows ($X$) that specify the DAR vehicle itineraries, (2) DAR vehicle capacities ($Q$) which represent the number of on-board passengers for the DAR vehicle after it leaves any node, and (3) DAR service times ($B$) which correspond to the time at which the DAR vehicles begin their service at any node, whether loading at a pick-up nodes or unloading at a drop-off nodes. In addition to passengers’ (requests’) related variables such as: (4) passengers times at each node ($T$), (5) passenger arc flows ($Y$) that indicate if a request is traveling using the DAR service, (6) passenger arc flows ($Z$) that indicate if a request is traveling using the fixed-route service, (7) passenger transfer variable ($\delta$) which indicate whether a passenger has transferred from DAR vehicle to the fixed-route service or not, (8) passenger transfer variable ($\theta$) which indicate whether a passenger has transferred from the fixed-route service to DAR vehicle or not, and (9) passenger transfer variable ($\omega$) which indicate the passenger transfer between different lines for the fixed-route service in case there are more than one line.

The proposed formulations can track accurate passengers’ movements and transfers through the entire network using the different passenger time, passenger flow and passenger transfer variables. Also, passengers can transfer between the different fixed-route service lines, but they cannot transfer between different DAR vehicles.

The following notations are used in the development of the two proposed formulations:

**Sets**

- $P$: Set of all pick-up nodes
- $D$: Set of all drop-off nodes
- $D_o$: Origin DAR vehicle depot
- $D_d$: Destination DAR vehicle depot
- $R$: Set of all requests
\( \mathcal{g} \) Set of all transfer locations
\( \mathcal{g}^{-} \) Set of transfer locations shared by more than one fixed-route line
\( \lambda \) Set of all replicated transfer nodes, \( \lambda = g_R \)
\( \lambda^{-} \) Set of replicated transfer nodes shared by more than one fixed-route line
\( \lambda_r \) Set of replicated transfer nodes related to request \( r \)
\( N \) Set of all nodes, \( N = P \cup D \cup D_o \cup D_d \cup \lambda \)
\( N_1 \) \( \quad P \cup D \)
\( N_2 \) \( \quad P \cup D \cup \lambda \)
\( N_3 \) Subset of pick-up/drop-off nodes which are associated directly with a transfer node
\( N_4 \) \( \quad P \cup D \cup g \)
\( A \) Set of all arcs connecting all nodes
\( A_1 \) Set of arcs \((i, j), \forall i \in N_1, j \in N_2 \)
\( A_2 \) Set of arcs \((i, j), \forall i \in N_2, j \in N_1 \)
\( A_3 \) \( A_1 \cup A_2 \)
\( A_4 \) Set of arcs \((i, j), \forall i \in \mathcal{g}, j \in \lambda \)
\( A_5 \) Set of arcs \((i, j), \forall i \in N_1, j \in N_4 \)
\( A_6 \) Set of arcs \((i, j), \forall i \in N_2, j \in N_1 \)
\( A_7 \) \( A_3 \cup A_6 \)
\( A_8 \) Set of arcs \((i, j), \forall i \in \mathcal{g}, j \in \lambda \)
\( K \) Set of all DAR vehicles
\( W_j \) Set of all fixed-route vehicles passing through transfer node \( j \)
\( W_{ij} \) Set of all fixed-route vehicles traveling through arc \((i, j)\)

Parameters

\( d_i \) Service duration for passengers at node \( i \) (boarding/alighting)
\( e_i \) Earliest time at which service can begin at node \( i \)
\( l_i \) Latest time at which service can begin at node \( i \)
\( h_r \) Parameter used to control the maximum ride time of request \( r \)
\( q_r \) Load of request \( r \)
\( T_k \) Maximum route duration time of DAR vehicle \( k \)
\( MT_r \) Maximum number of transfers for request \( r \)
\( Q_k \) Maximum capacity of DAR vehicle \( k \)
\( Q^w_{ij} \) Available on-board capacity for fixed-route vehicle \( w \) for arc \((i, j)\)
\( c^1_{ij} \) Operating cost for DAR vehicle \( k \) for traveling through arc \((i, j)\)
\( c^2_{ij} \) Passenger in-vehicle travel cost using DAR service for arc \((i, j)\)
\( c^3_{ij} \) Passenger in-vehicle travel cost using fixed-route service for arc \((i, j)\)
\( \tau \) Transfer penalty cost
\( s_j \) Stopping time at node \( i \) for DAR and fixed-route vehicles
\( s^1_i \) Transferring time between DAR and fixed-route vehicles at node \( i \)
\( s^2_i \) Transferring time between different fixed-route vehicles at node \( i \)
\( d^w_{ij} \) Departure time of fixed-route vehicle \( w \) from node \( i \) to node \( j \)
Arrival time of fixed-route vehicle $w$ at node $i$ from node $j$ \\

Travel time for arc $(i, j)$ \\

\[
\alpha_i^r = \begin{cases} 
1 & \text{if node } i \text{ is the origin of request } r \\
-1 & \text{if node } i \text{ is the destination of request } r \\
0 & \text{otherwise}
\end{cases}
\]

Large positive number \\

**Variables**

\( T_i^r \) Continuous variable indicates the time of request $r$ at node $i$ \\

\( s_i \) Continuous variable indicates the time at which DAR vehicle starts service at node $i$ \\

\( B_k^i \) Continuous variable indicates the time of DAR vehicle $k$ at depots \\

\( x_{ij}^k \) Binary variable denotes whether DAR vehicle $k$ travels from node $i$ to node $j$ \\

\( y_{ij}^r \) Binary variable denotes whether arc $(i, j)$ is traversed by request $r$ using DAR service \\

\( z_{ij}^w \) Binary variable denotes whether arc $(i, j)$ is traversed by request $r$ using fixed-route service \\

\( \delta_j^r \) Binary variable indicates if request $r$ transfers from DAR service to fixed-route service at transfer node $j$ \\

\( \theta_j^r \) Binary variable indicates if request $r$ transfers from fixed-route service to DAR service at transfer node $j$ \\

\( \omega_j^r \) Binary variable indicates if request $r$ transfers between different fixed-route lines at transfer node $j$ \\

\( \Omega_i^k \) Auxiliary variable indicates the sequence of node $i$ on DAR vehicle $k$ path

### 4.2 Mathematical Formulation of Model 1

In the first mathematical formulation the same network structure used by Hall et al. (2009) was used. In addition, the modifications and new variables we have added to the problem. The problem is formulated over a directed graph $G(N, A)$, in which $N$ is the set of nodes and $A$ specifies the set of arcs. Each arc $(i, j) \in A$ connects nodes $i$ and $j$ for $i, j \in N$, $i \neq j$ with a link travel time $t_{ij}$. Figure 2 shows a network representation for two requests and two transfer locations. Each physical transfer node $i \in g$ contains one replicated node for each request, so the total number of replicated transfer nodes is equal to $(\lambda=gR)$. For example, a replicated node $a_2$ represents the original transfer node $a$ and request 2. The arcs shown in this figure represents the binary variable $(X)$ for one DAR vehicle. Binary variable $(Y)$ for one request has the same arcs as $(X)$ for one DAR vehicle except any arcs connected to the DAR vehicle depots (i.e. sets 1, 6, 7, 8, and 9).

The set of nodes, $N$, is divided into five subsets: pick-up nodes ($P = \{1, \ldots, R\}$), drop-off nodes ($D = \{R+i, \ldots, 2R\}$), the origin depot node ($D_{do} = \{0\}$), the destination depot node ($D_{d} = \{2R+1\}$), and the set of all replicated transfer nodes ($\lambda=gR$). We
define \( R \) as the set of all requests and \( K \) as the set of all DAR vehicles. Every DAR vehicle has maximum capacity \( Q_k \) and maximum route duration time \( T_k \). Every request is associated with two nodes \((i, R^+i)\), where \( i \in P \) and \( R^+i \in D \) and has load \( q_i \). Each node \( i \in N \) is associated with a time window \([e_i, l_i]\), where \( e_i \) and \( l_i \) represent the earliest and latest time at which the service may begin at node \( i \), respectively. Also, a service duration \( d_i \) for loading/unloading passengers is associated with each node \( i \in N \). The passenger load \( q_i \), is positive at pick-up nodes and negative at drop-off nodes.

\[
\begin{align*}
\text{Min} & \sum_{k \in K} \sum_{j \in A_3} c_{ij}^k x_{ij}^k + \tau \sum_{r \in R} \sum_{j \in A} (\delta_{ij}^r + \theta_{ij}^r) + \tau \sum_{r \in R} \sum_{j \in A} \omega_{ij}^r \\
& + c_{ij}^r \sum_{r \in R} \sum_{j \in A_3} q_i y_{ij}^r t_{ij} + c_{ij}^n \sum_{r \in R} \sum_{j \in A_4} \sum_{w \in W_j} q_i z_{ij}^w t_{ij}
\end{align*}
\]  

(1)
\[
\sum_{k \in K} \sum_{j \in N} x_{ij}^k = 1 \quad \forall i \in P \cup D \setminus N_3
\] (2)

\[
\sum_{k \in K} \sum_{j \in N} x_{ij}^k \leq 1 \quad \forall i \in N_3
\] (3)

\[
x_{0j}^k = 1 \quad \forall k \in K
\] (4)

\[
x_{i,2R+1}^k = 1 \quad \forall k \in K
\] (5)

\[
\sum_{j \in N} x_{ij}^k - \sum_{j \in N} x_{ij}^k = 0 \quad \forall i \in N_2, k \in K
\] (6)

\[
\sum_{j \in N_2} y_{ij}^r - \sum_{j \in N_2} y_{ij}^r = \alpha_i^r \quad \forall r \in R, i \in N_1 \setminus N_3
\] (7)

\[
\sum_{j \in N_2} y_{ij}^r + \sum_{j \in \lambda \cup W_{ij}} z_{ij}^{rw} - \sum_{j \in N_2} y_{ij}^r - \sum_{j \in \lambda \cup W_{ij}} z_{ij}^{rw} = \alpha_i^r \quad \forall r \in R, i \in \lambda_r
\] (8)

\[
\sum_{r \in R} q_r y_{ij}^r \leq \sum_{k \in K} Q_k x_{ij}^k \quad \forall ij \in A_3
\] (9)

\[
\sum_{r \in R} q_r z_{ij}^{ww} \leq Q_{ij}^w \quad \forall ij \in A_4, w \in W_{ij}
\] (10)

\[
B_j \geq B_i + s_i + t_{ij} - M(1 - \sum_{k \in K} x_{ij}^k) \quad \forall ij \in A_3
\] (11)

\[
T_{ij}^r \geq T_{ij}^r + s_i + t_{ij} - M(1 - y_{ij}^r) \quad \forall ij \in A_3, r \in R
\] (12)

\[
\sum_{j \in \lambda \cup W_{ij}} z_{ij}^{rw} d_{ij}^w \geq T_{ij}^r + s_i^j - M(1 - \sum_{j \in \lambda \cup W_{ij}} z_{ij}^{rw}) \quad \forall j \in (\lambda \cup N_3), r \in R
\] (13)

\[
T_{ij}^r \geq \sum_{j \in \lambda \cup W_{ij}} z_{ij}^{rw} p_{ij}^{w} + s_i^j - M(1 - \sum_{j \in \lambda \cup W_{ij}} z_{ij}^{rw}) \quad \forall j \in (\lambda \cup N_3), r \in R
\] (14)

\[
\tau_{i,R+i} \leq T_{R+i}^r - (T_{ij}^r + s_i) \leq h_r(t_{i,R+i}) \quad \forall i \in P
\] (15)

\[
e_i \leq T_{i}^r \leq l_i \quad \forall i \in P \cup D
\] (16)
\begin{align}
B_j & \geq \beta_k^j + s_j + t_{oj} - M(1 - x_{oj}^k) \quad \forall j \in N_2, k \in K \\
\beta_{2R+1}^k & \geq B_i + t_{i(2R+1)} - M(1 - x_{i2R+1}^k) \quad \forall j \in N_2, k \in K \\
\beta_{2R+1}^k & - \beta_0^k \leq T_k \quad \forall k \in K \\
T_r^r - B_i & \leq M(1 - \sum_{j \in N_2} y_{ij}^r) \quad \forall r \in R, i \in N_2 \\
T_r^r - B_i & \geq -M(1 - \sum_{j \in N_2} y_{ij}^r) \quad \forall r \in R, i \in N_2 \\
\sum_{i \in N_1} y_{ij}^r + \sum_{e \in \lambda} \sum_{w \in W_j} z_{ew}^{rw} & \leq \delta_{ij}^r + 1 \quad \forall r \in R, j \in \lambda_r \\
\sum_{i \in N_1} y_{ji}^r + \sum_{e \in \lambda} \sum_{w \in W_j} z_{ew}^{rw} & \leq \theta_{ij}^r + 1 \quad \forall r \in R, j \in \lambda_r \\
\sum_{i \in \lambda} \sum_{j} z_{ij}^{rw} + \sum_{e \in \lambda} \sum_{v \in W_j} z_{jv}^{rv} & \leq \omega_{ij}^r + 1 \quad \forall r \in R, j \in \lambda \\
\Omega_k^i - \Omega_k^j + |N| \times x_{ij}^k & \leq |N| - 1 \quad \forall i, j \in N_2, k \in K
\end{align}

The objective function (1) is to minimize the sum of different cost terms. The first term is the operating costs for the DAR vehicles. The second term is the transfer penalty costs for passengers transferring between DAR and fixed-route services. The third term is the transfer penalty costs for passengers transferring between the different fixed-route lines. The fourth term is the passengers’ in-vehicle travel costs for using the DAR service. The last term is the passengers’ in-vehicle travel costs for using the fixed-route service. Constraint (2) ensure that DAR vehicles visit every request node (pick-up or drop-off) exactly once, with the exception of nodes associated directly with a transfer node as it might be served by the fixed-route service instead of the DAR service. Constraint (3) ensure that the pick-up and drop-off nodes associated directly with a transfer node might be served by the DAR service or not, in case the fixed-route service is used to serve that node. Constraints (4) and (5) guarantee that the route for each DAR vehicle \( k \) will start from the origin depot and end at the destination depot. Constraint (6) ensures that the DAR vehicle leaves each node it arrives. Constraint (7) represents the passenger flow balancing condition on pick-up and drop-off nodes, with the exception of nodes associated directly with a transfer node. While constraint (8) is the passenger flow balancing condition on transfer nodes and also the pick-up and drop-off nodes associated directly with a transfer node. On-board capacity limits are guaranteed by
constraints (9) and (10) to meet the passenger demand using the DAR and fixed-route services, respectively. Scheduling of the DAR vehicles is considered in constraint (11). Similarly, for each request, the timing is considered in constraint (12).

Constraints (13) and (14) ensure the timing of request \( r \) with the fixed-route service timetables. Constraint (15) limits the allowable ride time of each request. The parameter \( h_r \) allows us to control the upper limit of the ride time of a particular request. Constraint (16) ensures that all nodes are serviced within their time window. Constraints (17) and (18) ensure the time of service for the DAR vehicles at the origin and destination depots. Constraint (19) ensures that the duration of each vehicle itinerary is less than the maximum route duration time for the DAR vehicle. Constraints (20) and (21) assures that the service time for request \( r \) is equal to the service time of the DAR vehicle. Constraints (22 - 24) are related to the different types of transfers for each request. Constraint (22) used to know whether request \( r \) transfers from DAR vehicle to fixed-route lines at transfer nodes. The same for the transfers from fixed-route lines to DAR vehicles or the transfers between different fixed-route lines in constraints (23) and (24), respectively. Constraint (25) is used to prevent sub-tours for the DAR vehicles itineraries.

4.3 Mathematical Formulation of Model 2

The size of the problem presented in Model 1 increases quickly as a function of the number of solved requests and the number of transfer locations. The replication of every transfer location for every request is mainly done to capture the multiple visits of the DAR vehicles to every transfer location. The maximum expected number of visits for any transfer node is equal to the number of requests in the network. This is true and necessary to be done for the DAR arc flows variable (\( X \)) as we may expect multiple visits for the DAR vehicle to the transfer node. However, this is not necessary to be done for the passenger arc flows variable (\( Y \)) as the request is expected to visit any transfer node just once. Also, from a practical view, there is no need for any request to make multiple visits at any transfer location. This modification can reduce the number of variables in the model, hence improve the computational time. The previous modification results in changing the network representation and the number of variables for variables (\( Y, T, \delta, \theta, \) and \( \omega \)).

\[
\begin{align*}
\text{Min} & \sum_{k \in K} \sum_{i \in A_3} c_{ij} x_{ij}^k + \tau \sum_{r \in R} \sum_{j \in g} \left( \delta_j^r + \theta_j^r \right) + \tau \sum_{r \in R} \sum_{j \in g} \omega_j^r \\
& + c_{ij}^3 \sum_{r \in R} \sum_{j \in g} q_r y_{ij}^r t_{ij} + c_{ij}^3 \sum_{r \in R} \sum_{j \in g} \sum_{w \in W} q_r z_{ij}^w t_{ij} \\
\sum_{k \in K} \sum_{j \in N} x_{ij}^k &= 1 \quad \forall i \in P \cup D \setminus N_3
\end{align*}
\]
\[
\sum_{k \in K} \sum_{j \in N} x_{ij}^k \leq 1 \quad \forall i \in N_3 \quad (28)
\]

\[
\sum_{j \in N} x_{oij}^k = 1 \quad \forall k \in K \quad (29)
\]

\[
\sum_{j \in N} x_{i2R+1}^k = 1 \quad \forall k \in K \quad (30)
\]

\[
\sum_{j \in N} x_{ij}^k - \sum_{j \in N} x_{ij}^k = 0 \quad \forall i \in N_2, k \in K \quad (31)
\]

\[
\sum_{j \in N_2} y_{ij}^r - \sum_{j \in N_2} y_{ij}^r = \alpha_r^f \quad \forall r \in R, i \in N_1 \setminus N_3 \quad (32)
\]

\[
\sum_{j \in N_4} y_{ij}^r + \sum_{j \in g \in W_{ij}} z_{ij}^{rw} + \sum_{j \in N_4} y_{ij}^r - \sum_{j \in g \in W_{ij}} z_{ij}^{rw} = \alpha_r^f \quad \forall r \in R, i \in g \quad (33)
\]

\[
\sum_{r \in R} q_r y_{ij}^r \leq \sum_{k \in K} Q_k x_{ij}^k \quad \forall ij \in A_3 \quad (34)
\]

\[
\sum_{r \in R} q_r z_{ij}^{rw} \leq Q_{ij}^w \quad \forall ij \in A_3, w \in W_{ij} \quad (35)
\]

\[
B_j \geq B_i + s_1 + t_{ij} - M(1 - \sum_{k \in K} x_{ij}^k) \quad \forall ij \in A_3 \quad (36)
\]

\[
T_{ij}^r \geq T_{ij}^r + s_1 + t_{ij} - M(1 - y_{ij}^r) \quad \forall ij \in A_3, r \in R \quad (37)
\]

\[
\sum_{j \in g \in W_{ij}} z_{ij}^{rw} d_{ij}^w \geq T_{ij}^r + s_i^1 - M(1 - \sum_{j \in g \in W_{ij}} z_{ij}^{rw}) \quad \forall j \in (g \cup N_3), r \in R \quad (38)
\]

\[
T_{ij}^r \geq T_{ij}^r + s_j^1 - M(1 - \sum_{j \in g \in W_{ij}} z_{ij}^{rw}) \quad \forall j \in (g \cup N_3), r \in R \quad (39)
\]

\[
t_{i,R+1} \leq T_{i,R+1}^r - (T_{i,R+1}^r + s_i) \leq h_r(t_{i,R+1}) \quad \forall i \in P \quad (40)
\]

\[
e_i \leq T_i^r \leq l_i \quad \forall i \in P \cup D \quad (41)
\]

\[
B_j \geq \beta_{o}^k + s_j + t_{oj} - M(1 - x_{oij}^k) \quad \forall j \in N_2, k \in K \quad (42)
\]

\[
\beta_{2R+1}^k \geq B_i + t_{i2R+1} - M(1 - x_{i2R+1}) \quad \forall j \in N_2, k \in K \quad (43)
\]
\[
\beta_{2R+1}^k - \beta_0^k \leq T_k \quad \forall \ k \in K
\]

\[
T_i^r - B_i \leq M(1 - \sum_{j \in N_4} y_{ij}^r) \quad \forall \ r \in R, i \in N_2
\]

\[
T_i^r - B_i \geq -M(1 - \sum_{j \in N_4} y_{ij}^r) \quad \forall \ r \in R, i \in N_2
\]

\[
\sum_{i \in N_1} y_{ij}^r + \sum_{e \in g} \sum_{w \in W} z_{je}^{rw} \leq \delta_{ij}^r + 1 \quad \forall \ r \in R, j \in g
\]

\[
\sum_{i \in N_2} y_{ji}^r + \sum_{e \in g} \sum_{w \in W} z_{ej}^{rw} \leq \theta_{ij}^r + 1 \quad \forall \ r \in R, j \in g
\]

\[
\sum_{i \in g} \sum_{w \in W} z_{ij}^{rw} + \sum_{e \in g} \sum_{v \in W} z_{je}^{rw} \leq \omega_{ij}^r + 1 \quad \forall \ r \in R, j \in \bar{g}
\]

\[
\Omega_i^k - \Omega_j^k + |N| \times x_{ij}^k \leq |N| - 1 \quad \forall \ i, j \in N_2, k \in K
\]

\[
\sum_{j \in \lambda} (\delta_{ij}^r + \theta_{ij}^r) + \sum_{j \in \lambda} \omega_{ij}^r \leq MT_r \quad \forall \ r \in R
\]

\[
\sum_{i \in N_1} y_{ij}^r \leq 1 \quad \forall \ r \in R, j \in g
\]

Most of the constraints in Model 2 have the same purpose and formulation as in Model 1. However, due to the modifications in the network structure for variables \((Y, T, \delta, \theta, \omega)\) some changes are necessary to consider one node only for each transfer location for these variables. These modifications can be noticed in constraints (32)-(35), (37)-(39), and (45)-(49). Constraint (51) controls the maximum number of transfers for each request. Constraint (52) is added to ensure that each request will visit any transfer location only once.

### 4.4 Comparison Between the Two Models

The Modifications made in Model 2 were to reduce the number of binary and continuous variables \((Y, T, \delta, \theta, \omega)\), hence the required computational time. For a better explanation about the reduction in the number of variables, imagine we have a network of one bus line \((l=1)\) which consist of six bus stops \((g=6)\) and three bus vehicles \((w=3)\) passing every one hour on each bus stop. Also, we have one DAR vehicle \((k=1)\), one origin depot, and one destination depot. Figure 3 gives a
comparison of the total number of variables for the previously explained network for both models as a function of the number of requests. It is obvious that the number of variables in Model 1 increases more rapidly as the number of the requests increases compared with Model 2.

![Graph](image.png)

**Figure 3.** The Number of Total Variables for the Two Models as a Function of the Number of Requests

### 5 Results

The solution algorithm described in this study was developed in Matlab (Matlab Inc., Natick, MA, USA), and the optimization solver was the academic version of IBM CPLEX 12.6.1. The experiments were run on a 3.6-GHz i7-4960x computer. The formulation was tested using randomly generated instances over a square \([-20, 20] \times [-20, 20]\). In all instances, the coordinates of the pick-up and drop-off points were randomly chosen and one distance unit corresponds to one driving time unit. One bus line is located in this area with four bus stops and the average length between two consecutive bus stops was 14. We assume a planning horizon of 6 hours and the width of the time window was taken as 15 min.

Up to four DAR vehicles were used with a maximum capacity $Q_k = 8$ passengers. The unit DAR operating cost per unit length is $c^l = 1$ and the maximum DAR vehicle route duration is 680 min. We set $c^d = 0.035$, $c^l = 0.03$, and $\tau = 0.02$. The maximum computational time was set to 200 min (≈ 3.33 hours) to check the computational ability and the calculations were interrupted if the computational time exceeded the time limits. Each instance was identified with a letter and a number. The letter referred to the instance set and the number referred to the number of requests. The number of passengers in each instance was selected randomly from 1-3 passengers.

Table 1 shows the computational time required to solve the randomly generated instances. Tests have been performed on both models and the final computational
time was reported. For both models, the solving times increase rapidly and the results show that the increase is higher for Model 1 as the complexity increases faster compared with Model 2 as shown previously in Figure 3. We were able to solve only 4 requests using Model 1 within the predefined computational time limits. Due to the lower number of variables in Model 2, up to 7 requests were solved optimally.

Table 1 Computational Results for Instances Using Both Models

<table>
<thead>
<tr>
<th>Instance</th>
<th>No. of Requests</th>
<th>No. of DAR Vehicles</th>
<th>Computational Time</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Model 1</td>
</tr>
<tr>
<td>C2</td>
<td>2</td>
<td>2</td>
<td>1.7 sec</td>
</tr>
<tr>
<td>C3</td>
<td>3</td>
<td>3</td>
<td>15 sec</td>
</tr>
<tr>
<td>C4</td>
<td>4</td>
<td>4</td>
<td>16 min</td>
</tr>
<tr>
<td>C5</td>
<td>5</td>
<td>4</td>
<td>--</td>
</tr>
<tr>
<td>C6</td>
<td>6</td>
<td>4</td>
<td>--</td>
</tr>
<tr>
<td>C7</td>
<td>7</td>
<td>4</td>
<td>--</td>
</tr>
</tbody>
</table>

Table 2 shows the total costs for both the operator and the users for all the generated instances for both models. The operator’s costs (Oper.) include the DAR vehicles operating costs. The users’ costs consist of the in-vehicle travel time costs for using DAR and bus services in addition to the transfer penalty costs. To demonstrate the advantages of the integration between the DAR and the existing fixed-route services, we calculate the optimal solution when only a DAR service is provided using the model for the DAR service in our previous study (Saeed and Kurauchi, 2015). In the case of DAR only, the operator’s costs are the DAR vehicles operating costs and the users’ costs is the in-vehicle travel time costs for using the DAR service.

Table 2 Components of the Total Costs for Different Optimal Solutions

<table>
<thead>
<tr>
<th>Instance</th>
<th>Model 1</th>
<th></th>
<th>Model 2</th>
<th></th>
<th>DAR Only</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total Costs</td>
<td>Oper. Costs</td>
<td>User Costs</td>
<td>Total Costs</td>
<td>Oper. Costs</td>
<td>User Costs</td>
</tr>
<tr>
<td>C2</td>
<td>16.28</td>
<td>10.42</td>
<td>5.86</td>
<td>16.28</td>
<td>10.42</td>
<td>5.86</td>
</tr>
<tr>
<td>C3</td>
<td>27.11</td>
<td>15.19</td>
<td>11.92</td>
<td>27.11</td>
<td>15.19</td>
<td>11.92</td>
</tr>
<tr>
<td>C4</td>
<td>21.30</td>
<td>7.45</td>
<td>13.85</td>
<td>21.30</td>
<td>7.45</td>
<td>13.85</td>
</tr>
<tr>
<td>C5</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>32.52</td>
<td>18.57</td>
</tr>
<tr>
<td>C6</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>40.55</td>
<td>22.42</td>
</tr>
<tr>
<td>C7</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>53.96</td>
<td>33.47</td>
</tr>
</tbody>
</table>

As shown in Table 2, for Model 1 as the number of requests increased for instances (C5, C6, and C7), we couldn’t get an optimal solution within the predefined computational time limits due to the complexity of Model 1. Also, the total costs for Models 1 and 2 for all the generated instances are lower than the total costs for the DAR service only. This means that integrating the DAR with the existing bus service will reduce the total welfare costs (i.e. summation operator’s and users’ costs). By comparing the results of the two models, instances (C2, C3, and C4) have the same values for all the cost components. However, as shown in Table 1, the
required computational time for Model 1 is much higher compared with Model 2 due to the higher number of variables in Model 1.

Comparing the values of the users’ costs in the three optimal solutions, it’s clear that the difference in the users’ costs between the DAR only and Models 1 and 2 are small. This indicates the capability of the integrated service in reducing the total costs while maintaining a suitable level of service for the travellers.

6 Conclusions

In this paper, we have presented two exact formulations for a service integrate the flexible services (e.g. DAR) with the existing fixed-route services (e.g. bus). The aim was to minimize the total costs of the operator and users. We attempt to decrease the number of variables by making some modifications to the network structure in Model 2. The results show that Model 2 was able to solve a higher number of requests compared to Model 1. In addition, the results show the capability of the integrated service in reducing the total costs while maintaining a suitable level of service for the travellers. We further aim to test our model on real-life data. In addition, including the waiting time costs for the travellers before picked up by DAR vehicles and during waiting for the fixed-route vehicles. Also, extend the current model to include the transfers between different DAR vehicles.

References