Transportation Research Part B 77 (2015) 38–75

Contents lists available at ScienceDirect

Transportation Research Part B

journal homepage: www.elsevier.com/locate/trb

Review

Planning, operation, and control of bus transport systems: A literature review

O.J. Ibarra-Rojas, F. Delgado, R. Giesen, J.C. Muñoz *

BRT Centre of Excellence, Department of Transport Engineering and Logistics, Pontificia Universidad Católica de Chile, Vicuña Mackenna 4860, Macul, Casilla 306, Código 105, Santiago, Chile

A R T I C L E   I N F O

Article history:
Received 24 June 2014
Received in revised form 11 February 2015
Accepted 7 March 2015
Available online 2 April 2015

Keywords:
Literature review
Bus transport systems
Planning
Operation
Real-time control

A B S T R A C T

The efficiency of a transport system depends on several elements, such as available technology, governmental policies, the planning process, and control strategies. Indeed, the interaction between these elements is quite complex, leading to intractable decision making problems. The planning process and real-time control strategies have been widely studied in recent years, and there are several practical implementations with promising results. In this paper, we review the literature on Transit Network Planning problems and real-time control strategies suitable to bus transport systems. Our goal is to present a comprehensive review, emphasizing recent studies as well as works not addressed in previous reviews.

© 2015 Elsevier Ltd. All rights reserved.

Contents

* Corresponding author.
E-mail addresses: oibarrar@uc.cl (O.J. Ibarra-Rojas), fdb@ing.puc.cl (F. Delgado), giesen@ing.puc.cl (R. Giesen), jcm@ing.puc.cl (J.C. Muñoz).

http://dx.doi.org/10.1016/j.trb.2015.03.002
0191-2615/© 2015 Elsevier Ltd. All rights reserved.

1. Introduction

City growth in terms of surface and population is one of the most important global trends of the last century. Too often the speed of this phenomenon has prevented an organic or well-planned urban development. Instead, cities worldwide are suffering from long travel times, severe traffic congestion, pollution, road accidents, etc. Public transport is considered an important backbone of sustainable urban development, since it should allow more efficient movements across a city. However, public transport systems often struggle to provide a good level of service at an affordable cost for the public administration and the user: these systems need to provide a high level of service in order to be attractive to the non-captive user, while at the same time affordable for the low income segments of the population. To achieve this, public systems inevitably face the need of subsidies that are increasingly contested by society.

These systems vary quite a bit around the globe. Very often multi-modal, they provide different integration opportunities across services: while some cities provide fare integration citywide, in others, all services compete as independent alternatives. While some cities have relied on the operations of private companies, others maintain this responsibility within a public agency. Planning, operating and controlling a public transport system, which is the focus of this document, is very challenging. Several actors with different goals are involved: the authorities, users, non-users, and operators. And not all users or non-users are identical since their traveling needs vary significantly in space and time. They are also different in their socioeconomic characteristics, which affects their choices: gender, age, income, knowledge of the system, and disabilities. Also, users and non-users interact in the city in a space that is increasingly limited: road congestion and limited vehicular capacity implies that each traveler's decision will affect the experience of many others. The urban context in which all this activity happens is very dynamic and often unpredictable, so such key elements as demand and travel times follow inherently stochastic time-dependent patterns.

The last decades have seen the development of new technologies aimed at improving the information available for planning and operating public transport systems: fare-box and Automated Fare Collection systems (AFC); Automatic Passenger Counter systems (APC); and Automated Vehicle Location systems (AVL), and Geographical Positioning Systems (GPS), among others. These tools are increasingly being installed in public transport systems. Their use has allowed a better understanding of the impact of decisions and an improved performance. The large amount of data that these systems provide has become an important asset that could be exploited more intensively.

The primary trade-off faced in the planning and operating tasks is between the level of service faced by the user and the operating costs for agencies. Users expect to reach their destinations quickly where the total travel time is decomposed into access and egress to/from the stops, waiting, in-vehicle traveling, and transferring. The user also expects the trip to be reliable and hopefully comfortable. For this model to be realistic, these dimensions should be adequately weighted in terms of each user's behavior (Ortúzar and Willumsen, 2001; Raveau et al., 2011; Cepeda et al., 2006; Schmöcker et al., 2011). In case of operators, they are interested in a profitable system where costs of vehicles usage and drivers wages are low.

The planning process spans every decision that should be taken before the operation of the system, and is known as the Transit Network Planning problem denoted as TNP. Due to its complexity, TNP is commonly divided into the following sub-problems that span tactical, strategical, and operational decisions (Desaulniers and Hickman, 2007; Ceder, 2007).

- Transit Network Design (TND): Defines the lines layouts and associated operational characteristics such as rolling stock types and space between stops in order to optimize specific objective functions such as the minimization of the weighted sum of operators' and users' costs. Notice that in this strategic process, frequencies must also be preliminarily set, but they are later adjusted in the Frequency Setting problem.
- Frequency Setting (FS): Characterizes the periods of operation based on demand patterns (morning peak, morning non-peak, afternoon peak, and so on) and determines the number of trips per hour needed to satisfy the passenger demand in each planning period.
- Transit Network Timetabling (TNT): Defines arrival and departure times of buses at all stops along the transit network in order to achieve different goals such as: meet a given frequency, satisfy specific demand patterns, maximize the number of well-timed passenger transfers, and minimize waiting times. In some cases, the number of trips is given while other problems may also determine the number of trips based on vehicle capacity and demand patterns.
- Vehicle Scheduling Problem (VSP): Determines the trips-vehicles assignment to cover all the planned trips such that operational costs based on vehicle usage are minimized.
- Driver Scheduling Problem (DSP): Defines daily duties that cover all the scheduled trips and minimize the cost of driver wages. A solution of the DSP must satisfy specific labor regulations for drivers such as minimum/maximum work length, maximum working time without a rest, and daily rest for all drivers.
- Driver Rostering Problem (DRP): Given a set of generic duties defined over a certain time horizon (e.g., a month) for the drivers assigned to a particular depot, the DRP assigns these duties to the available drivers to their work schedules, called rosters, such that labor regulations are satisfied and driver wages are minimized.

The interdependence between the sub-problems of the TNP is represented in Fig. 1. For example, different frequencies may imply different vehicle schedules and driver duties which strongly influence operational costs. Therefore, an integral approach considering all decisions towards solving the TNP would be desirable. However, each sub-problem may be complex to solve thus, the usual approach to obtain a solution for the TNP is to implement sequential methodologies. These methods are iterative procedures that solve one or more sub-problems in each iteration, taking as input the current solutions obtained for previous sub-problems.

Even when a solution for the planning process is given, operation of transport systems is affected by uncertainty of travel times, demand patterns and disturbances such as: public protest activities or parades; vehicle breakdowns; weather conditions; and activities of specific days. To address these situations, real-time control strategies are implemented to guarantee an efficient service during the operation of the system. The most widely implemented strategies are the following:

- Station control: are strategies that take decisions such as hold vehicles in some stops or skip specific stops in order to increase service regularity, reduce travel or waiting times, and allow passenger transfers.
- Inter-station control: decisions that should be taken between stops (also called stations) of the transit networks, i.e. while the bus is traveling between control stops, in order to achieve an efficient operation. Some strategies belonging to this category are operating speed control and traffic signal priority.

The previous control strategies have greatly benefited by the advances in technology and the development of monitoring tools such as AVL, APC, and GPS. Using these tools, decision makers are able to know the actual behavior of the transit network and implement real-time control strategies.

Each one of the decision levels-strategic, tactical, and operational-as well as real-time decisions impact the efficiency of the system which is an important element in achieving a sustainable public transportation system in urban areas. Indeed, the efficient operation of transport systems not only influences the levels of pollutants, fuel use, and noise, but it can also encourage interaction between individuals and reduce the segregation of communities by establishing communication channels between them. Because of these potential benefits, developing efficient tools for the planning and operation of transport systems remains a challenging research area which requires the careful considerations of different characteristics in the context of different cities.

In this paper, we review the literature on the planning, operation, and control of urban transport systems based on buses. Our goal is to provide a comprehensive classification of the studies presented in the literature by meeting two particular

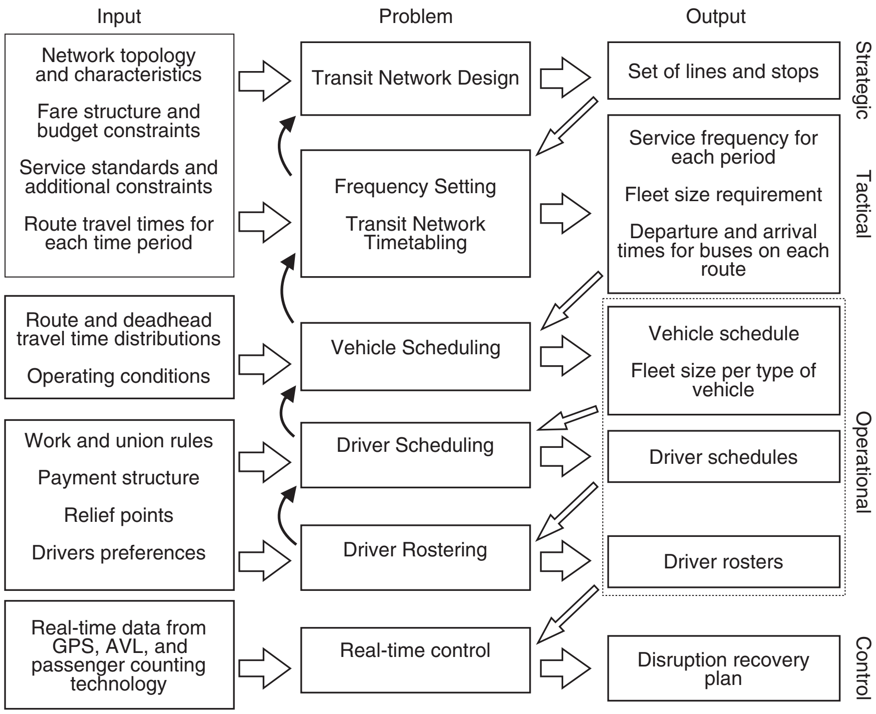

Fig. 1. Interaction between stages of the planning process and real-time control strategies.

goals: (i) complement previous reviews emphasizing foundational studies, recent papers, and some work previously overlooked and (ii) try to incorporate most relevant approaches to address the aforementioned problems. In this review, we complemented our previous knowledge of the field by using the SCOPUS search engine to identify new studies, and the paper they use as references. This method proved to be a successful methodology for the purposes of this review. Some of the journals more extensively used include: *Transportation Research* (all parts), *Transportation Science, Public Transport, Transportation Research Record, Journal of Transportation Engineering, European Journal of Operational Research, and Operations Research*. Occasionally, we will refer to rail-based works that are applicable to the case of buses.

The structure of this paper is as follows. Section 2 presents a literature review for the Transit Network Design Problem. Section 3 focuses on tactical decisions embedded in Frequency Setting and Transit Network Timetabling. Section 4 is based on the sub-problems of the planning process that minimize operational costs considering vehicles' usage and driver wages. Section 5 presents integration approaches for sub-problems of the TNP. Section 6 regards studies of different real-time control strategies. Finally, Section 7 presents some conclusions drawn from our literature review regarding different areas of further research.

2. Strategic planning decisions

Long-term decisions in the Transit Network Planning process are the focus of the Transit Network Design which *determines the lines, types of vehicles, and stop spacing to meet population's movement requirements*. A representation of an urban zone can be defined based on a network $N(V,A)$ where the set of nodes V represents either specific points or a geographical zone called centroid while arcs $(i,j) \in A$ represent a transportation mode between nodes i and j thus, a line l is a set of connected arcs. Common inputs are the following: potential stops; estimated travel times for each arc $(i,j) \in A$; available budget; types of buses and their capacities; and an origin–destination matrix denoted as OD which represents the demand along the nodes in the network N. Then, each element $(i,j) \in$ OD denotes the number of passengers that need to travel from i to j during a planning period (the demand of a centroid node $i \in V$ can be properly estimated proportionally to the distance between the centroid i and the real origin/destination points).

In real life, passengers choose from different routes (that may consist of different trip legs covered by different lines) to travel from a specific origin to a specific destination in order to optimize their own criteria. Fig. 2 shows an example with four routes that cover a specific origin-destination pair $(i,j) \in OD$: (i) a direct trip via line 2; (ii) a trip starting with line 2 and then, transferring to line 1 at transfer point T_1; (iii) a trip starting along line 3 and then, transferring to line 1 at transfer point T_2; (iv) a direct trip via line 4. The previous passenger choices are embedded in an optimization problem called transit assignment problem and it leads to an endogenous passenger demand for the bus network. Even though most authors

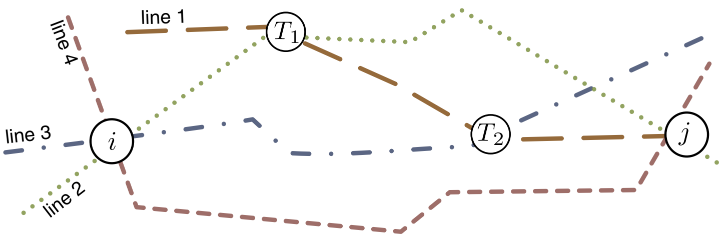

Fig. 2. Transit lines that passengers could use to travel from origin i to destination j.

assume an inelastic demand, i.e. demand independent of the level of service, the demand for each line will depend on the level of service provided. In addition, user route choices depends on their individual valuation of different attributes such as: access, waiting, travel time, transfers and comfort. These weights may not coincide with those used by the system designer. For example, it is not uncommon for authorities to minimize total travel time in the system weighting different stages of the trip identically, while users are well known to weight waiting and walking heavier than traveling inside the vehicle. Also, the design may not give any value to comfort (just assume a fixed capacity per vehicle) while we know it heavily affects users preference once vehicle occupation approximates capacity. Moreover, users are not homogeneous in their preferences; i.e. different users assign different weights to these attributes.

Another challenge of this problem is the presence of congestion in which the level of service of a user is affected by the decisions of other travelers in the system. Other users may cause increasing waiting time, discomfort inside a vehicle or preventing entering a vehicle whose capacity has been reached. In these cases the minimization of the sum of the total costs of all users may not represent the interest of some individual users which may be forced to take lines, that would be best for the system as a whole, but will not represent their best alternative. Thus, under the presence of congestion, we should recognize that individual choices might not coincide with the system optimal. A common assumption to simplify these decisions are homogenous passengers taking the minimum cost route (all or nothing).

From the user's perspective, a transit network should cover a large service area, be highly accessible, offer numerous direct-through trips, not deviate much from the shortest paths, and should globally be able to meet the demand. From the agency's perspective, the transit network must stay within the current budget, be efficient in obtaining revenues, and be able to satisfy the demand. Hence, TND considers elements such as the service area, line coverage, spacing of stop, directness of the lines, line lengths, policy headway (separation time between consecutive trips), loading standards, and road speeds (see Fan and Machemehl, 2004).

The TND problem can be modeled as a discrete optimization problem to determine the set of lines for a given OD matrix and predicted travel times per link on the network. An alternative to this approach is to represent the TND through analytical models which represent continuous approximation approaches where passenger demand is represented as a continuous function over a geographical space instead of an OD matrix (see Clarens and Hurdle, 1975; Daganzo, 2010; Ouyang et al., 2014). Using these approaches, decisions are made in regard to the optimal spacing between lines and their frequencies. The solution of a continuous approximation is based on an idealized representation of the city which considers specific grid structures such as rectangular, circular, hub-and-spooks, and so on. In this section, we will present studies that complement reviews of Desaulniers and Hickman (2007), Guihaire and Hao (2008a), Kepaptsoglou and Karlaftis (2009), Farahani et al. (2013).

2.1. Continuous approximations for the Transit Network Design

As stated by Kepaptsoglou and Karlaftis (2009), analytical approaches develop relationships between components of the public transportation network represented with idealized structures. The first studies focused on determining the line spacing, stop spacing, and line frequency, assuming demand is uniformly distributed along the service area (Holryod, 1967; Byrne and Vuchic, 1972). Other approaches were developed to consider time dependent demand (Salzborn, 1972), space dependent demand (Byrne, 1975; Black, 1979), and time-and-space dependent demand (Hurdle, 1973). Later, Chien and Schonfeld (1997) define a TND problem to determine the stop locations as well as the lines, and their headways in order to minimize total user and operator costs. To avoid oversimplifying urban areas, the authors propose an analytical model which assumes a rectangular grid and zone dependent passenger demand.

Chien and Spasovic (2002) consider elastic demand, i.e. demand dependent on the level of service to determine the line and stop spacing, headways, and fares. Their objective function is to maximize the social welfare, i.e. consumer surplus plus the operators' profit. The authors assume irregular service regions, many-to-many demand patterns, and vehicle capacity constraints.

Daganzo (2010) propose a transit network structure for an idealized uniform demand square region, combining the grid and the hub-and-spoke concepts. The experimental results suggest the following: a hub-and-spoke structure becomes more attractive in cities where infrastructure is expensive; Bus Rapid Transit systems are a good option for cities with medium to high density, recommending it over metro as long as the city has enough suitable streets; user costs dominate over operational costs and both decline with low demand densities; and bus stop spacing and inter line distance should be kept over a critical threshold. Later, Estrada et al. (2011) implemented the methodology of Daganzo (2010) to generate lines for the city of Barcelona. Moreover, the authors develop a simulation stage to map the transit network in the city, obtaining a more accurate estimation of costs and level of service.

More recently, Medina et al. (2013) formulate an optimization problem to simultaneously determine the stop density in a bi-directional corridor and the lines' frequency for several periods. The objective is to minimize the user costs and operating costs based on waiting time, in-vehicle travel time, fleet operating costs, and stops installation. Since the proposed mathematical formulation is a non-linear optimization program, they propose a sequential approach consisting of two steps: (i) determining the optimal headway replacing the stop density decision variable with the analytic expression obtained from the first-order conditions of the optimization problem; and (ii) finding the density of stops along the corridor. The proposed approach is implemented to analyze a corridor in the transit network of Santiago, Chile. Numerical results show that the proposed set of stop locations results in a reduction of total costs by about 20% and longer headways than the current ones (an increase of more than 60%), especially in the off-peak period.

To address different densities across a city, Ouyang et al. (2014) propose a hierarchical system in which a city-wide network of few lines with long bus stop spacing and long distances between lines is complemented by more lines visiting local high density neighborhoods. The authors minimize the generalized cost based on buses' idle times, total traveled distance, total vehicle usage time, walking time for passengers, waiting times at stops, in-vehicle travel times, and penalties for transfers. Numerical results show that the heterogeneous network configurations produce lower generalized costs than the homogeneous grids.

Kim and Schonfeld (2014) propose a probabilistic analytical model to design and synchronize a bus transit network that must provide service from a major terminal to multiple regions. The decisions which need to be made in this problem include the type of service provided for particular regions (fixed lines and a fixed schedule or demand responsive lines and a flexible schedule), the vehicle size, the number of zones, frequency, fleet type, and slack times. A GA is developed to solve examples of the problem and numerical results show that the solutions leading to an optimal synchronization have either common headway or integer-ratio headways.

2.2. Discrete optimization approaches for the Transit Network Design

To overcome passenger demand endogeneity in discrete optimization versions of the TND, several approaches have been considered. One approach is decomposing the problem into sequential subproblems; e.g. define potential lines first and then, solve the transit assignment problem for the proposed network and iterate hoping for convergence. Another approach to achieve a user-operator equilibrium consists in modeling the TND as a bi-level optimization problem, where the design of lines is carried out in a "first level" and the transit assignment used to determine the passenger flow in the network is determined in a "lower-level" (e.g., Chiou, 2005). A third approach models the passenger assignment decisions in the TND using non-linear constraints or non-linear objective functions, leading to non-convex problems which can be solved through heuristic algorithms. In this section, we present studies addressing these three types of approaches.

2.2.1. Sequential approaches and simplified cases for the Transit Network Design

One of the early approaches to solve TND is to decompose it into line generation and transit assignment stages. For example, Marwah et al. (1984) assign the demand to lines of the network in a first stage and then design a set of potential bus lines to select the ones minimizing the number of transfers. Another example is van Nes et al. (1988), where a heuristic procedure starts with a given set of potential lines: the one with the highest number of direct trips is then selected, and the frequency on that line is increased until it reaches the budget limit and the fleet size limit. Similar decompositions are proposed by Baaj and Mahmassani (1991, 1992, 1995) to solve the TND with the objective of maximizing the number of direct trips and minimizing transfer times.

Mathematical formulations for the Transit Network Design are usually intractable by exact approaches. However, it is possible to solve them in an exact way by considering cases such as small instances (Wan and Lo, 2003), single-line design cases (Guan et al., 2003), simplifications that reduce the size of the transit network (Lownes and Machemehl, 2010), and determine just some decisions, for example, stop spacing (Furth and Rahbee, 2000). Then, advanced meta-heuristics are commonly implemented to tackle TND problems, some times, formulated with bi-level models to consider transit assignment decisions.

2.2.2. Metaheuristic algorithms for the Transit Network Design

Although TND is often decomposed to solve each stage through heuristic algorithms (see Ceder and Wilson, 1986; van Nes et al., 1988), it is important to define the feedback between the different stages trying to integrate the decisions of the line design with those of the transit assignment. This can be naturally embedded in Genetic Algorithms (GA) (see Chakroborty, 2003; Szeto and Wu, 2011). For example, Chakroborty (2003) solves the line generation and Frequency Setting stages using GA consisting of two main modules: (i) line exploration through random line generation and movements of cross-over and mutation, and (ii) line evaluation through a fitness function that determines the demand for the current lines, the direct trips, and the average travel time, in order to get a measure of the set of lines. We stress that the fitness function of the GA algorithms may be complex to evaluate since it may contain another optimization problem. Then, Argwal and Mathew (2004) propose the parallelization of the fitness function of a GA to solve a case of the TND with the objective of minimizing the costs based on travel time and round trip distances with bounded frequency, limited load factor, and demand satisfaction constraints. The performance of the GA is improved, since the computational time is reduced by increasing the number of processors.

Bagloee and Ceder (2011) also address a population-based approach for the TND with additional considerations. Given a road network, an exogenous fixed transit demand, and a budget, the problem is to derive a comprehensive transit plan (including lines and assignments of vehicles of different types) so that the total passengers' discomfort is minimized. The authors propose a heuristic approach, consisting of three components: (i) locating the stops based on proximity to highly concentrated transit demand places a clustering factor of nodes; (ii) generating a set of candidate lines passing through those stops; and (iii) running a search algorithm over the candidate lines to seek a good solution. This search procedure is encoded by means of a GA and equipped with Ant-System (AS) collective points as the search engine. The solution approach is tested on instances up to 903 nodes, 2975 links, and 5394 OD pairs.

Yu et al. (2012) extend the direct traveler density model of Yang et al. (2007) to design a transit network that maximizes the number of direct travelers carried per unit length of a line (passengers/km). The optimization approach consists of the following three stages, with feedback between them: (i) generate potential lines that satisfy directness, passenger demand, and line length constraints; (ii) lay main lines into the transit line network according to the maximum traveler density; and finally (iii) place branch lines in the transit network, including skeleton lines and main lines according to the maximum traveler density. The authors implement an Ant Colony Optimization algorithm (ACO) to design the transit network while a transit assignment problem based on minimum travel distance is solved to obtain the passengers flows. The proposed approach is tested using data from a Chinese city and numerical results show that the optimized transit line network is an improvement, in terms of transfers, on the existing transit network.

Fan and Machemehl (2006b) assumed fixed demand, i.e. demand not dependent on the level of service. The authors addressed the TND to minimize the weighted sum of operator cost, user cost, and unsatisfied demand cost subject to constraints such as headway bounds, load factor constraints, limited line length, and fleet-size bounds. A Simulated Annealing algorithm is implemented to solve the proposed problem. Fan and Machemehl (2006a) and Fan and Machemehl (2008) enhance the problem definition of Fan and Machemehl (2006b) by considering elastic demand, i.e. demand dependent on the level of service. These studies implemented a Network Analysis Procedure (NAP) consisting of assigning transit trips, determining line frequencies, and computing performance measures. Fan and Machemehl (2006a) develop a GA implementing the NAP as the fitness function while Fan and Machemehl (2008) design a local search algorithm consisting of (i) implementing an heuristic procedure to generate a feasible solution and (ii) implementing an iterative local search to generate the lines and their frequency as well as the NAP to re-assign the demand based on the current lines.

Cipriani et al. (2012) address the TND with elastic demand to define lines, frequencies, and vehicle sizes maximizing the total welfare. The authors propose a solution approach consisting of two stages: (i) implementing a heuristic algorithm to generate potential lines and their frequencies and (ii) a GA that recombines lines to generate new population individuals while the fitness function evaluates them using a probabilistic modal split model which determines the mode choice behavior of users and a hyper-path transit assignment model which determines the route choice behavior of users.

Recently, Szeto and Jiang (2012) implement three versions of the Artificial Bee Colony algorithm (ABC) (see details of ABC in Karaboga et al. (2012)) to solve the TND, minimizing the weighted sum of the number of transfers and the total travel time of the users. The ABC algorithms, together with a stop-sequence improvement heuristic, are used to design lines, whereas a neighborhood search heuristic is implemented to solve the Frequency Setting problem, allowing the exploration of different potential lines. Numerical results show that the frequency of each line found by the best proposed algorithm is reduced by 5.5%, without increasing the fleet size, compared with the other two population-based algorithms proposed. However, the study lacks any comparison with other algorithms in the literature. Another ABC is implemented by Nikolić and Teodorović (2013) to solve the TND with static demand that minimizes the number of transfers and the total travel time. The authors define a passenger as satisfied if two or less transfers are needed in his/her trip. They solve test instances and compare their results with previous approaches, such as Baaj and Mahmassani (1991) and Chakroborty and Dwivedi (2002). Experimental results show that the ABC improves (compared with the current operation) the percentage of demand satisfied without transfers, and the satisfied demand with one transfer, in three out of four instances. Improvements are obtained also for the average travel time. Later, Nikolić and Teodorović (2014) extend their previous approach outlined by Nikolić and Teodorović (2013) considering elastic demand and the minimization of the weighted sum of the total number of unsatisfied passengers, the total travel time of all passengers, and the fleet size.

Nayeem et al. (2014) develop two versions of a GA with elitism to solve the TND with static demand which minimizes the weighted sum of unsatisfied passengers, the total number of transfers, and the total travel time of all served passengers. Moreover, the authors also implemented an alternative algorithm that increases the population size after each iteration. The later study present a comparison between the proposed algorithms and the ones by Baaj and Mahmassani (1991), Chakroborty and Dwivedi (2002), Fan and Machemehl (2008), Nikolić and Teodorović (2013). Numerical results show that the GA with elitism outperforms the other approaches compared in the study in all three criteria in most of the tested instances.

2.2.3. Bi-level approaches for the Transit Network Design

Gao et al. (2004) address the TND where the upper-level define lines minimizing the weighted sum of the average travel time, average waiting time, and operational cost while the lower-level is a transit assignment problem. The authors propose an algorithm based on a sensitivity analysis to solve some numerical examples. Later, Fernández et al. (2008) present a Transit Network Design problem in the context of Santiago, Chile. The TND has the objective of minimizing the total cost based on the lines' travel times, vehicle operation, and users' travel times. Due to endogenous characteristic of demand in the Network Design problem, the authors propose a heuristic approach consisting of: (i) determining the potential network topologies and then, (ii) measuring each topology by solving the bi-level problem which defines the Frequency Setting at the upper-level and the transit assignment decisions at a lower-level.

Recently, Szeto and Jiang (2014) formulate the TND where the upper-level defines the bus lines minimizing the number of passenger transfers and the lower-level problem is the transit assignment problem with capacity constraints. As in the previous study of the authors for a single level problem (Szeto and Jiang, 2012), a Bee Colony Algorithm is also implemented to solve the optimization problem. The proposed approach is tested on transits networks in Tin Shui Wai, Hong Kong and Winnipeg, Canada.

Re-design of transit networks may be useful for old transport systems. For example, Fan and Machemehl (2011) state that significant spatial redistribution and demographic changes have been taking place in most U.S. cities, making the land use patterns increasingly decentralized. Thus, current transit networks are not necessarily suitable for cities with multi-centered and spatially dispersed trip patterns. To overcome this situation, the authors propose a bi-level optimization model to solve the Transit Network Re-Design problem. The upper-level determines network structures minimizing the weighted sum of use cost, operator cost, and unsatisfied demand, while the lower-level is a transit assignment problem. In this formulation, the improvement of the current minimum travel time for each OD pair is modeled as a hard constraint within the upper-level problem. To solve the problem, the authors develop a GA and numerical results show that the spatial equity increases the total travel cost in the network.

2.2.4. Robust and stochastic Transit Network Design

In practice demand is uncertain. To tackle this issue, robust optimization, chance-constrained optimization, and stochastic optimization have been implemented. For example, Chen and Yang (2004) address a TND to minimize the total travel time considering equity issues and demand uncertainty. The authors assume that the passenger demand fits a Normal distribution probability and they propose two different stochastic formulations: (i) one where they minimize the expected value of the total travel time and (ii) a stochastic version minimizing the total travel time considering chance-constraints associated with equity performance. To solve the problems the authors developed an iterative solution approach consisting of a GA algorithm to solve the TND and a simulation procedure to simulate the demand uncertainty. Numerical results show that there is a trade-off between the spatial equity measure and the expected value of the total travel time.

Amiripour et al. (2014) determine a set of lines to be implemented for an entire year considering seasonal demand patterns with a probability of occurrence. The authors propose a mathematical formulation to minimize the expected value of the weighed sum of passengers' total waiting time, unused seat capacity, unsatisfied demand (passengers with more than a specific number n of trip legs), and the fleet size. Constraints of the problem for each scenario are headway bounds, limited fleet size, maximum number of transfers, and maximum unsatisfied demand. To solve the proposed formulation the authors develop GA algorithm where the fitness function solves a transit assignment problem.

Stochastic travel times are considered by Yan et al. (2013), which addressed a robust optimization version of the TND minimizing the weighted sum of the expected value of the operator cost and its variability. To solve the problem, the authors implemented a solution methodology consisting of the following four components: (i) a k-shortest path algorithm to generate candidate lines; (ii) a Simulated Annealing algorithm to solve the transit assignment problem; (iii) a Monte Carlo simulation to define the passenger flow loading; and (iv) a probit-type discrete choice model embedded in the simulation to determine the passengers' choices.

2.2.5. Multi-objective Transit Network Design

Recent studies of TND problems focus on the analysis of the compromises between different evaluation criteria. For example, Fan and Machemehl (2004) implement a decomposition approach to solve the TND. However, a contribution of that study is an extensive experimentation stage considering different weight parameters for an objective function based on users' and operators' costs. Using this experimentation, the authors analyze the trade-offs between different objectives since weighted objective functions can be used to approximate the set of Pareto optimal solutions. Mauttone and Urquhart (2009) propose a multi-objective optimization approach for the TND minimizing users' costs (based on in-vehicle time, waiting time, and transfer time) and the fleet size. The authors propose a GRASP algorithm for the multi-objective optimization problem. Numerical results on instances presented by Mandl (1979), show that the proposed heuristic is capable of obtaining more Pareto optimal solutions that using a weighted objective function. Recently, Chew et al. (2013) formulate a bi-objective version of the TND that minimizes the user costs and operator costs subject to constraints such as a limited number of lines, limited number of line nodes, and a limited percentage of passengers using several transfers. To solve the proposed approach, a GA was designed and tested on a Benchmark data set.

Tables 1 and 2 summarizes our literature review for the TND. The second and third columns present the objective and constraints considered in the formulation addressed in each paper. The fourth column exhibits the solution approach.

Table 1
Literature review for the Transit Network Design problem.

Authors (year)	Objective	Constraints of the system	Solution method	Case
Clarens and Hurdle (1975)	Min total users' and operators' costs	Vehicle capacity	Analytical	Example
Marwah et al. (1984)	Min passenger riding-time and operation costs	Demand satisfaction	Decomposition & Heuristic	Test
Ceder and Wilson (1986)	Min excess travel time, transfer and waiting time	Minimum frequency	Heuristic	Examples
van Nes et al. (1988)	Max number of direct trips	Budget and fleet size	Decomposition & Heuristic	Example
Baaj and Mahmassani (1991, 1992, 1995)	Max number of direct trips	Fleet size and headway bounds	Decomposition & Heuristic	Example
Chien and Schonfeld (1997)	Min total users' and operators' costs	Vehicle capacity	Analytical	Example
Furth and Rahbee (2000)	Min walking distance, riding delay, and operator costs	Maximum stop distance	Dynamic programming	Real
Chakroborty and Dwivedi (2002)	Min unsatisfied demand, total travel time, and passengers with more than two transfers	Route feasibility	GA	Examples
Chien and Spasovic (2002)	Min consumers' surplus and operators' profit	Vehicle capacity	Analytical	Example
Chakroborty (2003)	Min passenger waiting time	Fleet size, capacity, policy headway, and maximum transfer time	Decomposition & GA	Example
Guan et al. (2003)	Min in-vehicle travel time and number of transfers	Demand satisfaction	Solver	Benchmark
Wan and Lo (2003)	Min operating costs	Capacity and bounded frequency	CPLEX	Example
Argwal and Mathew (2004)	Min cost based on travel times and trip distances	Bounded frequency, load factor, and demand satisfaction	Parallel GA	Real
Chen and Yang (2004)	Min total travel time	Spatial equity and budget	GA & Simulation	Example
Fan and Machemehl (2004)	Min weighted sum of user and operator costs	Headway bounds, load factor bounds, fleet size, trip length, and number of lines	Decomposition & Heuristic	Examples
Gao et al. (2004)	Min weighted sum of user and operator	Minimum frequency	Sensitivity analysis	Example
Chiou (2005)	Min travel and construction costs	Link capacity	Gradient-based method	Examples
Fan and Machemehl (2006a)	Min total users' and operators' costs	Headway bounds, load factor bounds, line length, and fleet size	GA	Example
Fan and Machemehl (2006b)	Min unsatisfied demand, users' costs, and operators' costs	Load factor, headway bounds, line length, and fleet size	SA	Example
Yang et al. (2007)	Max direct travels per unit length	Bus capacity, minimum direct travelers, and stop spacing	ACO	Examples
Fan and Machemehl (2008)	Min total users' and operators' costs	Headway bounds, load factor bounds, line length, and fleet size	TS	Example
Fernández et al. (2008)	Min cost based on travel times, vehicle operations, and users' travel time	Load factor and demand assignment	Heuristic	Real
Mauttone and Urquhart (2009)	Min users' cost vs. minimum fleet size	Frequency bounds and load factor	GRASP	Example

Finally, the fifth column presents the type of instances used for the experimental stage. In that column, there are several categories where "Example" represents small instances designed by the authors or simplified networks of real systems; "Real" represents the implementation in real transit networks; "Test" means that the authors use instances that are generated randomly or based on real contexts; and "Benchmark" represents comparing the approach with instance sets from the literature. The previous format to summarize our review was used in each section of this paper except for the Real-Time Control strategies.

3. Tactical planning decisions

Tactical decisions associated to the TNP are related to improving the level of service and reducing the operational costs. In this planning horizon, which is normally few months, the fleet size and characteristics is fixed. In addition, most lines and the origin-destination demand matrix are considered given. In this context, the remaining questions are: (i) the frequency and fleet (number and type of vehicles) assigned to different lines in each period; (ii) the timetables for low frequency lines; and (iii) the design of operational strategies to be considered during the execution.

3.1. Frequency Setting problem

In the previous section, we address the TND problem which usually defines preliminary frequencies. However, the operation of a transport system is dynamic, and so the service policies must be adjusted to the demand behavior and system

Table 2
Literature review for the Transit Network Design problem (continue).

Authors (year)	Objective	Constraints of the system	Solution method	Case
Lownes and Machemehl (2010)	Min weighted sum of user, operator, and unmet demand costs	Capacity and fleet size	TS	Example
Daganzo (2010)	Min total users' and operators' costs	Number of corridors, space between them,	Analytical	Examples
Bagloee and Ceder (2011)	Min passenger's discomfort	Demand satisfaction and budget	Decomposition & GA	Example
Estrada et al. (2011)	Min total users' and operators' costs	Number of corridors, space between them, and systems capacity	Analytical	Simulation
Fan and Machemehl (2011)	Min weighted sum of user, operator, and unmet demand costs	Trip length, number of lines, headway bounds, Load factor, fleet size, and spatial equity	GA	Examples
Szeto and Wu (2011)	Min weighted sum of number of transfers and total travel time	Fleet size, frequency bounds, number of Intermediate stops, and maximum travel time	GA	Real
Cipriani et al. (2012)	Min total users' and operators' costs and line length	Bus capacity, minimum frequency,	GA	Real
Roca-Riu et al. (2012)	Min planner and user costs	Demand satisfaction and speed bounds	TS	Real
Szeto and Jiang (2012)	Min weighted sum of number of transfers and total travel time	Fleet size, frequency bounds, number of Intermediate stops, and maximum travel time	ABC	Real
Yu et al. (2012)	Max trip density	Demand satisfaction	ACO	Real
Chew et al. (2013)	Min users' costs vs Min operators' costs	Number of stops and lines and percentage of passengers using n transfers	GA	Benchmark
Medina et al. (2013)	Min total users' and operators' costs	Bus capacity, stop capacity, and equivalence of stop density in each direction	Decomposition & Analytical	Real
Nikolić and Teodorović (2013)	Min weighted sum of number of transfers and total travel time	Demand satisfaction	ABC	Benchmark
Yan et al. (2013)	Min expected value of operator cost and its variability	Fleet size, bus capacity, line feasibility, and travel time reliability	SA	Examples
Amiripour et al. (2014)	Min total waiting time, fleet size, total differences from shortest paths, and total empty-seat hours	Budget, maximum unsatisfied demand, headway bounds, number of transfer, and bounded deviation from short paths	GA	Examples
Kim and Schonfeld (2014)	Min total users', operators, and transfer costs	Bus sizes and number of zones	Analytical	Examples
Ouyang et al. (2014)	Min users' generalized costs	Fleet size	Analytical	Examples
Nayeem et al. (2014)	Min unsatisfied passengers, number of transfers, and travel times	Demand satisfaction	GA	Real
Nikolić and Teodorović (2014)	Min weighted sum of unsatisfied passengers, total travel time, and fleet size	Demand satisfaction, link capacity, and vehicle capacity	ABC	Benchmark
Szeto and Jiang (2014)	Min transfers	Bounds for frequency, capacity and the fleet size	ABC	Real

performance, i.e. it is necessary to make changes in the network design, frequencies, and fare structure, among other things. The Frequency Setting problem (denoted as FS) *determines the number of trips for a given set of lines L to provide a high level of service in a planning period*. Since frequency changes influence the passengers' perception of the level of service it may lead to an increment or decrement of the system usage.

A simple version of the problem is to determine the frequency f^l (buses/hour) for each line $l \in L$ such that travel times are minimized and the fleet size is bounded. This problem can be formulated as follows.

$$[P] = \left\{ \min \sum_{i \in V} \sum_{j \in V} D_{ij} T_{ij}^l : \sum_{l \in L} \lceil t^l \cdot f^l \rceil \leqslant F \right\} \tag{1}$$

where T_{ij}^l is the travel time between i and j via line l, D_{ij} is the demand for the origin–destination pair (i,j), t^l is the cycle time of line l, and F is the maximum fleet size.

Early works on Frequency Setting when we assume a fixed demand-line assignment were based on analytic models (see Newell, 1971; Salzborn, 1972; Scheele, 1980; Han and Wilson, 1982) or heuristic solution methods (see Furth and Wilson, 1981). These seminal approaches were later extended including uncertainty in the demand observed in each line or other considerations.

Hadas and Shnaiderman (2012) address the minimization of a total cost based on empty seats and not-served demand. Thanks to GPS, APC, and AVL tools, the authors defined probability distributions for travel times and passenger demand. Based on this information, they defined an analytical optimization approach that determines frequencies and vehicle sizes. Implementing the proposed approach in examples shows that the most significant cost reduction is obtained in cases with low-level of service. Li et al. (2013) also consider stochastic parameters such as demand, arrival times, boarding/alighting times, and travel times. The authors define a stochastic optimization approach to find the optimal frequency that minimizes the sum of the expected value of the company profit and the waiting time costs for passengers. The authors develop a GA to solve the problem and compare their approach with traditional Frequency Setting models of Newell (1971) and Ceder (1984). The authors state that the headways obtained are usually larger than those using the approach by Newell (1971) and shorter than those using Ceder (1984). They claim that these moderate headways provide a better balance between the bus operational costs and the passengers' satisfaction.

Verbas and Mahmassani (2013) and Verbas et al. (2015) extend the model presented by Furth and Wilson (1981) considering demand variation along time and line route. The authors re-define a line as a set of "line patterns"-consisting of subset of stops of the normal line-for which a frequency must be set in order to properly satisfy demand patterns in specific time intervals. The variation of demand is modeled by assuming temporal and spatial heterogeneity of ridership elasticity with respect to frequencies. The problem is formulated with a non-linear program which minimizes the weighted sum of ridership and wait time savings over all stops, lines, and time intervals subject to constraints such as budget, fleet size, headway bounds for each line pattern, and bounds for load factors. Verbas and Mahmassani (2013) test the model on an example in order to analyze the impact of the constraints. Numerical results show that increasing the fleet size may lead to reduce operational costs since vehicles may be assigned to low-cost, high-ridership line patterns. Verbas et al. (2015) test the model on large instances in order to analyze the impact of temporal/spatial demand elasticities and the authors state that demand is not temporally stable.

In the previously mentioned approaches, transit assignment decisions are not considered. To address these them, Martínez et al. (2014) propose a MILP to jointly determine the frequency and demand over each line. The problem considers discrete frequencies, uncapacitated vehicles, and limited fleet size. The major contribution of this study is the modeling of transit assignment decisions assuming that each passenger choose his/her minimal travel time path. The proposed formulation has a special structure that allow decomposition into independent problems where frequencies are fixed (but this is not explored by the authors). The model is tested on the real transit network of Rivera, Uruguay and numerical results show that the model diminishes the travel times 13% compared with the current operation of the system. Since the model is intractable for large instances, the authors develop a Tabu Search heuristic which is tested on the transit network of Montevideo, Uruguay where solutions lead to improvements of 1.7% compared with the current operation.

Other approaches including transit assignment decisions in FS are based on bi-level optimization models. Their hierarchical nature allows determining frequencies at an upper-level and transit assignment decisions at a lower-level.

3.1.1. Bi-level approaches for the Frequency Setting

Constantin and Florian (1995) define the upper-level as determining the frequencies which minimize the total travel and waiting times while at the lower-level a transit assignment problem is modeled. Then, they develop a projected sub-gradient algorithm to solve instances based on the transit networks of Stockholm, Sweden; Winnipeg, Canada; and Portland, OR, U.S.

dell'Olio et al. (2012) address the problem of determining the vehicle capacity and the needed frequency in order to satisfy a given demand. The upper-level considers the optimization of the sum of user and operator costs while the lower-level represents transit assignment decisions. To solve the optimization problem, the authors propose an iterative heuristic approach consisting of three steps at each iteration: (i) generate an initial set of frequencies; (ii) solve the transit assignment model using commercial software; and (iii) implement the Hooke–Jeeves algorithm to find new frequencies and determine vehicle capacities. The numerical result of implementing the proposed approach in a sub-network in Santander, Spain quantify the benefits from using an heterogeneous fleet.

Demand uncertainty incorporates a new element into the FS problem since passengers may not board the first arriving bus even though the aggregated frequency is enough to carry all passengers. Yu et al. (2010) propose a formulation that minimizes the total travel time of passengers subject to fleet size constraints and considering users' route choices. The upper-level determines the bus frequencies while the lower-level solves a transit assignment problem assuming uncertain passenger arrivals at bus stops. To solve the proposed problem, the authors implement a heuristic algorithm consisting of two procedures: (i) a GA to compute potential frequencies and (ii) a label-marking method that assigns the demand to the current network configuration according to the minimization of the total expected cost. The proposed solution approach is tested on the transit network of Dalia, China. Numerical results show that it is possible to reduce travel times by 6% compared with the current operation of the system. Moreover, the authors compare the operation of four private companies under two scenarios: (i) an integrated system, i.e. companies sharing vehicles; and (ii) independent operation. Numerical results show a 13% reduction of the total travel time in the integrated case.

Yoo et al. (2010) present a FS model based on a non-cooperative Stackelberg game. The upper-level maximizes the demand, subject to fleet size and frequency constraints. At the lower-level a capacity-constrained stochastic user equilibrium assignment model is solved considering variable demand and transfer delays. To solve the proposed formulation the authors develop an iterative procedure consisting of two stages: (i) solving the lower-level for a given frequency; and (ii) determining new frequencies using a gradient projection method.

Huang et al. (2013) present a FS problem with uncertain demand to minimize the weighted sum of the operating costs and travel times variance. In this study, the upper-level determines the frequencies while the lower-level computes the mean and variance of the passenger flow through each link of the transit network. The proposed problem is solved through

GA and implemented on the transit network of Liupanshui, China which consists of 16 lines, 270 stops, and 260 buses. Numerical results show that the total cost can be reduced by about 6% compared with the current operation.

Verbas and Mahmassani (2015) propose a bi-level solution method which: (i) determines frequencies at the upper-level maximizing waiting time savings considering constraints such as budget, fleet size, vehicle load, and headway policy; and (ii) runs a simulation algorithm at the lower-level modelling the demand response to the new Frequency Setting. The proposed approach is tested on the transit network of Chicago. Compared with the base scenario, the method obtain a reduction of 45.7% in waiting times, and an increasing of 9% in the number of boardings.

3.1.2. Frequency Setting to coordinate different transport modes

Coordination of different lines or transport modes is rarely considered in FS problems, even when coordination affects costs that are imparted to transit operators as well as its users (Sun and Hickman, 2004). In this matter, Shrivastava and Dhingra (2002) and Shrivastava et al. (2002) develop a non-linear integer formulation for the FS of feeder lines connecting trunk lines (with given schedules) in order to minimize transfer times at connection stops and the operational cost. Assumptions of the problem are fixed demand and bounds for transfer waiting times, fleet size, and load factors. Since the problem is intractable by commercial software, the authors develop a GA taking into account the load factor as a quality measure. The proposed algorithm is tested on the transit network of Mumbai, India and numerical results show that there is a trade-off between the load factor and transfer waiting time. The later GA is also implemented by Verma and Dhingra (2006) and Shrivastava and O'Mahony (2006, 2009) in sequential approaches that heuristically generate lines and then coordinate them.

More recently, Sivakumaran et al. (2012) propose a continuous approximation approach which determines the frequency of feeder lines minimizing the weighted sum of wait times at feeder stops, their transfer wait at the trunk stop, and the feeder operating cost. The demand is modeled through a time-independent density function which varies gradually along distance. The authors finds out that by dispatching feeder vehicles in coordination with the given trunk schedule, total user cost can significantly diminish while little or no extra cost is imparted to the operator of the feeder lines. Moreover, they

Table 3
Literature review of the Frequency Setting problem.

Authors (year)	Objective	Constraints of the system	Solution method	Case
Newell (1971)	Min total waiting time	Capacity constraints	Analytical	None
Salzborn (1972)	Min total waiting time and fleet size	Passenger arrival rates	Analytical	None
Scheele (1980)	Min travel time	Capacity and fleet size	Heuristic	Real
Furth and Wilson (1981)	Max costumer utilization	Budget, fleet size, maximum headway	Heuristic	Example
Han and Wilson (1982)	Min maximum load factor	Passenger assignment choice, capacity, and fleet size	Decomposition & Heuristic	Real
Ceder (1984)	Min bus runs	Load profiles and vehicle capacity	Heuristic	Example
Constantin and Florian (1995)	Min expected value of travel and waiting times	Fleet size	Sub-gradient algorithm	Real
Shrivastava and Dhingra (2002)	Min transfer time and operator costs	Bounds for transfer waiting times,	GA	Real
Shrivastava et al. (2002)		Fleet size, and load factor		
Mohaymany and Amiripour (2009)	Min expected waiting time	Fleet size	Simulation	Example
Yoo et al. (2010)	Max demand	Fleet size, frequency, and vehicle capacity	Iterative heuristic	Example
Yu et al. (2010)	Min total travel time	Fleet size	GA	Real
Hadas and Shnaiderman (2012)	Min empty seats and unfulfilled demand	Headway bounds	Analytical	Example
dell'Olio et al. (2012)	Min total users' and operators' costs	Bus capacity and limited number of lines that a bus can cover	Heuristic	Real
Huang et al. (2013)	Min operators' costs and travel time variance	Fleet size and minimum number of trips	GA	Real
Li et al. (2013)	Min expected value of company's profit and waiting time costs	Bus capacity	GA	Test
Martínez et al. (2014)	Min travel time	Demand satisfaction and fleet size bound	CPLEX & TS	Real
Sivakumaran et al. (2012)	Min weighted sum of waiting time, transfer time, and operator cost	Demand satisfaction and vehicles capacity	Analytical	None
Wu et al. (2014)	Min weighted sum of waiting time, transfer cots, drivers/vehicles costs	Frequency bounds	GA	Example
Verbas and Mahmassani (2015)	Min waiting time	Budget, fleet size, and headway bounds, and load factor	Heuristic	Real
Verbas and Mahmassani (2013)	Min weighted sum of ridership and waiting time savings	Budget, fleet size, and headway bounds, and load factor	KNITRO	Example&Test
Verbas et al. (2015)				

extend the analysis to jointly determine the headways for trunk and feeder lines. In this case, it is shown that schedule coordination can often be Pareto improving which benefits both users and operators.

Wu et al. (2014) propose a non-linear formulation for the FS with static demand in order to coordinate feeder bus lines with trunk lines in BRT corridors. The model considers vehicles' capacity and frequencies' bound. The objective is to minimize the weighted sum of users' and operators' cost including waiting time, transfer cots, in-vehicle waiting time, drivers' wages, vehicle operation, and vehicle maintenance. To solve the problem, the authors develop a GA and test it on two kinds of scenarios: (i) coordination of all-stop BRT and bus; (ii) coordination of all-stop BRT, limited-stop BRT, and bus. Numerical results show that in case (i), passenger costs reduce significantly through coordinated operation. However, operation costs may increase caused by the larger number of vehicles needed after implementing coordinated operation. In case (ii), operators' costs are diminish by shorter cycle times but waiting times and transfer times are reduced only when the stopping stops are carefully coordinated.

Other approaches that consider, at some level, the design of lines and Frequency Setting decisions include the design of short-turn lines and limited-stop lines which satisfy the passenger demand more efficiently. We explore these strategies in Section 3.3. Now, we summarize our literature review of the Frequency Setting problem in Table 3.

3.2. Transit Network Timetabling problem

There are two types of transit systems operation: (i) frequency-based operation where a frequency f (buses/hour) should be met with hopefully a regular service, i.e. buses are expected to pass every $\frac{60}{f}$ minutes; and (ii) timetable-based operation where specific departure and arrival times for selected stops are set for all trips. Commonly, the TND and FS problems assume a frequency-based operation where it is possible to estimate the waiting times, fleet size, and waiting times at transfers stops in terms of the lines' frequencies. However, the timetable-based operation arise from the need of consider specific operation characteristics such as satisfy known demand patterns, synchronize different bus lines, and maximize the number of well timed transfers.

The Transit Network Timetabling problem, denoted as TNT, is the first decision level that *determines the schedules of the lines, i.e. the departure and arrival times for a set of trips along the stops of the transit network in order to optimize a specific objective function*. An example of a timetable solution is given in Table 4, where the pair (a, d) represents the arrival time a and departure time d for each stop in the line. Notice that the headways (separation time between consecutive trips) are different, and thus the transport service is not regular. The relevance of generating a timetable relies on the fact that inadequate and/or inaccurate timetables not only confuse the passengers but also reinforce the bad image of public transit as a whole (Ceder, 2007).

The objectives of TNT involve different criteria such as satisfying headway bounds, satisfying specific demand patterns, maximizing synchronization events of the bus lines, and minimizing other measures such as load unevenness, headways unevenness, travel times, vehicle/driver schedules costs, and/or passenger waiting times (see reviews of Desaulniers and Hickman, 2007; Guihaire and Hao, 2008a).

An important aspect of the planning process is to consider heterogeneous conditions of the system along the day. A common approach to handle the latter issue is to divide the day into smaller planning periods with low variability of demand and travel times. The definition of these planning periods is not straightforward. For example, a method used for generating run-time values consists of the following steps (see Salicrú et al., 2011): purging and screening atypical trips based on the consideration of confidence intervals for median trips; segmenting the day into time bands based on a classification algorithm; creating run-time values based on criteria derived from statistical analysis; and adjusting and validating run-time values using micro simulations. Once the smaller planning periods are defined, deterministic timetabling problems could be implemented.

3.2.1. Transit Network Timetabling to meet specific demand patterns

The efficiency of timetables is strongly related to passenger demand, which may be highly variable during the day and even within short planning periods. Examples of how to create timetables with balanced passenger loads are shown by Ceder (2007), where the rate of passenger arrivals is assumed to be constant (based on data recollection) in small planning periods and departure times are generated through analytic procedures. It is unusual to address load and headway evenness in a timetabling problem. However, Ceder et al. (2013) propose to minimize the deviation from the desired passenger load while trying to maintain even headways using buses with different sizes. A handicap is the determination of the headway to be used in the timetable. Hence, the authors develop a heuristic approach to determining the desired headway to satisfy the

Table 4
Timetable of three trips of a line l that cover 5 stops.

Trips	stop 1	stop 2	stop 3	stop 4	stop 5
1	(7:00, 7:01)	(7:05, 7:06)	(7:11, 7:11)	(7:18, 7:18)	(7:25, 7:26)
2	(7:10, 7:11)	(7:15, 7:16)	(7:21, 7:21)	(7:28, 7:28)	(7:35, 7:36)
3	(7:15, 7:16)	(7:20, 7:21)	(7:26, 7:26)	(7:33, 7:33)	(7:40, 7:41)

demand based on the following three strategies: (i) maximizing the size of the bus; (ii) minimizing the size of the bus; and (iii) selecting the bus whose capacity is closer to the average passenger load. Moreover, the desired headway is used in a heuristic algorithm to determine the timetable. Tests are made with a real case study considering three planning periods of (08:00 to 14:00, 14:00 to 19:00, and 19:00 to 24:00) and buses with capacities of 20, 50, and 75 passengers. Numerical results show that the objectives are indeed in conflict.

Li et al. (2010) address the periodic timetabling problem in a multi-modal system in order to maximize the total user net utility. One of its major contributions is that the activity and travel choices of transit passengers are considered explicitly in terms of departure time choice, activity/trip chain choices, activity duration choice, and mode choice, through a supply–demand equilibrium model. The authors propose a bi-level mathematical programming problem with supply–demand equilibrium constraints and implement heuristic procedures to solve small instances leading to headway values that improve the user's net utility.

Punctuality is important for a transport system since delays of departures lead to losing transfers, and giving the service a bad image. Palma and Lindsey (2001) assume that each passenger has an ideal boarding time and incurs a varying schedule delay cost from traveling earlier or later. Then, the authors develop an analytical model considering two steps: (i) a demand allocation problem is solved to assign individuals to line runs then, (ii) an optimization process is solved to set departure times in such a way as to minimize the passengers' total schedule delay costs. The minimization of the user's inconvenience is also studied by Mesa et al. (2013). Constraints of the system are the limited number of trips in a given planning period, limits to the fleet size, and with respect to the vehicles' capacities. The authors propose a mathematical formulation based on the ρ-median problem, and design a clustering algorithm to reduce the number of requests (origin-destination trips with preferred departure time). This reduction allows the authors solving randomly generated instances (up to 500 requests and four vehicles) using CPLEX's solver.

3.2.2. Transit Network Timetabling to minimize waiting times

Passenger transfers are present in all transport systems. Then, the minimization of waiting time costs at transfer stops may be a reasonable objective for the TNT. For example, Klemt and Stemme (1988) propose a quadratic semi-assignment formulation for the TNT minimizing transfer waiting times in order to schedule a given number of trips. The authors develop a constructive process in which trips are scheduled one at a time with consideration to transfer synchronization. Later, Domschke (1989) designed B&B, local search, and SA algorithms that outperform the solutions obtained by Klemt and Stemme (1988). Chakroborty et al. (1995) present a non-linear mathematical formulation for the TNT minimizing the total waiting time (including transfers). Constraints of the problem are fleet size, arrival times bounds, maximum headway value, and maximum transfer times. The authors develop a Genetic Algorithm (GA) which is tested on examples of the problem obtaining high quality feasible solutions in short time.

Daduna and Voß (1995) propose a timetabling problem to synchronize arrival times at transfer zones minimizing the waiting time incurred. Moreover, the authors consider alternative objectives such as a weighted objective function based on different transfers and the maximum waiting time at a transfer zone. A mathematical model based on quadratic semi-assignment is used to formulate the TNT. Due to its complexity, a heuristic is used to compute initial solutions, which are then improved by a Simulated Annealing algorithm and different versions of Tabu Search (TS).

Bookbinder and Désilets (1992) address a TNT that minimizes transfer costs. The authors assume stochastic travel times and a unique headway value for each line, i.e. the first departure is the only decision variable while the rest of departures time are computed from it. The proposed solution approach combines a simulation procedure with a mathematical formulation like the one presented by Klemt and Stemme (1988). Wong et al. (2008) address a similar objective but considering headway bounds and deterministic travel times. The authors propose an optimization-based heuristic that iteratively solves linear sub-problems and then fixes/releases some variables until a feasible solution is obtained. This solution algorithm is able to solve instances of up to four lines and 16 transfer stops. Even when the latter study is focus on rail systems, the considered characteristics are suitable for general transit networks.

Castelli et al. (2004) address a timetabling problem to minimize the weighted sum of passenger costs (based on time spent in the system) and operational costs (based on vehicles usage). The authors define a mathematical formulation including transfer coordination decisions. Since the proposed model has a large number of decision variables, the authors develop a Lagrangian relaxation method to solve generated instances considering a planning period of one hour.

Schröder and Solchenbach (2006) state that minimizing transfer waiting times may lead to risky passenger transfers due to the usual delays in bus arrivals at stops. Then, they define a TNT considering the following types of transfers and their costs based on the waiting time incurred: (i) impossible transfer; (ii) almost transfer; (iii) risky transfer; (iv) acceptable transfer; and (v) patience transfer. Then, a mathematical formulation to shift trips (of a given timetable) is defined in order to minimize the total transfers cost. A few instances are solved using a commercial solver, and numerical results show that trip shifting leads to significant improvements in transfer coordination.

Guihaire and Hao (2010b) address the TNT with an objective function based on quantity and quality of transfers, evenness of line headways, feet size and length of the deadheads. The authors consider a constraint that bounds the deviation from an initial timetable. This constraint allows the authors to develop a preprocessing stage to define feasible intervals for departure times. Then, the information of the preprocessing stage is used to implement shifting movements embedded in a Tabu Search algorithm to solve generated instances of the problem.

As analyzed in Zhao and Zeng (2008), uncertainty of travel times influence the chances of achieving well-timed passenger transfers. For this reason, Liebchen and Stiller (2012) implement buffer times to generate delay resistant timetables. The authors define an optimization problem to minimize the expected total weighted delay for planned trips considering headway bounds and fleet size constraints. To solve the problem, the authors develop a heuristic algorithm which is capable of obtaining optimal solutions for problems that can be modeled through series-parallel graphs. Although this is implemented in periodic timetables for railway systems, the main idea of implementing buffer times can also be used for general transit networks with aperiodic timetables.

3.2.3. Transit Network Timetabling to maximize the number of synchronization events

The synchronization of different lines can be used to bus congestion at common stops or allow passenger transfers. For example, Ceder and Tal (2001) and Ceder et al. (2001), address a TNT focused on maximizing the number of pairwise simultaneous arrivals at common nodes in order to benefit passenger transfers. The problem consider headway bounds and an initial frequency. Since the problem is intractable by commercial solvers, the authors design a heuristic procedure to generate timetables for some examples. Later, Liu et al. (2007) redefined the synchronization measure of Ceder and Tal (2001) as the ratio of the number of lines where there are vehicles arriving simultaneously at a connection stop to the number of all lines passing through the same stop. Then, a nesting Tabu Search is designed to obtain feasible solutions for small instances (up to eight bus lines and three synchronization nodes).

More recently, Ibarra-Rojas and Rios-Solis (2012) redefine a synchronization event as the arrival of two trips belonging to different lines at a specific stop such that the separation time between these arrivals is within a specific time window. These synchronization events are used to benefit passenger transfers and to reduce bus congestion at common stops. The authors consider headway bounds and specific departure time for the first/last departure of each line in order to guarantee an almost regular service over the entire planning period. The problem is modeled with a Mixed integer linear formulation that is intractable by commercial solvers, even for small instances. Besides, the NP-hardness of the timetabling problem is proved. Then, an Iterated Local Search Algorithm (ILS) is designed to solve large instances of the problem.

As we mentioned before, the division of a day into smaller planning periods based on demand and travel times variability is not straightforward (see Salicrú et al., 2011). Moreover, the divisions are not necessarily the same for different lines. On the basis of this fact, Ibarra-Rojas et al. (2015) extend the approach of Ibarra-Rojas and Rios-Solis (2012) to consider multiple planning periods. The authors define a synchronization event between trips belonging to different periods as the pair wise arrival at synchronization nodes such that the separation between arrivals is between minimum and maximum that are function of the headways of the related trips. Then, they define the optimization problem to determine a timetable for the entire day maximizing the number of synchronization events. The size of the problem increases significantly when an entire day is considered compared with a single-period case. Thus, the authors develop metaheuristics such as multi-start ILS and multi-start Variable Neighborhood Search (VNS) to solve the proposed problem. Numerical results of implementing the metaheuristics on generated instances show that, even when optimality is not guaranteed, the system synchronization obtained with the proposed multi-period approach significantly improved the system synchronization obtained by merging the timetables of a single period case.

3.2.4. Multi-objective Transit Network Timetabling

As we saw in the previous sections, there are different criteria for timetabling problems and it is possible that some of those criteria are in conflict. Then, multi-objective optimization approaches can be used to properly represent the trade-offs between different objective functions.

Kwan and Chang (2008) present a bi-objective formulation for the TNT: minimizing the cost of the number of transfers and minimizing the cost caused by deviations from an initial timetabling. The authors implement the Non-dominated Sorting Genetic Algorithm II (NSGA II) (Deb et al., 2002) and multi-objective EA, and local search procedures to solve the problem formulation. Hassold and Ceder (2012), study the TNT problem with the objective of minimizing the expected passenger waiting time (a level of service feature) and the empty seat penalty (operational efficiency). The main idea of the study is that the disadvantages of the even-headway and even-load approaches for timetabling can be avoided by using more than one vehicle size on the same line. The authors design a network flow bi-objective formulation assuming known potential departure times, and determine the vehicle type to cover all trips. They implement a multi-objective label-correcting algorithm to solve the problem. Numerical results for a case of study in New Zealand show significantly savings in passenger waiting times but also acceptable passenger loads on all vehicles.

Since it is difficult to compare efficient solutions for multi-objective optimization problems, an alternative way to model the preferences of the decision maker is to implement fuzzy approaches (see an introduction to fuzzy optimization in Lodwick and Untiedt (2010)). For example, Tilahun and Ong (2012) minimize the waiting time of different transfers which are in conflict. The TNT is formulated as a fuzzy multi-objective problem considering a set of lines with a unique daily trip and deterministic travel times. The authors implement a GA to solve small instances consisting of 10 intersecting lines.

Table 5 summarizes our literature review and shows the details of the studies presented in this section.

Table 5
Literature review of the Transit Network Timetabling problem.

Authors (year)	Objective	Constraints of the system	Solution method	Case
Klemt and Stemme (1988)	Min transfer waiting times	Number of trips	Heuristic	Example
Domschke (1989)	Min transfer waiting times	Number of trips	SA & B&B	Example
Bookbinder and Désilets (1992)	Min waiting times	Even headways	Simulation and mathematical programming	Example
Daduna and Voß (1995)	Min transfer waiting times	Fulfill demand	TS	Examples
Chakroborty et al. (1995, 1997)	Min passengers total waiting time	Fleet size, stopping times bounds, and maximum headways and travel times	GA	Example
Ceder et al. (2001)	Max number of simultaneous arrivals	Headway bounds	Heuristic	Example
Ceder and Tal (2001)				
Palma and Lindsey (2001)	Min total delay costs	Desired boarding times	Analytical	None
Ceder (2003)	Demand and number of trips	Headway bounds	Heuristic	Example
Castelli et al. (2004)	Min passenger transfer waiting time	Frequency bounds	Lagrangian heuristic	Test
Schröder and Solchenbach (2006)	Min waiting times costs	Deviation from a timetable	CPLEX	Real
Ceder (2007)	Min bus runs	Load profiles and vehicle capacity	Heuristic	Example
Liu et al. (2007)	Min transfer time	Headway bounds	Nesting TS	Example
Wong et al. (2008)	Min waiting time costs	Headway bounds	Lagrangian heuristic	Real
Zhao and Zeng (2008)	Min travel time costs	Headways, fleet size, line length, and load factor	SA & TS	Benchmark
Kwan and Chang (2008)	Max syncs vs. min deviation from an initial timetable	Frequency, limited load and drivers, and bounded run time	NSGA2, TS, & SA	Example
Guihaire and Hao (2010a)	Max number and quality of transfers	Vehicle and driver schedule	TS	Real
Li et al. (2010)	Min cost of fulfilling demand considering uneven headways	Fleet size	Hooke-Jeeves	Example
Guihaire and Hao (2010b)	Min waiting time costs	Vehicle and driver schedules	TS	Test
Hassold and Ceder (2012)	Min waiting times vs. min empty seat penalty	Potential departure times	Heuristic	Test
Ibarra-Rojas and Rios-Solis (2012)	Max number of synchronizations	Headway bounds, departure dispersion along the planning period	ILS	Test
Liebchen and Stiller (2012)	Min expected waiting delay	Budget of buffer times	Heuristic	Example
Tilahun and Ong (2012)	Multi-objective min of waiting times for different transfers	Arrivals precedence	GA	Example
Mesa et al. (2013)	Min user inconvenience	Number of trips and fleet size	Cluster pre-processing& CPLEX	Test
Ceder et al. (2013)	Min load discrepancy vs. Min headway unevenness	Deviation from given headway and demand satisfaction	Heuristics	Real
Ibarra-Rojas et al. (2015)	Max synchronization events	Headway bounds for multiple planning	ILS and VNS	Test

3.3. Design of operational strategies to improve the efficiency in a corridor

Unfortunately, a priori frequency calculation is not enough to guarantee providing good service and meeting the demand in high-demand corridors due to different demand patterns and the unbalanced distribution of the passengers along the corridor. In these cases, the following operational strategies are implemented (see Furth and Day, 1985).

- Short-turn lines: some buses serving a line make shorter cycles in order to concentrate on areas of greater demand. This is useful when it is desired to bolster capacity along a given stretch of the line.
- Dead-heading: empty vehicles return to the line starting point in the low-demand direction in order to begin another run as quickly as possible in the high-demand direction, thus increasing the latter direction's frequencies.
- Limited-stop lines: lines that visit only a subset of the stops on a line where the passenger demand is concentrated. These lines allow diminishing travel times, but passenger transfers must be considered.

The goal of using the previous strategies is to provide an optimal frequency using spatial-time considerations, i.e. to adjust the frequency along the corridor and over the day based on the demand behavior. For example, peak hours may lead to a large amount of people traveling from zones near the beginning of the lines to the Central Business District (CDB). However, choosing these "zones" depends on many aspects, such as passenger demand, travel times, and stop separation, among others. The major drawback to this approach is that the perceived frequency at different stops of the same corridor will be different and some passengers may experiment discomfort if they are not well informed about the special lines.

3.3.1. Short-turning and deadheading strategies

Early studies of short-turn lines, such as Jordan and Turnquist (1979), Furth (1986), focus on designing lines for specific zones, where the main decisions are the number of zones, the first and last stops of each zone, and the number of buses serving each zone. Characteristics of the system are stops outside the zone that can be served for alighting and boarding, asymmetric boundaries that depend on the direction of the trip during any given period, and branching corridors. The problems are solved through dynamic programming. Later, some studies have aimed at fleet size minimization through determining the frequency of short-turning lines and deadheading interlining options for trips using heuristic algorithms (see Furth, 1987; Ceder, 1989).

More recently, Delle Site and Filippi (1998) address the problem of determining the stops for short-turning lines, the fare (in terms of the traveled length), the vehicle size, and the frequency for each type of line, considering an initial OD matrix. The objective is minimizing the sum of two types of costs: (i) passengers' costs based on waiting times, walking distances, transfers, and fares; and (ii) operators' costs based on fleet size, distance traveled by the buses, fixed costs, and drivers costs. The authors design a solution methodology based on solving sub-problems to find the frequency and fares for a feasible combination of stops where turn-backs can occur. This solving stage is implemented within a global optimization algorithm based on random searches. The proposed approach is tested on generated instances and numerical results show that short-turn lines are useful for minimizing the costs of both passengers and the operator, especially in peak periods.

Cortés et al. (2011) address the integration of short turning and dead-heading strategies. The goal is defining lines to satisfy a demand profile in a more efficient way minimizing operators' and users' costs. An interesting characteristic is that the authors also allow trips to arrive at stops using a probabilistic arrival rate. Thus, the assumption of deterministic travel times could be avoided. The proposed model considers stop-to-stop demand and generalizes the formula that define the optimal rate at which vehicles are dispatched, proportionally to the square root of the passenger arrival rate (see Welding, 1964; Newell, 1971; Salzborn, 1972).

3.3.2. Limited-stop lines

From the users' perspective, limited-stop lines improve the level of service by reducing travel times due to fewer stops. From the operators' perspective, it makes more likely to meet the demand using fewer vehicles by performing shorter bus cycles. These lines have proven to be attractive in transit systems in Bogota, Colombia, Santiago, Chile, New York, U.S., Montreal, Canada, Los Angeles, CA. However, selecting the proper stops for limited-stop lines is a challenging research area, where several aspects should be considered. For example, passengers' changing from normal to limited-stop lines, waiting times along the line, estimation of the running times, frequency regularization, and stop separation.

Tétreault and El-Geneidy (2010) propose a model to estimate running times considering limited-stop lines. This model isolates the effects of passenger activity and actual stops made by the current line at skipped stops in order to estimate the run time for the limited-stop line and the normal line. Besides, it considers direction of the line, stops covered, time of day, type of day, delay at the first stop, passenger activity, and climatic conditions. The model is tested on a corridor in the transit network of Montreal considering three different scenarios: (i) a normal line and a limited-stop line that covers only transfer stops; (ii) a normal line and a limited-stop line that covers the stops in the first quartile of passenger activity; and (iii) a normal line and a limited-stop line that covers stops with the most activity based on OD data. The authors analyze all scenarios measuring run time savings and walking distances to bus stops. Numerical results show that significant running time savings are obtained in all scenarios. The accuracy of the previous study is evaluated by El-Geneidy and Surprenant-Legault (2010), where a before-and-after approach is used to measure the actual changes in running times and on-time performance along two lines of the transit system in Montreal, Canada.

Even when significant benefits can be achieved by implementing limited-stop lines, only a few studies address the design of such lines. An example is presented by Afanasiev and Liberman (1982). The authors develop a method that provides a chain of linked bus runs. The objective is to fulfill the passenger demand and the lines can be of three types: all stops; every other stop (limited-stop); direct trip from origin to destination (express line). The construction of the chains of limited-stop lines is based on a model of the passenger demand pattern that determines the beginning and the end of the line along with its frequency. Tests on a transit network in the city of Chimkent show that travel times decrease on average 30% and gasoline consumption decreases between 30% and 40% compared with no implementation of the proposed lines.

Ulusoy et al. (2010) address the design of short-turn and limited-stop lines, along with their Frequency Setting, to minimize the total costs based on waiting times, transfers, and vehicle usage. The problem formulation considers constraints such as frequency bounds, vehicles capacity, and fleet size. The complexity of the problem lies in the combinatorial nature of the set of potential patterns that should be considered in order to define the limited-stop lines, i.e. more potential line patterns should be explored as the number of stops increases. The authors design a heuristic search procedure that modifies the pattern of limited-stop lines after these patterns are evaluated by solving a mathematical formulation that minimizes the total costs. The proposed approach allows reducing the fleet size, vehicle-miles traveled, and total costs. Although the case study is a railway transit system in the northeastern region of the United States, the problem definition can be adapted to bus lines.

Leiva et al. (2010) propose an optimization problem which choose limited-stop lines and determine their frequencies, minimizing a cost function based on wait time, travel time, and operators' costs. This methodology assumes that the origin–destination matrix and the set of feasible limited-stop lines are given. The authors study the following scenarios: (i) no capacity constraints and considering transfers; (ii) considering both capacity constraints and transfers; (iii) considering different vehicle types. Numerical results show that value of the cost function decreased more than 10% for all scenarios compared with no implementation of limited-stop lines. Moreover, the authors stated that the grater the trip length and demand variability (particularly when the demand is concentrated in a few OD pairs), the higher the benefits of implementing limited-stop lines. The impact of the different parameters of the previous approach is presented by Larrain et al. (2010).

Recently, Freyss et al. (2013) study a system based on three types of metro stations: AB stations where all the trains stop, A stations where only one-half of the trains stop, and type B stations where the other one-half of the trains stop. The authors implement a continuous approximation model to determine the optimal AB station density for the lines (stations/m) based on the local conditions (regular station density and affluence) and to compare the total cost (including in-vehicle travel time, waiting time, transferring costs, and operational costs) involved with both stop-skipping operation and with standard service. This operational strategy is quite suitable for transit systems without overtaking lanes, as many BRT corridors worldwide. The strategy is tested on the metro system in Santiago, Chile, and a numerical analysis leads to identifying characteristics less favorable for stop-skipping operations, such as short lines with few stations, low initial station density, and low/medium frequency lines.

Chiraphadhanakul and Barnhart (2013) assume that only one limited-stop line may run in parallel with the normal line. Moreover, the passenger assignment is a linear function of the frequency share and in-vehicle travel time savings. For each frequency allocation, the approach finds an optimal limited-stop line using a mixed integer program together with an algorithm for reducing the size of the problem. Then, it selects the frequency allocation that maximizes the total user welfare which is defined as total in-vehicle time savings minus total increase in waiting time. The proposed approach is tested on instances considering the peak period of 7 a.m. to 9 a.m. for 178 high-frequency bus lines. Numerical results show that line frequency and average trip length are highly correlated with the total user welfare attained by the optimal solutions.

As we have seen in this section, the design of different types of lines such as short-turning and limited-stop lines is a promising and mostly uncovered research area.

4. Operational planning decisions

As we mentioned in Section 1, operational planning span short-term decisions focus on minimizing the cost related with the usage of vehicles, fuel consumption, and drivers' wages. It is divided into three problems: (i) Vehicle Scheduling Problem (VSP), (ii) Driver Scheduling Problem (DSP), and (iii) Driver Rostering Problem (DRP). The usual approach is to consider strategical and tactical planning decisions as input and then, determine a better way of using the agencies' resources in order to reduce costs.

4.1. Vehicle Scheduling Problem

In the cases of frequency-based operation, it is possible to estimate the fleet size as the smallest following integer to the product between the cycle time and the frequency:

$$\text{Fleetsize} = \sum_{l \in L} \lceil t^l \cdot f^l \rceil \tag{2}$$

Parameters t^r and f^r are, respectively, the cycle time and the frequency for line $l \in L$. A characteristic that influences the fleet size in frequency-based systems is the type of vehicles used. This is because determining the proper frequency to satisfy the passenger demand depends on vehicles capacities. For example, Wei et al. (2013) address a bi-level optimization approach to jointly solve the Vehicle Scheduling Problem and the vehicles procurement scheme. The upper-level determines the number of vehicles to be purchased of each type satisfying passenger demand and pollutant regulations while the lower-level determines the vehicle schedules minimizing the cost of vehicle fees and mileage.

In the case of timetable-based operation, Formula (2) lacks of sense since the vehicle schedule will be defined based on departure times, arrival times, and travel times. This leads to the Vehicle Scheduling Problem (denoted as VSP) that *determines the trips that are assigned to each bus so that all trips can be carried out as scheduled and costs based on vehicle usage are minimized*. Important elements to define the VSP are the following.

- Number of depots which vehicles may depart from. If there are multiple depots, it is important to define whether a vehicle must return to the same depot from which it departed, or whether vehicles can share all depots.
- Number of different fleets with different capacities and operating costs.
- Resting points where vehicles may remain until the next trip (e.g. a street with low vehicles flow, a parking lot, and so on). A depot may be used as a resting point but a resting point does not necessarily require special infrastructure.
- Operating conditions such as the following: (i) inter-routing, i.e. a vehicle could be assigned to a trip of one line and then, be assigned to a trip of a different line; (ii) deadhead implementation, i.e. empty trips traveling from one point to another increasing the availability of vehicles at specific points where they are more required. Allowing these conditions may prevent under-utilization of a fleet.

The more flexible the system is, the greater the potential benefits but also the greater the difficulty of exploring the solution space. Now, we review different approaches for Vehicle Scheduling Problems.

4.1.1. Single-Depot Vehicle Scheduling Problem

The simplest version of the VSP is known as the Single-Depot Single-Type VSP (denoted as SDVSP). This problem determines the vehicles' schedules so that all trips are covered by vehicles departing (returning) from (to) the same location, called the depot. A vehicle schedule is composed of vehicle blocks, where each block represents the departure from the depot to serve a sequence of trips and then return to the depot. The SDVSP admits several types of representations, such as assignment models, transportation models, and network flow models (see reviews in Freling et al., 2001b; Bunte and Kliewer, 2009). Moreover, the SDVSP can be efficiently solved by auction algorithms (see auction algorithm details in Bertsekas, 1992).

Although the SDVSP can be easily solved, practical considerations may lead to more intractable problems. For example, Baita et al. (2000) present several approaches to solve the VSP considering the following practical elements: it is possible to perform deadheading; it is possible to fuel a vehicle after finishing a trip; and it is possible to hold a bus at the end of the trip to wait for the starting time of the following assigned trip. Moreover, the authors considered criteria such as minimizing the fleet size, minimizing the number of lines that a bus is assigned to, minimizing the deadhead costs, and minimizing the idle times of buses waiting for the next trip. The authors provided a pool of solutions for operators and compare three solution approaches: (i) solving a mathematical formulation with a solver considering a weighted objective function; (ii) implementing a logic programming method considering a lexicographic order for the different criteria; and (iii) implementing a multi-objective GA to find Pareto optimal solutions.

The proper division of an operational day into smaller planning periods based on demand behavior is an important aspect in the implementation of deterministic optimization approaches. In this matter, Zhoucong et al. (2013) propose a clustering-based method to generate short planning periods with low variability of travel times. Then, the SDVSP is solved considering the partition of the day made by the clustering approach. The authors implement the proposed methodology in a transit network in Shanghai. Numerical result show that a proper division of the day leads to a more accurate representation of vehicles' usage costs.

4.1.2. Multi-Depot Vehicle Scheduling Problem

Another version of the VSP is the Multi-Depot Vehicle Scheduling Problem (MDVSP), where vehicles can depart from different locations. This assumption leads to complex formulations, such as multi-commodity network flow problems. The MDVSP is intractable since it is NP-hard (proved by Bertossi et al., 1987).

Exact approaches towards solving large instances of MDVSP have been presented in recent studies. For example, Löbel (1998) address the MDVSP minimizing the costs of vehicle pullouts from depots subject to constraints such as fleet size, depot capacity, and vehicle-depot compatibility. The authors develop a Column Generation approach (CG) based on Lagrangian relaxations of a multi-commodity network flow formulation. The proposed solution approach is tested on large instances of the problem based on a German transit network. Numerical results suggests that Lagrangian relaxation is a key element to solve large instances, and that the combination of CG with heuristics exploiting the solutions of the solved restricted LPs leads to high quality feasible solutions within few iterations and few hours of computational time.

Kliewer et al. (2006) develop a time-space formulation that significantly reduces the size of the network which allow solving large instances using commercial solvers. Hadjar et al. (2006) develop a Branch & Bound (B&B) approach that combines CG, variable fixing, and Branch & Cut (B&C). By implementing the proposed solution approaches, the authors succeed in solving randomly generated instances up to six depots and 750 trips. Other authors propose heuristic methods, such as schedule-first/cluster-second approaches, shortest-path-based heuristics, and ILS algorithms (see details of ILS in Daduna and Paixão, 1995; Laurent and Hao, 2009). Pepin et al. (2009) compare B&C, CG, Lagrangian Heuristic, Tabu Search, and Large Neighborhood Search (LNS). Computational results on randomly generated instances show that the truncated Column Generation method performs the best when enough computational time is available while the Large Neighborhood Search method is the best alternative when looking for good quality with the MDVSP case, the consideration of different vehicle types also leads to intractable problems where the usual solution approaches are heuristic procedures (Gintner et al., 2005).

Route time constraints for vehicles' schedules have been also considered for MDVSP. For example, Haghani and Banihashemi (2002) bound the length of the blocks of vehicle schedules in order to represent vehicles' fuel consumption. Thus, every bus may run at most t consecutive units of time before re-fuel and then, the vehicle is unavailable for another t' units of time (t' is considered to be a known parameter). The authors enhance the quasi-assignment formulation to allow buses returning to the depot so they can be re-fueled. The proposed solution approaches are exact and heuristic methods based on two procedures: (i) Solving iteratively the classical MDVSP quassi-assignment formulation, removing infeasibilities of running time constraints at each iteration, and (ii) reducing the search space by merging vehicle blocks and removing variables with a slim probability of being chosen. Another example is shown by Wang and Shen (2007) where every vehicle's running time after fueling is limited. Thus, the vehicle must be re-fueled after taking several trips and the minimal fueling time must be satisfied. The optimization problem aims to minimize a hierarchical objective function consisting of two levels:

(i) the number of vehicles and (ii) the total deadhead time. To solve the proposed problem, the authors develop an ACO algorithm which is tested on several examples.

More recently, Hassold and Ceder (2014) present a multilayer network flow formulation for the MDVSP minimizing vehicle costs based on vehicles' usage and empty seats. A set of Pareto-optimal timetables, including recommended type-of-vehicle for each trip, is given as an input. Considering these potential timetables, the problem must build a unique timetable minimizing operators' costs. The authors analyze different levels of operation's flexibility such as: (i) allow substitution of vehicles, i.e., trips may be executed by different vehicle types (with enough capacity) than what was initially assigned in the timetable; and (ii) using partial timetables to build a unique timetable. The authors test their model on two bi-directional lines of the transit network in Auckland, New Zealand yielding a reduction up to 15% of operating costs compared with the classical MDVSP. Moreover, they analyze the impact of the solution on the level of service, in terms of waiting times and empty seats.

4.1.3. Robust Vehicle Scheduling Problem

Vehicles schedules are strongly affected by congestion and drivers' behavior. Then, no consideration of future disruptions may lead to a risky operation and to penalty fees which have to be paid by transportation agencies if punctuality falls below a contracted minimal level. An option to tackle this handicap is to define robust schedules which obtain solutions resistant to uncertainty of parameters. For example, Kramkowski et al. (2009), implement buffer times in the SDVSP to get robust solutions for the problem minimizing the average delay of the planned timetable. To solve the problem, the authors develop an SA for noisy environments with different neighborhood operators. Numerical results show that the use of minimal buffer times before every timetabled trip decreases the expected length of propagated delays (caused by a trip delay and the planned schedule) to a great extent, while introducing a larger buffer times (for example, five minutes), results in almost zero propagated delays but the planned cost also increases.

Naumann et al. (2011) present an optimization problem to reduce the sum of planned and disruption costs caused by delays in the original schedule. The authors present a time–space formulation based on Kliewer et al. (2006), where arcs are added to allow buses to wait a certain amount of time before running the next scheduled trip, thus making a more robust schedule. Obviously, diminishing the disruption costs may increase the planned costs. Then, the authors obtain Pareto optimal solutions which are compared with the optimal solutions from the deterministic approach and the implementation of buffer times to diminish delay disruptions. The proposed approach is used to solve small size (up to hundreds of trips) real instances.

Wei et al. (2012) address a robust MDVSP minimizing the operating costs including mileage expenses for their travel. The delay of a trip caused by uncertain travel times is modeled as a random variable within a specific interval and a known estimated value. Some constraints of the systems include: they limit the use of backup vehicles; they must guarantee that a vehicle has adequate fuel to cover a series of trips between two adjacent fueling activities; and they require compatible line-vehicle pairs. To solve several examples of the proposed problem, the authors implemented an ACO.

In order to achieve a punctual service, Yan et al. (2012) propose a VSP minimizing the deviation from an initial schedule at specific control points. The travel times are represented as stochastic variables with a known probability distribution and the behavior of drivers is modeled in order to meet punctuality. To achieve this, the authors assume that once a bus in operation deviates from its schedule at a time control point, the bus driver will adjust the bus speed-by either speeding up or slowing down on the next segment-in order to ensure schedule adherence at the next time control point. Moreover, the authors propose the concept of remaining impact after recovery, in order to describe the schedule deviation that cannot be recovered after the driver's adjustment. These characteristics are modeled in a Monte Carlo simulation to determine the proper bus scheduling that guarantees adherence to the timetable based on minimizing penalties for early/late arrivals at control points.

4.1.4. Dynamic Vehicle Scheduling Problem and Vehicle Re-Scheduling Problem

Other alternatives to tackle parameter uncertainty in VSP problems are: (i) the Dynamic Vehicle Scheduling Problem (DVSP), which assumes that it is possible to obtain accurate estimations in short planning periods and then, *re-defines vehicle-trip assignments in a rolling time horizon preventing trip delays*; and (ii) the Vehicle Re-Scheduling Problem (VRSP), which *re-defines the trip-vehicle assignment when a disruption (such as a vehicle breakdown, a traffic accident, a medical emergency, and road work) occurs, in order to minimize the impact on the original schedule*.

Huisman et al. (2004) address a robust optimization approach for the Dynamic version of the MDVSP considering different scenarios for estimation of travel times. The authors implement a cluster-reschedule heuristic, where trips are clustered first and they are assigned to different depots using the static Vehicle Scheduling Problem, and then, dynamically reschedule the trips per depot. The proposed heuristic is tested on real-life instances and numerical results show that the gap between this cluster-reschedule heuristic and a lower bound on the overall problem is less than 4%, which is reasonably small.

The main goal of VRSP is to backup canceled trips to avoid losing passengers but, there may be consequences, such as the use of extra vehicles or drivers. Thus, the optimization approach consists in minimizing the costs arising when a disruption occurs or the timetable is changed. Although the previous problem definition may lead to important practical implementations, there have been only a few studies addressing VRSP (see Li et al. (2007, 2009)). For example, Li et al. (2007) present mathematical formulations for the single-period case and they develop a parallel auction algorithm to solve the proposed formulation. Using this algorithm, they are able to find, in minutes, optimal solutions for instances of up to 1300 remaining trips and 40 backup trips. Another study of the VRSP is presented by Li et al. (2009), where the authors define a mathematical formulation for the single-depot case. Although the original VSP can be solved efficiently through the auction algorithm, the authors in this study show that their VRSP is indeed NP-hard. Then, they use a Lagrangian relaxation based heuristic to find feasible solutions for instances of up to 700 trips.

We summarize our review of Vehicle Scheduling Problems in Table 6. Another review of VRPS including bus transport systems is presented by Visentini et al. (2013).

4.2. Driver Scheduling Problem

The Driver Scheduling Problem (denoted as DSP), also called as duty scheduling, *defines the generic daily duties to cover a set of vehicle blocks-consisting of timetabled trips-in order to minimize the cost of duties and satisfy labor regulations constraints*. Thus, a solution of this problem determines starting and ending time of the work load for each driver. Common constraints for DSP are the following (see a review of the DSP in Wren and Rousseau, 1995): limited duration of the work load; meal breaks for drivers; limited amount of continuous work without break; and starting times for duties. The DSP considers different objective functions such as minimizing the number of duties, minimizing idle times, minimizing the penalty of constraints violation, minimize the uncovered duties, and minimizing cost of potential disruptions (robustness) among others.The DSP is one of the last problems in TNP to be solved by sequential approaches thus, the feasible space of the DSP is often restricted by solutions of the previous sub-problems of the TNP. This characteristic makes difficult even finding feasible solutions for the DSP that satisfy all of its constraints.

Common formulations for the DSP are based on set partitioning and set covering problems. In these kind of formulations, the model must be fed with all possible duties. The chosen duties must cover all trips at the minimum cost. Due to the large number of variables this entails, these formulations are intractable by exhaustive enumeration techniques. Some approaches to solve these formulations are reducing the number of duties being considered (e.g., Smith and Wren, 1988), using Column Generation techniques (e.g., Desrochers and Soumis, 1989), and obtaining dual bounds through heuristics that are then fed into a B&B algorithm (e.g., Mingozzi et al. (1999)).

4.2.1. Column Generation approaches for the Driver Scheduling Problem

A common handicap for CG approaches for the DSP is the need of solving sub-problems such as the resource-constrained shortest path problem (RCSPP) which determines the columns that should be added to the problem in further iterations of

Table 6
Literature review for the Vehicle Scheduling Problem.

Authors (year)	Objective	Constraints of the system	Solution method	Case
Löbel (1998)	Min fleet size costs	Fleet size, depot capacity, and vehicle-depot compatibility	CG	Test
Baita et al. (2000)	Min the fleet size, number of lines assigned to each bus, deadhead costs, and idle times	Fueling, bus-line compatibility,	Mathematical programing & GA	Examples
Haghani and Banihashemi (2002)	Vehicle operation costs based on parking, run, and wait in street	line time, trip coverage	Heuristic	Test
Huisman et al. (2004)	Min vehicle costs	Single vehicle-depot assignment and vehicle-depot compatibility	Cluster reschedule heuristic	Real
Gintner et al. (2005)	Min vehicle costs	Trip coverage	Variable fixing heuristic	Test
Hadjar et al. (2006)	Min vehicle costs	Fleet size	CG, variable fixing, and cutting planes	Test
Kliewer et al. (2006)	Min vehicle costs	Trip coverage	Reformulation & solver	Test
Li et al. (2007)	Min delay cost	State of disrupted system	Parallel auction algorithm	Test
Wang and Shen (2007)	Fleet size and deadhead costs	Fleet size and begin and ending at the same depot	ACO	Examples
Kramkowski et al. (2009)	Robust min of delay penalties	State of disrupted system	SA	Real
Laurent and Hao (2009)	Min cost based on vehicle usage	Trip coverage and depot capacity	ILS	Benchmark
Li et al. (2009)	Min weighted sum of trip cancellation, operating, and schedule disruption costs	State of disrupted system	Lagrangian Heuristic	Test
Naumann et al. (2011)	Min planned costs vs. Min disruption costs	Trip coverage	Solver	Test
Wei et al. (2012)	Min operating costs	Limited backup fleet, feeling, and line-vehicle compatibility	ACO	Examples
Yan et al. (2012)	Min the deviation from initial schedule	Bounded slack time	Simulation	Examples
Wei et al. (2013)	Min number of vehicles to be purchased	Depot capacity, passenger demand, fueling, and pollution regulations	GA	Examples
Zhoucong et al. (2013)	Min fleet size and operational costs	Trip coverage, and even headway	Formula	Examples
Hassold and Ceder (2014)	Min vehicle usage costs	Potential timetables	Solver	Real

the main algorithm. For this reason, some studies have developed methods to reduce the calls to solve the RCSPP. For example, Fores et al. (2002) propose creating a set of valid shifts considering labor regulation constraints, and Then, the CG is applied within a Branch & Bound framework to select the best set of candidates from the set of potential shifts. Unfortunately, the proposed approach does not guarantee finding an optimal solution. Later, Chen and Shen (2013) defined a set of potential efficient shifts (i.e. columns) through an algorithm that takes advantage of characteristics of the problem. Then, the RCSPP is solved until no columns with reductions of cost exist in the shift pool. This idea allows of reducing the computational time of the CG to solve the DSP. The experimentation stage considers instances of up to 701 trips, 55 vehicle blocks, and two depots. Numerical results show that the improved CG algorithm outperforms the traditional CG.

Boschetti et al. (2004) extend the method used by Mingozzi et al. (1999) for the DSP but focus on the multi-depot case of the DSP. By implementing Lagrangian Relaxation (LR) and CG, the authors are able to handle the constraints and generate bounding procedures, leading to a more efficient implementation of the proposed CG approach. Although there are several formulations in the literature, Portugal et al. (2009) states that the classical set partition/set covering model simplifies some of the specific business aspects and issues of real problems, making it difficult to use these models as automatic planning systems because the schedules obtained must be modified manually in order to be implemented in real situations. Then, the authors use the actual observations of a transport network agency in Portugal to develop several mathematical formulations with evaluation criteria such as total "real" cost, not covered work by number of pieces, and not covered work duration of pieces. The authors conclude that decision makers prefer to have partial solutions which are easier to adjust instead of having complete solutions with too many duties and multiple trip-duty assignments.

The homogeneity of drivers' schedules, i.e. similar schedules for all days for each driver, is considered by Steinzen et al. (2009). A set covering formulation is defined and the authors develop a solution approach consisting of two stages: (i) solving the linear relaxation of the formulation, through CG and LR; and (ii) implementing local branching procedures to obtain integer solutions which keep the optimal drivers' costs but improve the regularity of the drivers' schedules. The proposed approach is tested on real-world and randomly generated instances. Numerical results show that homogeneity of drivers' schedules could be improved allowing low increments of operating costs.

4.2.2. Heuristics for the Driver Scheduling Problem

Since large instances of the DSP are intractable by commercial solvers and exact approaches may lead to unacceptable computational time, some authors propose implementing heuristic algorithms (e.g., Kwan et al., 1999, Shen and Kwan, 2001). For example, Zhao (2006) divide the entire day into "morning" and "afternoon" problems. Each problem is solved separately and then combined to obtain a solution for the entire day. The proposed algorithms are based on identifying critical relief opportunities that must be retained in the bus schedule, in order to reduce the number of pieces of drivers' duties. Kecskeméti and Bilics (2013) compares three solution methodologies: a CG, an Evolutionary Algorithm (EA), and a combined CG-EA (hybrid). Numerical results show that although CG yields the best results in terms of the solutions' quality, the EA and the hybrid algorithm outperforms the CG approach in terms of running time. Although, the hybrid algorithm takes more execution time compared with the EA, the running time is reasonable and overcomes the quality of the EA.

Chen and Niu (2012) assume that drivers are compatible with the duties belonging to specific periods of the day. For example, a driver might be assigned only to morning trips if that driver lives too far away from the depot. The authors formulate the problem as a 0–1 integer program with the objective of minimizing the cost of the idle time of drivers for a circular bus line. Then, they develop a Tabu Search algorithm which is able to solve instances of up to 168 trips and 26 drivers. Chen et al. (2013) address the DSP where meal breaks for drivers should be within a specific time window of the day. Since "meal break" constraints are hard constraints, the number of feasible duties is significantly smaller than in the case where meal breaks could be assigned at any time of the day. Therefore, the authors implement heuristic procedures based on exhaustive search among key relief opportunities. Results shows that it is possible to find feasible solutions for instances of up to 100,000 potential shifts using the proposed approach.

Tóth and Krész (2013) states that it could be necessary to obtain several solutions based on different parameter values in order to compare these solutions and select the most suitable one for the transit network operation. To achieve this, the authors propose a two step greedy solution approach. The first step produces rough shifts, containing only the trips (both timetabled and deadhead trips) and the traveling activities of the driver. The second phase produces the complete shifts, containing all the obligatory activities, and fills the idle times with the proper idle activities. Experimentation in a transit network in Hungary shows that the proposed solution approach is able to obtain feasible solutions for instances of up to 40 lines, 120 buses, and 160 drivers.

Yen and Birge (2006) combine robust and recovery frameworks to tackle the DSP in airline transportation. The objective is minimizing the expected value of future actions due to disruptions in the original schedule. The authors develop a branching algorithm to identify expensive flight connections and find alternative solutions. Moreover, this algorithm allows branching on multiple variables without invalidating the optimality of the algorithm, and it adds buffer times to solutions becoming less sensitive to disruptions. Computational experiments show that fewer flight changes are needed, compared with the deterministic approach. Takahashi et al. (2008) address re-scheduling decisions for DSP problem, in which, after a disruption occurs, the problem is to minimize the violation of the constraints, such as the number of drivers, train coverage, meal times, continuous shifts, and work time bounds. A TS is designed to solve the problem and the experimentation stage in a real example shows that the solutions generated surpass (in terms of the satisfaction of each constraint) the ones obtained by the manual process employed by the agency.

Table 7
Literature review for the Driver Scheduling Problem.

Authors (year)	Objective	Constraints of the system	Solution method	Case
Smith and Wren (1988)	Min duty costs	Number of duties and uncovered work	Mathematical programming	Test
Desrochers and Soumis (1989)	Min cost of driver schedules	Labor regulations	CG	Real
Kwan et al. (1999)	Min duty costs	Duty coverage	GA	Test
Lourenço et al. (2001)	Multi-objective min of uncovered work, infeasibility, number of duties, and number of vehicle changes	Duty coverage	GRASP, TS & GA	Test
Shen and Kwan (2001)	Minimize soft-constraints violation	Max working time, limited continuous work without break, and meal breaks	TS	Benchmark
Dias et al. (2002)	Min duty costs	Trip coverage and labor regulations constraints	GA	Test
Fores et al. (2002)	Min duty costs or number of shifts or min cost with minimum number of shifts	Number of shifts, meal break constraints	IP & Heuristic	Test
Mingozzi et al. (1999)	Min costs of duties	Duty coverage and initial set of duties	B&B	Test
Boschetti et al. (2004)	Min duty costs	Working time, number of drivers	LR & CG	Test
Abbink et al. (2005)	Min duty costs	Labor regulations and variation within the duties	CG, LR, & Heuristic	Test
Yen and Birge (2006)	Min expected value of future actions after a disruption	Labor regulations	CG	Test
Zhao (2006)	Min number of duties	Working time, meal breaks, No. vehicles changes	Heuristic	Test
Takahashi et al. (2008)	Min constraints violation after disruption	Labor regulations and number of drivers	TS	Example
Portugal et al. (2009)	Min weighted sum of duty costs and non-covered work	Limited duties and minimum work covered	CPLEX	Test
Steinzen et al. (2009)	Min duty costs	Trip coverage and same starting and ending depot	Local branching	Test
Chen and Niu (2012)	Min drivers' idle time	Working time without break, total working time and working time and break duration	TS	Test
Chen and Shen (2013)	Min duty costs	Labor regulations, number of shifts	CG	Test
Chen et al. (2013)	Min number of shifts and operational costs	Labor regulation constraints and meal breaks in specific time interval	IP & Heuristic	Test
Kecskeméti and Bilics (2013)	Min duty costs	Spread time, change over bounds, rest time, and duty types	IP & EA	Test
Li and Kwan (2003)	Min working time vs Min ratio of total worked time to spread-over vs Min number of duties	Duty coverage and initial set of duties	GA	Benchmark
Tóth and Krész (2013)	Min cost of driver schedules	Labor regulations	Greedy heuristic	Real

Table 8
Example of cyclic roster consisting of four sequences of work for a planning period of five days.

Sequences	Day 1	Day 2	Day3	Day 4	Day 5
$s^{(1)}$	$\text{duty}_1^{(1)}$	$\text{duty}_2^{(1)}$	$\text{duty}_3^{(1)}$	$\text{duty}_4^{(1)}$	$\text{duty}_5^{(1)}$
$s^{(2)}$	$\text{duty}_1^{(2)}$	$\text{duty}_2^{(2)}$	$\text{duty}_3^{(2)}$	$\text{duty}_4^{(2)}$	$\text{duty}_5^{(2)}$
$s^{(3)}$	$\text{duty}_1^{(3)}$	$\text{duty}_2^{(3)}$	$\text{duty}_3^{(3)}$	$\text{duty}_4^{(3)}$	$\text{duty}_5^{(3)}$
$s^{(4)}$	$\text{duty}_1^{(4)}$	$\text{duty}_2^{(4)}$	$\text{duty}_3^{(4)}$	$\text{duty}_4^{(4)}$	$\text{duty}_5^{(4)}$

Multi-objective DSP have been also tackled through metaheuristics. For example, Lourenço et al. (2001) consider criteria such as drivers' costs, the number of work pieces (parts of a duty between two breaks) not covered, total number of duties, one-piece-work duties, and the number of vehicle changes. The authors propose multi-objective versions of the Greedy Randomized Adaptive Search Procedure (GRASP), TS, and GA to solve the problem. Li and Kwan (2003) propose a fuzzified weighted objective function to represent several measures of driver schedules, such as total working time, ratio of total worked time to spread-over, and number of work pieces. Then, the authors develop a GA with fuzzified criteria in the fitness function to solve Benchmark instances of the problem.

The previous studies summarized in Table 7, could be used to design decision support tools for decision makers. Software such as HASTUS, IMPACTS, TRACS II, and GOAL BUS have implemented operations research techniques in order to achieve a more efficient operation (see Lourenço et al., 2001). However, there are still interesting open research areas to explore such as multi-objective and re-scheduling approaches.

4.3. Driver Rostering Problem

The Driver Rostering Problem (denoted as DRP) *determines the assignment of drivers to the daily duties yielded by the DSP solution for a specific planning period, e.g. a month*. This assignment, called roster, must comply with labor rules and the company's regulations. These regulations are not only important for the workers themselves, but also for the employer, since worker satisfaction lowers the incidence of illness, accidents and absenteeism. The constraints that are often considered are the following: days-off for drivers, specific day-off such as holy days and weekends, limited consecutive work days for drivers, not assigning a late night duty just before an early morning duty, limits for working hours within periods of specific length, equity in the drivers' rosters and rotation for each driver, i.e. a driver should not be assigned always to the worst period of the day.

There are two types of rosters: the cyclic and the non-cyclic roster.The cyclic roster scheme assign the same schedule of duties to different drivers shifted by few days. Table 8 shows an example of four sequences of duties covering five days which can be used to create cyclic rosters. The non-cyclic case relaxes the similarity of rosters for different drivers by allowing different sequences of duties per driver.

Formulations such as set partitioning and set covering can be used to model the DRP. Due to the difficulty of obtaining feasible solutions considering all types of constraints, mathematical formulations define hard constraints such as duty coverage and the usual objectives are based on minimizing wages costs, minimizing idle times, and minimizing the penalization of violated soft constraints (such as the agency's and the drivers' preferences) among others. In the following sections, we will present some of the recent developments related with the DRP.

4.3.1. Cyclic Driver Rostering Problem

To solve the cyclic DRP in railway systems, Hartog et al. (2009) propose to divide the problem into two modules: (i) creating a pattern where a duty (early duty, late duty, rest day, and so on) is assigned to each day; and (ii) assigning specific duties to different places in the roster. These modules are formulated as assignment problems with additional constraints and the goal is to minimize the total sum of the penalties, which are determined by undesirable combinations of duties, days off, etc. The authors implement the solver of CPLEX 9.1 to solve instances of up to 188 drivers over a planning period of 13 weeks.

Obtaining dual bounds is an important contribution to efficiently solve complex integer optimization problems. Caprara et al. (1998) implement Lagrangian relaxations to obtain dual bounds for the DRP that minimizes driver costs subject to hard constraints related to the rests for each roster. The dual bounds are obtained by relaxing the rest of the constraints. Moreover, the authors design a Lagrangian heuristic to find feasible solutions using information from dual bounds. The proposed solution approach is able to find solutions with gaps of nearly 0.1% for real size instances (up to 900 duties) in only hours of computational time.

4.3.2. Non-cyclic Driver Rostering Problem

Nurmi et al. (2011) divide the DRP into two sub-problems solved sequentially: (i) the days-off scheduling problem, which assigns rest days between working days to drivers over a given planning horizon; and (ii) the shift scheduling problem, which assigns drivers to shifts specifying the starting time and duration of the shifts. The objective of the problem is minimizing the penalization of violated drivers' and operators' preferences. The solution approach is a population-based local search method which is tested on a Finnish transit network with 68 drivers. Numerical results show that the minimum number of drivers is at most 50 over the course of a week. Moreover, by implementing different penalization parameters, the authors are able to generate a pool feasible solutions.

Since the DRP has a significant impact on both agencies and drivers, it is natural to find multi-objective approaches to tackle it. For example Moz et al. (2009) propose a DRP to minimize the total overtime during the roster period and the number of drivers whose work load is below the level specified by the work contract. The authors implement two solution

Table 9
Literature review for the Driver Rostering Problem.

Authors (year)	Objective	Constraints of the system	Solution method	Case
Caprara et al. (1998)	Drivers' costs	Duty coverage and rest constraints	Lagrangian heuristic	Real
Hartog et al. (2009)	Min total sum of penalties	Labor regulations, cyclic roster and undesirable patterns	Heuristic	Test
Moz et al. (2009)	Min overtime vs. min number of drivers	Days off per week, specific days off, and length of consecutive work	EA & SPEA2	Real
Nurmi et al. (2011)	Min penalization of violated constraints	Duty coverage and work regulations	EA	Real
Montalva et al. (2012)	Max global driver welfare	Labor regulations and payment bounds for specific drivers	Analytical	Example
Respício et al. (2013)	Min overtime vs. min number of drivers	Days off per week, specific days off, and length of consecutive work	EA & SPEA2 and length of consecutive work	Real Test

algorithms: (i) an EA based on utopic and lexicographic individuals (called UGH); and (ii) the improved Strength Pareto Evolutionary Algorithm (called SPEA2, see Zitzler et al., 2002). The only difference between these algorithms is the way the Pareto front is approximated. Computational experiments for a specific case of a bus transit company in Portugal demonstrate that both algorithms can obtain good results while consuming an acceptable amount of CPU time. However, although SPEA2 is able to find a larger number of potentially non-dominated points, it is unable to reach regions of the objective space which are explored by the UGH. Recently, Respício et al. (2013) present a bi-objective approach to minimize the maximum workload of the drivers and the drivers' wages. Typical constraints, such as work length, rest days, and duty coverage, are considered. The authors hybridize Evolutionary Algorithms with search procedures in order to efficiently explore the solution space.

Recently, the preference of drivers for flexible shifts, i.e. different working loads for different days, has been studied. An analysis of different characteristics of drivers' shifts is presented by Miranda et al. (2008). The authors propose using a mixed logit model based on stated preference data for the choice of working shifts. The characteristics considered in this study are: days off, workday period, weekly distribution of working hours, use of split shifts, and weekly shift variability. The authors show that some drivers prefer heterogeneous of shifts, while one-third of them manifest a systematic preference for the typical shift structure (eight continuous working hours). Similarly, Montalva et al. (2012) show that flexible shifts may also be implemented to improve the productivity of drivers and increase their job satisfaction. Therefore, flexible shifts could be considered by optimization approaches which lead to new problem formulations for the DRP. Details of the studies of DRP presented in this review are in Table 9.

5. Integrating sub-problems of the Transit Network Planning process

Even when it is possible to obtain high quality solutions for each sub-problem of the TNP. Sequential approaches lead to solutions that not necessarily guarantee a cohesive solution for the planning problem as a whole. To overcome this situation, recent studies address the integration of two or more sub-problems of the TNP. There are two common approaches to integrate two or more problems: (i) solving partial integrated formulations that consider some characteristics of one problem while taking decisions of other subproblems, or/and iterative sequential approaches where the main idea is to explore the degrees of freedom of the integrated subproblems by iterations; and (ii) defining complete integrated formulations and/or solution approaches that jointly determine the decisions of the problems. Since complete integrations consider all degrees of freedom of each sub-problem, they are more difficult to define and to handle. Now, we show the different kind of approaches that have arisen in recent years for integrating the different sub-problems of the TNP (see Table 10).

5.1. Integrating Transit Network Timetabling and Vehicle Scheduling Problem

The integration of TNT and VSP defines a bridge between tactical and operational planning. Since SDVSP is easy to solve, it is straightforward to implement iterative approaches which modify the current timetable and then, solves the Vehicle Scheduling Problem. If a complete integration is desired, the departure times of trips would be decision variables. In this case, network flow formulations/algorithms for VSP would be difficult to implement since the model lacks a fixed network, instead, it deals with a set of potential ones depending on timetabling decisions. On the basis of the above, usual solution approaches are partial integrations. For example, Ceder (2001) proposes a solution method to integrate the TNT and the SDVSP consisting of the following steps: (i) generating timetables with average even loads using a graphical procedure; (ii) determining the minimum number of vehicles needed to cover those trips; (iii) implementing deadheading and

Table 10
Literature review of holding strategies for the real-time control process.

Authors (year)	Par	Objective	Cap	Control P	Veh	Method
Barnett (1974)	Det	$W_{first} + W_{extra}$	No	PSS	One	Heuristic
Abkowitz and Lepofsky (1990)	Det	W_{first}	No	PSS	One	Heuristic
Eberlein et al. (2001)	Det	W_{first}	No	PSS	Multi	Heuristic
Hall et al. (2001)	Stoch	W_{trans}	No	PMS	One	Analytical
Hickman (2001)	Stoch	$W_{\text{first}} + W_{\text{in-veh}}$	No	PSS	One	Analytical
Chandrasekar et al. (2012)	Det	W_{first}	Yes	PMS	One	Heuristic
Fu and Yang (2002)	Det	W_{first}	No	PMS	One	Preset rules
Zhao et al. (2003)	Stoch	$W_{\text{first}} + W_{\text{in-veh}}$	No	MSC	One	Heuristic
Puong and Wilson (2004)	Det	$W_{\text{first}} + W_{\text{in-veh}} + W_{\text{extra}}$	Yes	MSC	Multi	B&C
Zolfaghari et al. (2004)	Det	$W_{\text{first}} + W_{\text{extra}}$	Yes	SSC	Multi	Heuristic
Puong and Wilson (2008)	Det	$W_{\text{first}} + W_{\text{in-veh}} + W_{\text{extra}}$	Yes	MSC	Multi	Heuristic & B&C
Sun and Hickman (2008)	Det	$W_{\text{first}} + W_{\text{in-veh}}$	No	PMS	Multi	Heuristic
Xuan et al. (2011)	Stoch	W_{first}	No	PMS	Multi	Heuristic
Bartholdi and Eisenstein (2012)	Det	V_h	Yes	PMS	Multi	Math
Delgado et al. (2009, 2012)	Det	$W_{\text{in-veh}} + W_{\text{extra}}$	Yes	PMS	Multi	MINOS
Delgado et al. (2013)	Det	$W_{\text{first}} + W_{\text{in-veh}} + W_{\text{trans}}$	Yes	PSS	One	MINOS

departure time shifting procedures to reduce the costs of that vehicle schedule; and (iv) adjusting the initial timetable and obtaining the final vehicle schedules. The proposed approach is tested on some examples and the author states that the obtained solutions are efficient schedules from both the passenger and operator perspectives. Another example of partial integrations is presented by van den Heuvel et al. (2008). The authors address the integration of clock-face timetabling and the MDVSP. They develop a Tabu Search where the main iterations consist of modifying the timetable, and then optimizing the MDVSP. The proposed approach is tested on real life data sets of a Dutch transit network. Numerical results indicate that a significant reduction of operational costs can be achieved when there is more flexibility in the type and number of vehicles performing the service trips and some trips are split and evenly spread in time. Guihaire and Hao (2008b) implement a similar iterative sequential idea to integrate the TNT minimizing waiting times and the VSP. The authors propose an ILS, where each iteration implements trip-shifting operators and then, solves the VSP given the current timetable. The solution approach is used to solve some examples where the number of vehicles is reduced up to 26% while the number of feasible transfers increases up to 44%.

Liu and Shen (2007) integrate the TNT presented by Liu et al. (2007) and a MDVSP minimizing the number of buses and deadhead costs. To solve the problem, the authors develop a Bi-level Nesting Tabu Search which is implemented in a small example. Guihaire and Hao (2010b) propose a mathematical formulation for the integration of TNT and VSP minimizing a weighted objective function based on the following elements: (i) number and quantity of transfers; (ii) headway evenness; (iii) fleet size; and (iv) deadhead costs. The model consider a limited deviation from an initial timetable which allow the authors define feasible shifting procedures. These procedures are embedded in an Iteratel Local Search. The solution approach is tested on an instance based on real data and numerical results show the following: the number of vehicles drops by 28.35%; the number of transfers increases by 112.33%; the number of headways not fitting into the allowed variation margin diminishes by 75.72%; only the value of the deadheads objective deteriorates.

Ceder (2011) addresses an interactive heuristic approach to find timetables with even loads and even headways which take advantage of different vehicle types. The later approach is a combination of previously defined tools (Ceder, 2007) to compute timetables with even loads, timetables with even headways, and vehicle schedules with minimum operational costs.

Petersen et al. (2013) formulate a partial integration of TNT and MDVSP minimizing the weighted sum of the costs of vehicles usage and passenger transfers. The authors consider *metatrips* which are a sets of trips that contains an original trip of the initial timetable and also contains the "same trip" but with different departure times. Then, a feasible vehicle schedule is forced to assign one, and only one trip, of each metatrip to a vehicle, i.e. a large feasible space is considered for the MDVSP. The authors propose a Large Neighborhood Search metaheuristic to solve the proposed formulation. Numerical results show that the flexibility of the timetables (in the form of metatrips) allows savings of up to 20% in the passenger transfer waiting times compared with the current operation.

The weight parameters in a weighted objective functions must represent the planner's preferences for the different objectives. The later is an issue if two or more objectives are in conflict, as happens with TNT and VSP. Then, Ibarra-Rojas et al. (2014) propose a bi-objective optimization problem to jointly solve the SDVSP and the synchronization bus timetabling problem considering time windows for departure times and assuming constant demand. The objectives are maximizing the number of passengers benefited by well-timed transfers and minimizing the fleet size. The authors implement an ϵ-constraint algorithm to obtain Pareto optimal solutions; thus, they are able to measure the "cost" of a vehicle in terms of passengers transfers and vice versa. The experimental stage is performed using generated instances based on a transit network in Mexico. Numerical results show that in some instances using one more vehicle lead to significant improvements in passenger transfers.

5.2. Integrating Vehicle Scheduling Problem and Driver Scheduling Problem

In the early 1980's, the predominant approach towards integrating the VSP and the DSP was including Driver Scheduling considerations into the VSP, or vice verse (see Scott, 1985; Darby-Dowman et al., 1988; Tosini and Vercellis, 1988; Falkner and Ryan, 1992; Patrikalakis and Xerocostas, 1992). The common solution approach was implementing iterative sequential algorithms which take only one solution of the VSP and then, solve the DSP. More recently, exact and heuristic approaches have been developed considering an objective function based on operational costs including both vehicles usage and driver wages.

5.2.1. Exact approaches to integrate Vehicle Scheduling Problem and Driver Scheduling Problem

Haase et al. (2001) introduce a set partition formulation for DSP including side constraints for duty flow variables and bus itineraries. This formulation guarantees that an optimal vehicle schedule can always be derived later. To solve the problem, the authors design a Branch & Price (B&P) algorithm that relies on several acceleration strategies, e.g. a dynamic generation of bus count constraints and the appropriate substitution of partitioning constraints in order to reduce the column density. Borndörfer et al. (2002) also propose a CG based on Lagrangian relaxation to solve a linked multi-commodity network flow and set partitioning formulation for the integrated VSP and DSP. Their proposed approach is able to obtain feasible solutions for instances of up to 1457 trips and one depot. However, it takes 6.5 h to solve the largest instance. Other studies based on set partition/covering formulations, CG approaches, and Lagrangian heuristics, are presented by Freling et al. (2001a) and Freling et al. (2003).

The MDVSP has been integrated with the DSP by Gaffi and Nonato (1999). The authors propose a formulation based on a quasi-assignment model for the MDVSP and a set of linking constraints that ensure compatibility of the driver schedule. Constraints considered in the model are number of vehicle blocks per depot and the number of duties of a specific type. The authors assume that a driver must be assigned to the same vehicle for the whole planning period, and each work piece starts and ends at the same depot. Then, work pieces and vehicle blocks coincide, which makes the problem computationally much more tractable than it would have been without these assumptions. The authors develop CG approaches that allow solving instances of up to 257 trips and 28 depots using a computational time between two and six hours on average (but needing more than 24 h for the largest instance).

de Groot and Huisman (2004) propose to divide the MDVSP according to the following rules: (i) splitting the problem into several instances of SDVSP and DSP, i.e. assigning each trip to a depot; and (ii) splitting an instance into a predetermined number of smaller ones based on the number of trips and trip-vehicle assignments. The implement the proposed rules on large instances (up to 653 trips and 1.74 depots/trip) based on data sets of a transit network in Netherlands. Numerical results suggest that the effect of dividing these instances did not significantly deteriorate the quality of the solutions while the proposed approach outperforms the classical sequential method.

Borndörfer et al. (2008) develop a Lagrangian relaxation approach for the VDSP minimizing overall operating costs. The proposed approach is tested on instances of up to 1500 trips and three depots. Numerical results show a quick convergence to obtain optimal solutions. Mesquita and Paias (2008) formulate the VDSP as a partitioning/covering model. The authors show several conditions where just a subset of decision variables are required to be integers. Based on this information, they propose a solution approach consisting of: (i) solving a linear relaxation of the problem through CG; and (ii) implementing a B&B algorithm to find integer solutions. The solution approach is tested on instances of up to 400 trips and four depots (presented by Huisman et al., 2005). Numerical results indicate that the proposed approach seems to be more efficient than previous approaches presented by Borndörfer et al. (2008), Huisman et al. (2005).

5.2.2. Heuristic approaches to integrate Vehicle Scheduling Problem and Driver Scheduling Problem

The degrees of freedom of complete integrated formulations may lead to time consuming exact approaches. Then, a reasonable alternative is to develop metaheuristic algorithms in order to obtain high quality solutions in acceptable computational time (see Tosini and Vercellis, 1988; Falkner and Ryan, 1992; Patrikalakis and Xerocostas, 1992). For example, Valouxis and Housos (2002) address the VDSP considering limits on the daily work and the minimum daily shift breaks that a driver must have. Since the driver remains with the bus for the duration of the shift, rules for both drivers and buses must be satisfied when shifts are created and assigned to a particular bus. The previous characteristics of the system are modeled in a linear integer program and the authors implement a solution approach based on local search algorithms and CG to solve instances of up to 89 buses and 337 trip segments.

Steinzen et al. (2007) present a mathematical formulation for a multi-depot case of the VDSP minimizing the sum of vehicle and crew cost, considering constraints. The authors develop a EA consisting of: (i) generating potential population members exploring trip-depot assignments; and (ii) implementing a fitness function that determines the vehicle and driver schedules by combining CG with Lagrangian Relaxation. The proposed solution approach is tested on instances of up to 200 trips and four depots. Numerical results show reductions between 5.1% and 10.0% for the number of duties compared with the current operation.

Laurent and Hao (2008) propose a constraint programming formulation for the VDSP minimizing the number of vehicles and drivers subject to constraints for maximum spread time, minimum working time, and driver changeovers. The authors develop a GRASP algorithm which is implemented on instances up to 249 trips, 50 drivers, and 47 vehicles. Numerical results show, that even when optimality is not guaranteed by using the proposed GRASP, the obtained solutions are of better quality compared with the sequential method.

Another GRASP implementation is proposed by Leone et al. (2011). The authors address the VDSP with several hard constraints, such as limited spread-over, a limited number of breaks, upper bounds for idle time, limited working time, limited resting time, and compatibility between trips and drivers. Due to the large number of constraints in the proposed mathematical formulation, small instances are solved with a commercial solver, while the GRASP algorithm is used to obtain feasible solutions for instances of up to 400 trips.

To tackle parameters uncertainty in the planning process, Huisman et al. (2004) propose a dynamic approach for the multiple-depot VDSP, i.e. defining a schedule for the next l time units considering the following elements: (i) known travel times; (ii) decisions already made earlier that can no longer be changed; and (iii) estimation of future after the next l time units. The authors present mathematical formulations and solution algorithms for both single-depot and multi-depot cases. The optimization problems consist in minimize operational costs based on fixed vehicle costs and usage of vehicles and drivers. The authors propose several solution algorithms based on Lagrangian Relaxation based heuristics, CG, and Branch & Bound algorithms. Numerical results show that the proposed approaches surpass the static version of the integrated problems and it is suitable for the single-depot case, while it does not perform so well in the case of multiple depots.

5.2.3. Integrating Vehicle Scheduling Problem and Driver Scheduling Problem with flexible timetables

Gintner et al. (2008) define a sequential approach for VSP and DSP but considering multiple solutions for the VSP. The authors use the time-space network formulation developed by Kliewer et al. (2006) in order to obtain a pool of optimal vehicle schedules. Then, the authors solve the DSP by a solution approach based on LR and CG. Numerical results show significant savings compared with the classical sequential approach.

Kéri and Haase (2007a) and Kéri and Haase (2007b) address a VDSP minimizing the number of buses and crew costs subject to constraints for task covering, bus-driver coupling, and a flexible timetable. This flexibility is defined based on *flexible groups*, i.e, sets of trips that could be shifted together in order to keep the level of service in terms of waiting times. The authors design a solution algorithm based on CG and show that the proposed approach yields a much better solution in terms of the number of buses and the drivers' costs, but it requires much more time to solve the problem.

Recently, Kliewer et al. (2012) address the VDSP with the possibility of shifting scheduled trips within defined time windows, proposing a solution methodology based on Gintner et al. (2006) and Steinzen et al. (2010), where the main idea is to enhance these solution approaches so as to handle time windows. While flexibility is used to reduce operational costs, a penalty is employed for deviations from the initial timetabling so that the level of service does not deteriorate.

5.3. Integrating Driver Scheduling Problem and Driver Rostering Problem

Abbink et al. (2011) present an integer programming formulation for the DSP considering the following elements to guarantee the feasibility of the DRP: average duration of daily duties; maximum number of duties assigned to a driver base; limited number of duties with duration outside a time interval; and limited night (or weekend) duties. The authors propose a Lagrangian based heuristic combined with Column Generation to solve the problem. The proposed approach is tested on real data sets of a transit network in Netherlands and numerical results show savings up to 3 million euros (1% of efficiency gain compared with the current operation).

Xie et al. (2012) integrate the DSP and the cyclic version of the DRP maximizing the drivers' satisfaction. The authors they design a Simulated Annealing algorithm which is tested on two medium-size instances (of up to 256 drivers and 1013 duties) based on real data from a transit network in Germany. Numerical results show that the running time of the proposed integrated approach is not much longer than the running time of the sequential one (about 10 min for each instance). Besides, drivers' satisfaction is high for at least 71% of the drivers.

Recently, Xie and Suhl (2014) integrate the cyclic and non-cyclic versions of the DRP and the DSP. The problem considers hard constraints, such as maximum daily working time and number of days off; and soft constraints, such as minimum sum of days off, limited number of unpopular single-off activities, separation time between two double-off activities, and maximum overtime for all drivers. To jointly determine cyclic and non-cyclic rosters, the authors proposed a multi-commodity network flow formulation based on the model developed for the airline crew Rostering problem by Cappanera and Gallo (2004). The proposed formulation can be efficiently solved for instances up to 3273 duties and the authors present a comparison of sequential vs integrated approaches based on a German transit network where the integrated approach lead to better quality feasible solutions.

5.4. Integrating Vehicle Scheduling Problem, Driver Scheduling Problem, and Driver Rostering Problem

Shen and Xia (2009) present an iterative sequential approach to solve VSP, DSP, and DRP (denoted as VDRP). The authors address the context of Chinese transport systems where specific time windows for drivers' rest should be considered. Moreover, practical constraints include: a driver should be assigned to the same bus before and after having lunch; drivers can be assigned to at most two vehicles; and vehicles must be identified since compatible driver-vehicle pairs are given as an input. The transit network of the case study requires a line-by-line planning process and bus schedules are changed frequently; thus, time consuming algorithms should be avoided. The authors propose an iterative sequential heuristic algorithm consisting of the following steps: (i) solving the VSP with a local search based on n-opt operators; (ii) solving the DSP with a Tabu Search based on Shen and Kwan (2001); and (iii) constructing the driver rosters and allow to be modified through an user-friendly interface. Numerical results show that it is possible to find feasible solutions for instances up to 107 buses and 164 duties in minutes of computational time. Savings for vehicles costs and driver wages are near to 4.5% and 9.9%, respectively, compared with solutions obtained by human expertise.

Mesquita et al. (2011) present a non-linear multi-objective formulation for the VDRP which minimizes the number of drivers (cost criterion) and minimizes the maximum overtime assigned to a driver (preference criterion). Since the formulation has a large number of constraints and it is unlikely to be solved efficiently by commercial solvers, a preemptive goal programming based heuristic is designed to decompose the VDRP as follows: (i) solving a VDSP per day; and (ii) determining a roster for a given time horizon. Extensive experimentation using data from a bus company in Lisbon shows that the proposed iterative sequential approach easily obtain optimal solutions in short computing times where minimization of the number of drivers alone was concerned. However, 17.8% of the instances were not solved within a reasonable time limit when all the costs were considered in the objective function.

More recently, Mesquita et al. (2013) address a VDRP minimizing vehicle and driver costs subject to constraints regarding roster balancing and coverage of all daily duties. An important characteristic is that driver rosters must be compatible with predefined days-off patterns based on the requirements of a Portuguese public transit company. The problem considers a planning period of seven weeks, where a set of drivers must be assigned to a fleet of vehicles from a set of depots in order to cover given planned trips. The authors propose a mathematical formulation with elements of problems such as multi-commodity network flow, set covering, and covering-assignment. This formulation allows the authors to implement a Bender decomposition that outperforms the traditional sequential approach.

6. Real-time control strategies

Real-time control strategies are needed in transit networks since the absence of these procedures leads to undesired behavior of the transport system, such as vehicle bunching due to the stochastic nature of traffic flows and passenger demand at bus stops. This may cause an increment in the variance of the headways and a consequent worsening of both the magnitude and variability of average waiting times.

In terms of the spatial configuration of the different control strategies, they can be classified into the following categories: station control and inter-station control.

6.1. Station control strategies

In station control strategies decisions are taken at some stops along the line. The most common approaches are the following:

- Holding strategies are used to increase service regularity (see Turnquist and Blume, 1980) and/or reduce transfer time between lines (Hall et al., 2001; Delgado et al., 2013). There are two types of bus holding strategies to increase regularity: (i) headway-based which is commonly used for lines with short headways (e.g., less than 10 min); and (ii) schedule-based which is typical on lines with long headways for meeting a timetable. Also, holding strategies are sometimes implemented to avoid passengers missing a connection to another line.
- Stop-skipping strategies in which a bus might be asked to skip a stop as long as nobody inside the bus requests it, even if there are passengers waiting to board the bus.
- Boarding limits strategies can be considered a continuous version of the stop-skipping strategy. Here, a fraction of the passengers waiting to board the bus might be requested to wait for the next bus even when the bus has available capacity.

We present a review of station control strategies in the following section.

6.1.1. Headway-based holding strategies

The main idea of headway-based holding strategies is to provide reliable service in the form of evenly spaced arrivals at stops. For example, the early study of Abkowitz and Lepofsky (1990) addresses a real-time strategy minimize the total waiting time. This strategy holds buses at specific stops until a minimum headway time is achieved, in order to. An implementation is done on two lines of a transit network in U.S. Based on the numerical results, the authors conclude that some characteristics that favor the implementation of holding strategies are the following: frequency-based operation, randomness of passenger arrivals, and selection of control points based on boarding/alighting profiles. Fu and Yang (2002) present a formulation that minimizes headway variation in order to minimize the average waiting time at different stops. However, in-vehicle holding delay and bus capacity constraints are not considered. In such a case, the results demonstrate that the optimal policy is to employ holding at a single stop. Sun and Hickman (2008) state the same problem in two dimensions, with holding of multiple buses but at a defined set of control stops. They show that if holding is applied at various stations, bus headways can be regularized and greater cost reductions thereby achieved, compared with single station holding.

Recently, Daganzo (2009) propose an adaptive control scheme aiming to provide quasi-regular headways while maintaining as high a commercial speed as possible. The control strategy dynamically determines bus holding times at a line's control points based on real-time information about the arrival time of the previous bus. Although the method proves to be efficient under small disturbances, under large perturbations it reduces performance. Other approach based on providing regular service is presented by Xuan et al. (2011). The authors propose holding strategies to regularize headways while maximizing commercial speeds. The assumption of their approach is that information about the current bus and the leading bus is available. Based on numerical results, the authors state that their approach outperforms headway-based methods when it is used only to regulate headways.

Bartholdi and Eisenstein (2012) rules out a priori headway since they state that trying to adjust the trip regularity to a unique and static headway leads to the unsatisfactory performance of the system. Instead, they propose a holding technique depending on the state of the system in order to obtain a self coordination strategy where the natural headway arises from the system behavior over time. In this strategy, the holding time is computed based on the separation with the previous bus which reduces the management work and simplifies the job of the drivers. The proposed approach is tested on a real line and simulated scenarios. Numerical results suggest that it could be useful in cases with large disruptions or surge in ridership, by re-distributing the buses to equalize the headways. Similarly, with reduced traffic, the common headway is spontaneously reduced and service improves. Cats et al. (2011, 2012) evaluate different holding control strategies for improving service reliability. Moreover, transit performance is also evaluated from the operator perspective by considering the impacts of holding strategies on fleet operations and drivers management. The evaluation is based on a simulation process, and the numerical results show that there are potential benefits from implementing an evaluation holding strategy that regulates headways depending on both the preceding and succeeding buses, as proposed by Bartholdi and Eisenstein (2012).

6.1.2. Holding strategies to minimize waiting times

The previous studies assume that regular service satisfies the passenger demand in an efficient way, but there are cases where there is high variability of demand (such as passenger arriving as a bunch) and waiting times should be considered. In these cases, it is necessary to implement control strategies with variable headways, but based on passenger waiting times.

The early research of Barnett (1974) develop a holding model at a given control stop, where the sum of total waiting time plus the extra delay of passengers on board is minimized.

Eberlein et al. (2001) propose a rolling horizon for the holding problem minimizing the total waiting time and considering safety headways. The control strategy is implemented on a line of the transit network in Boston and numerical results show that the estimated reduction in waiting times to board the first vehicle ranges from about 7% to 31% in the presence of terminal schedule constraints and from about 21% to 50% in the absence of these constraints. Moreover, the result show that holding reduces dwell times and inter-station stopping times, and provides earlier arrival times which could be counter-intuitive at first glance, this can result in. Zhao et al. (2003) propose a control model based on negotiation between two agents: one aboard the bus, and the other at a stop. The optimization problem minimizes waiting times and considers stochastic passenger arrivals at stops. The approach is tested through simulation and numerical results indicate that the negotiation algorithm is robust for a variety of operating conditions such as stationary and different kinds of non-stationary passenger arrivals.

Zolfaghari et al. (2004) formulate an optimization problem which minimizes the waiting times of both of users who arrive at a stop and those who have to wait for more than one bus due to the activation of the capacity constraint. The authors consider vehicle capacity constraints but they do not consider the extra waiting time endured by passengers held at a stop. Puong and Wilson (2008) extend this case by including the latter factor in their objective function in the context of interruptions in train service. They propose a non-linear mixed-integer model in which dwell time is assumed to be constant at any given station. The problem is solved in a reasonable amount of time using a hybrid heuristic and Branch & Cut strategy.

The combination of holding strategies with other control tools has been addressed by Delgado et al. (2012). The authors develop a deterministic optimization model capable of executing two strategies: holding and boarding limits (limits the amount of passengers that can board a bus even when the bus is not at full capacity). The objective function considers the waiting time of the passengers at the bus stops, wait times on the bus due to the holding executed, and the extra waiting time experienced because of the capacity constraint of the bus, or the boarding limits strategy. The authors show that in high frequency and high demand scenarios, the application of both strategies reduced the total waiting time 69%, compared with the no-control scenario, while just applying the holding strategy reduced the total waiting time 61%. In high frequency scenarios where bus capacity is never reached, applying only holding or both strategies yields similar results, reducing the total wait times between 63% and 65%, compared with the no-control scenario. One important finding of their work is that the application of these strategies allows the buses to travel at a lower capacity, ensuring a better balanced load and improving the comfort of the passengers.

6.1.3. Holding strategies to minimize waiting times considering transfers

Another reasonable goal of implementing holding strategies is to allow well-timed passenger transfers. To achieve this, it should be noted that passengers may arrive as a bunch at transfer stops. Then, it is inadequate to model the estimated waiting time as the multiple of the square of the headway and half of the arrival rate. Hall et al. (2001) develop analytical models that determine the optimal holding times at transfer stations with general bus arrival time distributions. The authors test the proposed approach on generated instances based on real data of the transit network of Los Angeles, CA.

Delgado et al. (2013) implement holding strategies to benefit passengers who must walk a little to transfer from line B to another line A at a single stop. Since it is necessary to walk for such transfers, the arrival rate of passengers that will take line A after alighting from B is modeled as a trapezoidal function, i.e. this rate increases at the beginning of a time interval, remains constant at the middle of the interval, and decreases at the end of the interval. The goal is to minimize the following waiting time costs: (i) those of passengers already waiting at the stop; (ii) those of passengers on the held bus; and (iii) those of passengers transferring from B to A. The main assumptions are the arrival rate for both lines and the estimation of the arrival times for buses at each stop. The model is embedded in a rolling horizon approach that is tested through simulation.

Table 7 shows the studies presented in this review. The first column presents the citation; The second column "Par" indicates whether travel times and passenger demand are deterministic (Det) or stochastic (Stoch); The third column shows the objective of the problem that could be based on waiting time to take the first bus (W_{first}), in-vehicle waiting time ($W_{\text{in-veh}}$), extra waiting time of passengers that could not take the first bus (W_{extra}), or waiting time at transfer stops (W_{trans}); The column "Cap" indicates whether using capacity (Yes) or uncapacitated systems (No); The column "Control P" indicates the different types of stops, such Predefined Single Stop (PSS), Predefined Multiple Stops (PMS), Single Stop defined by the Control process (SSC), or Multiple Stops defined by the Control process (MSC); The column "Veh" shows the number of vehicles considered in the optimization process, which can involve a single (One) or multiple (Multi) vehicles; finally, the column "Method" presents the proposed solution approach.

6.1.4. Stop-skipping strategies

Different service patterns, i.e. trips that skip some of the stops of the original line, can be used to increase bus operating speeds. However, the problem is to determine which stops should be ignored considering different criteria, such as operational costs, average waiting time, and passenger awareness, among others. Early work on rail systems is presented by Suh et al. (2002), where a stop-skipping strategy is implemented to increase the speed of subway service. To achieve this, the authors use information from an OD matrix, the distances between stations, the headways, and the maximum link speeds. Numerical results implementing stop-skipping strategies yield average waiting time savings up to 12.9% compared with the current operation.

Instead of design limited stop lines in tactical planning. Fu et al. (2003) present a real-time operating strategy that provides a limited-stop line every other trip minimizing total waiting time costs. The authors develop a mathematical formulation that determines the line pattern depending on the previous two trips and the actual passenger demand. Then, this problem is solved using a rolling time horizon approach to define the stops to be skipped. Since the formulation determines the line pattern for only one trip, the problem can be solved in an exact way in a short time. The authors state that stop-skipping strategies are more effective on bus lines with high passenger demand, short headway, and moderate travel time variability.

Sun and Hickman (2005) propose a stop-skipping strategy to minimize passenger waiting times but allow passengers to alight at stops in the skipped segment. This stop-skipping strategy is formulated as a non-linear integer programming problem, with assumptions as uncertain passenger boarding and alighting times. The size of the instances are small enough to be solved through exhaustive methods, and it is tested through simulation. Numerical results suggest that the downtown-oriented passenger distribution pattern may present the most desirable condition for minimizing passenger waiting times.

A multi-objective optimization approach for stop-skipping strategies is addressed by Sidi et al. (2008). The authors present a decision support system to determine line patterns and their departure times after a disruption occurs and a time horizon is defined. The following different criteria are considered: trip regularity; transfer time; travel time; distance traveled; and the number of unsatisfied stops. Considered constraints are headway bounds, stop times ranges, and vehicle loads. The solution approach is based on fuzzy logic to represent the subjective nature of the decision maker's preferences.

Cortés et al. (2010) and Sáez et al. (2012) develop a hybrid predictive control formulation which jointly estimates bus speed and implements control strategies such as skipping stations or holding buses. This formulation considers a multi-objective case minimizing waiting times and minimizing the cost of in-vehicle travel times. The authors solve the problem through GAs and results of a simulation show savings of 20% and 10% in the total travel time for users when the proposed strategy is implemented. Later, Muñoz et al. (2013) provide a performance comparison of the previous approaches and Delgado et al. (2009, 2012). Different scenarios are considered and numerical results show that the approaches of Delgado et al. (2009, 2012) dominates when capacity constraints are binding.

The main drawback of station control strategies is negatively affecting some passengers (holding inside a bus, preventing from boarding). Also, in many urban/cultural context these strategies may not be feasible. The control strategies presented in the next section can be implemented instead or in addition to these strategies.

6.2. Inter-station control

Inter-station control strategies define decisions that should be taken between stops of the transit networks, i.e. while the bus is traveling between control stops. Some approaches that belong to this kind of strategy are the control of the operating speed and traffic signal priority.

6.2.1. Bus speed regulation strategies

Chandrasekar et al. (2012) develop a strategy based on controlling the bus's speed along its route. The main idea is that passengers will not notice that they are being held (speed up is not considered as a control strategy), as happens when a bus is held at a stop. By implementing the proposed approach, the authors show that bus speed control lead to a decrease in headway irregularity. For example, 80% of the cases studied saw a reduction of more than 10% in excess waiting time, and 60% of the cases had more than 20% reduction.

More recently, Daganzo and Pilachowski (2011) present a control strategy that continuously adjust the bus cruising speed based on a cooperative two-way to achieve a proper space between the front and back buses. Such a cooperative control is shown to be effective in preventing bunching.

6.2.2. Traffic Signal Priority strategies

Traffic Signal Priority Strategies (called as TSP) optimize the traffic flow by defining the combination of traffic lights (called phase) that favors a specific circulation at an intersection, considering priority calls of vehicles and estimation of arrival times. These strategies should consider feasibility constraints based on the queuing process along the entire network, i.e. we should not benefit a specific intersection if the following arcs possess full capacity (Liu et al., 2003). Such a strategy depends on three main elements: (i) delay estimation; (ii) weight determination (commonly based on the number of passengers in the vehicles); and (iii) the optimization problem to define the traffic signals. Studies are oriented to cyclic, i.e. repetitiveness of traffic lights phases over time, and non-cyclic cases where the traffic lights' phases are properly defined due to the actual traffic flow.

The benefits in terms of average network speed when implementing TSP are shown by Diakaki et al. (2007). The authors implement an extension of the traffic responsive urban control strategy considering bus priority signals. This strategy consists of four modules: (i) split control to minimize the risk of over saturation and queue spill back; (ii) cycle control to adapt the cycle duration to the currently observed maximum saturation level in the network; (iii) offset control which creates successive green lights considering vehicle queues; and (iv) Priority of public transportation. The strategy is tested on the Jerusalem network, obtaining average travel times savings of 65% compared with the current operation. Kraus et al. (2010) implement the previous strategy in Maxaé, Brazil and numerical results report improvements of up to a 25% increase in average network speed.

Simultaneous traffic lights and bus holding is addressed by van Oort et al. (2012). The goal is the minimization of a cost based on car delays, passenger delays, and bus stop cost which is relevant in cases where it is preferred to hold a bus at a stop instead of to continue its trip and immediately stop at an intersection or a queue due to red lights. Two strategies are proposed: (i) immediate departure where the intersection controller grants priority (e.g., extending a green light) upon request of the bus (approaching the intersection); and (ii) controlled departure which avoids a second stop of the bus at the end of the queue formed during a red light by holding the bus at the bus stop (near the intersection), while still granting priority to the bus lane. The analysis shows that controlled departure is the best strategy for small values of passenger car flows in the same link, and it is the recommended strategy when there are segregated bus lanes, as in the context of Bus Rapid Transit (BRT) systems.

Yang and Lou (2013) formulate a signal optimization procedure based on the interaction of two modules: (i) a GA to compute the phases of the traffic lights along with their sequences, and (ii) a microscopic simulation that analyzes the flow of individual vehicles and computes the total delays for a given signal timing plan obtained through the GA. The latter study consider both cyclic and non-cyclic TSP. Numerical results show that the improvement of the non-cyclic case over the cyclic one is a decrease of up to 6.5% in delays. Han et al. (2014) address a TSP strategy that dynamically defines the weights in the objective functions to reflect the changing necessities of the network under different conditions. To achieve this, the authors define a weighted objective function with two parameters, α and β, to represent the weight of the common vehicles and of the buses, respectively. Then, before solving the optimization problem, α and β are determined through a sensitivity analysis where the benefits and general traffic interferences are measured. On the basis of this analysis, the values of the parameters are fixed to strike a balance between delay reduction and delay increase.

Recently, He et al. (2014) consider a multi-modal case for the TSP strategy such that the signal timing is responsive to real-time actions of non-priority demand by allowing phases to extend and gap out. Moreover the authors compare the proposed approach with the state-of-practice transit signal priority (TSP), both under the optimized signal timing plans, using a microscopic traffic simulation based on a section of Speedway Blvd. in Tucson, Arizona. The later approach implements robust and responsive priority policies to resolve the conflict between multiple priority requests from an optimum prospective. Numerical results show that the proposed control model is able to reduce the average bus delay up to 24.9%, especially for highly congested conditions.

Real-time control strategies have been implemented to improve the daily operation of transit systems. Moreover, the increasing developments in data collecting tools are leading to a better understanding of the decision process of the system, and new technologies such as vehicle-to-vehicle communication, could be used to enhance these control strategies.

7. Conclusions

In this paper, we present a literature review on optimization approaches used for planning, operation, and control of bus transport systems. The types of problems studied were divided according to the planning horizon into strategic, tactical, and operational levels. We presented decision support techniques which are playing an important role on improving the efficiency of transit systems.

Although there has been a fruitful development of models and solution techniques to address relevant decision problems in bus transport systems, there are still many open research questions, such as the following:

- There are still opportunities for integrating problems that are solved separately, which forces a multi-criteria analysis approach. For instance, the integration of timetabling and operational planning has shown promising results, however only few authors have tackled this type of integrated problems.
- The urbanization process the world is facing creating several megapolis around the globe. In this context, bus service becomes increasingly slow and expensive since trips and congestion grows. The adoption of fare integration world-wide that forces some passengers to transfer opens the door to rethink the Transit Network Design by considering different variants of limited-stop lines and transfers coordination. These new types of lines could complement the more standard ones to improve mobility and accessibility city-wide.
- Another important aspect to handle is the natural uncertainty that transit systems face such as passenger demand and travel times. The increasingly amount of data generated through AVL, APC, and GPS allow us to consider these uncertainties to develop a more robust transit planning and operation. In this context, research opportunities to taking into account uncertainty are the development of robust, stochastic, and dynamic approaches for vehicle/Driver Scheduling problems.

Acknowledgments

This research was partially supported by CONICYT FONDECYT Projects 3140358 and 1120993, the Centro de Desarrollo Urbano Sustentable (CEDEUS), Conicyt/Fondap/15110020, the Bus Rapid Transit Centre of Excellence funded by the Volvo Research and Educational Foundations (VREF) We thank to the anonymous referees for their valuables comments to improve this review.

References

Abbink, E., Albino, L., Dollevoet, T., Huisman, D., Roussado, J., Saldanha, R., 2011. Solving large scale crew scheduling problems in practice. Public Transport 3, 149–164.
Abbink, E., Kroon, L., Vromans, M., 2005. Reinventing crew scheduling at netherlands railways. Interfaces 35, 393–401.
Abkowitz, M., Lepofsky, M., 1990. Implementing headway-based reliability control on transit routes. Journal of Transportationn Engineering 116, 49–63.
Afanasiev, L., Liberman, S.Y., 1982. Principles for orginizing express bus services. Transportation Research Part A 17, 343–346.
Amiripour, S., Ceder, A., Mohaymany, A., 2014. Designing large-scale bus network with seasonal variations of demand. Transportation Research Part C 48, 322–338.
Argwal, J., Mathew, T., 2004. Transit route network design using parallel genetic algorithm. Journal of Computing in Civil Engineering 18, 248–256.
Baaj, M., Mahmassani, H., 1991. An AI-based approach for transit route system. Journal of Advanced Transportation 25, 187–209.
Baaj, M., Mahmassani, H., 1992. AI-based system representation and search procedures for transit route network design. Transportation Research Record 1358, 67–70.
Baaj, M., Mahmassani, H., 1995. A hybrid route generation heuristic algorithm for the design of transit networks. Transportation Research Part C 3, 31–50.
Bagloee, S., Ceder, A., 2011. Transit-network design methodology for actual-size road networks. Transportation Research Part B 45, 1787–1804.
Baita, F., Pesenti, R., Ukovich, W., Favaretto, D., 2000. A comparison of different solution approaches to the vehicle scheduling problem in a practical case. Computers and Operations Research 27, 1249–1269.
Barnett, A., 1974. On controlling randomness in transit operations. Transportation Science 8, 101–116.
Bartholdi III, J.J., Eisenstein, D.D., 2012. A self-coördinating bus route to resist bus bunching. Transportation Research Part B 46, 481–491.
Bertossi, A., Carraresi, P., Gallo, G., 1987. On some matching problems arising in vehicle scheduling models. Networks 17, 271–281.
Bertsekas, D., 1992. Auction algorithms for network flow problems: a tutorial introduction. Computational Optimization and Applications 1, 7–66.
Black, A., 1979. Optimizing urban mass transit systems: a general model. Transportation Research Record 677, 41–47.
Bookbinder, J.H., Désilets, A., 1992. Transfer optimization in a transit network. Transportation Science 26, 106–118.
Borndörfer, R., Löbel, A., Weider, S., 2002. Integrierte Umlauf und Dienstplanung im Nahverkehr. Technical Report 02-10. ZIB, Konrad-Zuse Zentrum. Berlin, German.
Borndörfer, R., Löbel, A., Weider, S., 2008. A Bundle Method for Integrated Multi-depot Vehicle and Duty Scheduling in Public Transit, vol. 600. Springer, Berlin Heidelberg, pp. 3–24.
Boschetti, M., Mingozzi, A., Ricciardelli, S., 2004. An exact algorithm for the simplified multiple depot crew scheduling problem. Annals of Operations Research 127, 177–201.
Bunte, S., Kliewer, N., 2009. An overview on vehicle scheduling models. Public Transport 1, 299–317.
Byrne, B., 1975. Public transportation line positions and headways for minimum user and system cost in a radial case. Transportation Research Record 9, 97–102.
Byrne, B., Vuchic, V., 1972. Public transportation line positions and headways for minimum cost. In: Newell, F. (Ed.), Proceedings of the Fifth International Symposium on Traffic Flow and Transportation, pp. 347–360.
Cappanera, P., Gallo, G., 2004. A multicommodity flow approach to the crew rostering problem. Operations Research 52, 583–596.
Caprara, A., Toth, P., Vigo, D., Fischetti, M., 1998. Modeling and solving the crew rostering problem. Operations Research 46, 820–830.
Castelli, L., Pesenti, R., Ukovich, W., 2004. Scheduling multimodal transportation systems. European Journal of Operational Research 155, 603–615.
Cats, O., Larijani, A., Koutsopoulos, H., Burghout, W., 2011. Impacts of holding control strategies on transit performance: bus simulation model analysis. Transportation Research Record 2216, 51–58.
Cats, O., Larijani, A., Ólafsdóttir, A., Burghout, W., Andrésson, I., Koutsopoulos, H., 2012. Bus-holding control strategies: simulation-based evaluation and guidelines for implementation. Transportation Research Record 2274, 100–108.
Ceder, A., 1984. Bus frequency determination using passenger count data. Transportation Research Part A 18, 439–453.
Ceder, A., 1989. Optimal design of transit short-turn trips. Transportation Research Record 1221, 8–22.
Ceder, A., 2001. Efficient timetabling and vehicle scheduling for public transport. In: Voß, S., Daduna, J.R. (Eds.), Computer-Aided Scheduling of Public Transport, Lecture Notes in Economics and Mathematical Systems, vol. 505. Springer, Berlin Heidelberg, pp. 37–52.
Ceder, A., 2003. Designing public transport network and routes. In: Lam, W., Bell, M. (Eds.), Advanced Modeling for Transit Operations and Service Planning, Pergamon Imprint. Elsevier Science Ltd Pub., pp. 59–91, chapter 3.
Ceder, A., 2007. Public Transit Planning and Operation: Theory Modeling and Practice. Elsevier, Butterworth-Heinemann.
Ceder, A., 2011. Optimal multi-vehicle type transit timetabling and vehicle scheduling. Procedia Social and Behavioral Sciences 20, 19–30.
Ceder, A., Golany, B., Tal, O., 2001. Creating bus timetables with maximal synchronization. Transportation Research Part A 35, 913–928.
Ceder, A., Hassold, S., Dano, B., 2013. Approaching even-load and even-headway transit timetables using different bus sizes. Public Transport.
Ceder, A., Tal, O., 2001. Designing synchronization into bus timetables. Transportation Research Record: Journal of the Transportation Research Board 1760, 28–33.
Ceder, A., Wilson, N., 1986. Bus network design. Transportation Research Part B 20, 331–344.
Cepeda, M., Cominetti, R., Florian, M., 2006. A frequency-based assignment model for congested transit networks with strict capacity constraints: characterization and computation of equilibria. Transportation Research Part B 40, 437–459.
Chakroborty, P., 2003. Genetic algorithms for optimal urban transit network design. Computer-Aided Civil and Infrastructure Engineering 18, 184–200.
Chakroborty, P., Deb, K., Porwal, H., 1997. A genetic algorithm based procedure for optimal transit system scheduling. In: Proceedings of Fifth International Conference on Computers in Urban Planning and Urban Management, Mumbai, India. pp. 330–341.
Chakroborty, P., Deb, K., Subrahmanyam, P.S., 1995. Optimal scheduling of urban transit systems using genetic algorithms. Journal of Transportation Engineering 121, 544–553.
Chakroborty, P., Dwivedi, T., 2002. Optimal route network design for transit systems using genetic algorithms. Engineering Optimization 34, 83–100.
Chandrasekar, P., Long Cheu, R., Chin, H., 2012. Simulation evaluation of route-based control of bus operations. Journal of Transportation Engineering 128, 519–527.
Chen, A., Yang, C., 2004. Stochastic transportation network design problem with spatial equity constraint. Transportation Research Record 1882, 97–104.
Chen, M., Niu, H., 2012. A model for bus crew scheduling problem with multiple duty types. Discrete Dynamics in Nature and Society 2012.
Chen, S., Shen, Y., 2013. An improved column generation algorithm for crew scheduling problems. Journal of Information & Computation Science 10, 175–183.
Chen, S., Shen, Y., Su, X., Chen, H., 2013. A crew scheduling with chinese meal break rules. Journal of Transportation Systems Engineering and Information Technology 13, 90–95.

Chew, J., Lee, L., Seow, H., 2013. Genetic algorithm for biobjective urban transit routing problem. Journal of Applied Mathematics 2013, 15 pages.
Chien, S., Schonfeld, P., 1997. Optimization of grid transit system in heterogeneous urban environment. Journal of Transportation Engineering 123, 28–35.
Chien, S., Spasovic, L., 2002. Optimization of grid bus transit systems with elastic demand. Journal of Advanced Transportation 36, 63–91.
Chiou, S.W., 2005. Bilevel programming for the continuous transport network design problem. Transportation Research Part B 39, 361–383.
Chiraphadhanakul, V., Barnhart, C., 2013. Incremental bus service design: combining limited-stop and local bus services. Public Transport 5, 53–78.
Cipriani, E., Gori, S., Petrelli, M., 2012. A bus network design procedure with elastic demand for large urban areas. Public Transport 4, 57–76.
Clarens, G., Hurdle, V., 1975. An operating strategy for a commuter bus system. Transportation Science 9, 1–20.
Constantin, I., Florian, M., 1995. Optimizing frequencies in a transit network: a nonlinear bi-level programming approach. International Transactions in Operational Research 2, 149–164.
Cortés, C., Jara-Díaz, S., Tirachini, A., 2011. Integrating short turning and deadheading in the optimization of transit services. Transportation Research Part A 45, 419–434.
Cortés, C., Saéz, D., Milla, F., Nuñez, A., Riquelme, M., 2010. Hybrid predictive control for real-time optimization of public transport systems' operations based on evolutionary multi-objective optimization. Transportation Research Part C 18, 757–769.
Daduna, J., Paixão, J., 1995. Vehicle scheduling for public mass transit, an overview. In: Computer-Aided Transit Schedulig, pp. 76–90.
Daduna, J.R., Voß, S., 1995. Practical experiences in schedule synchronization. In: Daduna, J., Branco, I., Paixão, J. (Eds.), Computer-Aided Scheduling of Public Transport, Lecture Notes in Economics and Mathematical Systems, vol. 430. Springer, Berlin, pp. 39–55.
Daganzo, C., 2010. Structure of competitive transit networks. Transportation Research Part B 44, 434–446.
Daganzo, C., Pilachowski, J., 2011. Reducing bunching with bus-to-bus cooperation. Transportation Research Part B 45, 267–277.
Daganzo, C.F., 2009. A headway-based approach to eliminate bus bunching: systematic analysis and comparisons. Transportation Research Part B 43, 913–921.
Darby-Dowman, K., Jachnik, R., Mitra, G., 1988. Integrated decision support systems for urban transport scheduling: discussion of implementation and experience. In: Daduna, J., Wren, A. (Eds.), Computer-Aided Transit Scheduling, Lecture Notes in Economics and Mathematical Systems, vol. 308. Springer, Berlin, pp. 226–239.
Deb, K., Pratap, A., Agarwal, S., Meyarivan, T., 2002. A fast and elitist multiobjective genetic algorithm: NSGA-II. Transactions on Evolutionary Computation 6, 182–197.
Delgado, F., Contreras, N., Muñoz, J., 2013. Holding for transfers. In: Proceeding of the Transportation Research Board Meeting.
Delgado, F., Munoz, J., Giesen, R., 2012. How much can holding and limiting boarding improve transit performance? Transportation Research Part B 46, 1202–1217.
Delgado, F., Muñoz, J., Giesen, R., Cipriano, A., 2009. Real-time control of buses in a transit corridor based on vehicle holding and boarding limits. Transportation Research Record 2090, 59–67.
Delle Site, P., Filippi, F., 1998. Service optimization for bus corridors with short-turn strategies and variable vehicle size. Transportation Research Part A 32, 19–38.
dell'Olio, L., Ibeas, A., Ruisánchez, F., 2012. Optimizing bus-size and headway in transit networks. Transportation 39, 449–464.
Desaulniers, G., Hickman, M., 2007. Public transit. In: Barnhart, C., Laporte, G. (Eds.), Handbooks in operations research and management science, Transportation, vol. 14. North-Holland, pp. 69–128.
Desrochers, M., Soumis, F., 1989. A column generation approach to the urban transit crew scheduling problem. Transportation Science 23, 1–13.
Diakaki, C., Dinopoulou, V., Aboudolas, K., Papageorgiou, M., Ben-Shabat, E., Seider, E., Leibov, A., 2007. Extensions and new applications of the traffic-responsive urban control strategy: coordinated signal control for urban networks. Transportation Research Record 1856, 202–211.
Dias, T., de Sousa, J., Cunha, J., 2002. Genetic algorithms for the bus driver scheduling problem: a case study. Journal of the Operational Research Society 53, 1–12.
Domschke, W., 1989. Schedule synchronization for public transit systems. OR Spectrum 11, 17–24.
Eberlein, X.J., Wilson, N., Bernstein, D., 2001. The holding problem with real-time information available. Transportation Science 35, 1–18.
El-Geneidy, A., Surprenant-Legault, J., 2010. Limited-stop bus service: an evaluation of an implementation strategy. Public Transport 2, 291–306.
Estrada, M., Roca-Riu, M., Badia, H., Robusté, F., Daganzo, C., 2011. Design and implementation of efficient transit networks: procedure, case study and validity test. Transportation Research Part A 45, 935–950.
Falkner, J., Ryan, D., 1992. Express: set partitioning for bus crew scheduling in Christchurch. In: Desrochers, M., Rousseau, J.M. (Eds.), Computer-Aided Transit Scheduling, Lecture Notes in Economics and Mathematical Systems, vol. 386. Springer, Berlin, pp. 359–378.
Fan, W., Machemehl, R., 2004. Optimal Transit Route Network Design Problem: Algorithms, Implementations, and Numerical Results. Technical Report SWUTC/04/167244-1. University of Austin. Texas, USA.
Fan, W., Machemehl, R., 2006a. Optimal transit route network design problem with variable transit demand: a genetic algorithm approach. Journal of Transportation Engineering 132, 40–51.
Fan, W., Machemehl, R., 2006b. Using a simulated annealing algorithm to solve the transit route network design problem. Journal of Transportation Engineering 132, 122–132.
Fan, W., Machemehl, R., 2008. Tabu search strategies for the public transportation network optimizations with variable transit demand. Computer-Aided Civil and Infrastructure Engineering 23, 502–520.
Fan, W., Machemehl, R., 2011. Bi-level optimization model for public transportation network redesign problem. Transportation Research Record 2263, 151–162.
Farahani, R., Miandoabchi, E., Szeto, W., Rashidi, H., 2013. A review of urban transportation network design problems. European Journal of Operational Research 229, 281–302.
Fernández, J., de Cea, J., Malbran, R., 2008. Demand responsive urban public transport system design: methodology and application. Transportation Research Part A 42, 951–972.
Fores, S., Proll, L., Wren, A., 2002. TRACS II: a hybrid IP/heuristic driver scheduling system for public transport. Journal of the Operational Research Society 53, 1093–1100.
Freling, R., Huisman, D., Wagelmans, A., 2001a. Applying an integrated approach to vehicle and crew scheduling in practice. In: Voß, S., Daduna, J. (Eds.), Computer-Aided Scheduling of Public Transport, Lecture Notes in Economics and Mathematical Systems, vol. 505. Springer, Berlin, Germany, pp. 73–90.
Freling, R., Huisman, D., Wagelmans, A.P.M., 2003. Models and algorithms for integration of vehicle and crew scheduling. Journal of Scheduling 6, 63–85.
Freling, R., Wagelmans, A., Paixão, J., 2001b. Models and algorithms for single-depot vehicle scheduling. Transportation Science 35, 165–180.
Freyss, M., Giesen, R., Muñoz, J., 2013. Continuous approximation for skip-stop operation in rail transit. Procedia - Social and Behavioral Sciences 80, 186–210.
Fu, L., Liu, Q., Calamai, P., 2003. Real-time optimization model for dynamic scheduling of transit operations. Transportation Research Record 1857, 48–55.
Fu, L., Yang, X., 2002. Design and implementation of bus-holding control strategies with real-time information. Transportation Research Record 1791, 6–12.
Furth, P., 1986. Zonal route design for transit corridors. Transportation Science 20, 1–12.
Furth, P., 1987. Short turning on transit routes. Transportation Research Record 1108, 42–52.
Furth, P., Day, F., 1985. Transit routing and scheduling strategies for heavy demand corridors. Transportation Research Record 1011, 23–26.
Furth, P., Rahbee, A., 2000. Optimal bus stop spacing through dynamic programming and geographic modeling. Transportation Research Record 1731, 15–22.
Furth, P., Wilson, N., 1981. Setting frequencies on bus routes: theory and practice. Transportation Research Record 818, 1–7.
Gaffi, A., Nonato, M., 1999. An integrated approach to ex-urban crew and vehicle scheduling. In: Wilson, N. (Ed.), Computer-Aided Transit Scheduling, Lecture Notes in Economics and Mathematical Systems, vol. 471. Springer, Berlin, Germany, pp. 103–128.

Gao, Z., Sun, H., Shan, L., 2004. A continuous equilibrium network design model and algorithm for transit systems. Transportation Research Part B 38, 235–250.
Gintner, V., Kliewer, N., Suhl, L., 2005. Solving large multiple-depot multiple-vehicle-type bus scheduling problems in practice. OR Spectrum 27, 507–523.
Gintner, V., Kliewer, N., Suhl, L., 2008. A crew scheduling approach for public transit enhanced with aspects from vehicle scheduling. In: Hickman, M., Mirchandani, P., Voß, S. (Eds.), Computer-aided Systems in Public Transport, vol. 600. Springer, Berlin Heidelberg, pp. 339–359.
Gintner, V., Steinzen, I., Suhl, L., 2006. A time-space network based approach for integrated vehicle and crew scheduling in public transport. In: Proceedings of the EURO working group on transportation (EWGT) joint conferences 2006, pp. 371–377.
de Groot, S., Huisman, D., 2004. Vehicle and crew scheduling : solving large real-world instances with an integrated approach. In: Proceedings of the 9th International Conference on Computer-Aided Scheduling of Public Transport.
Guan, J., Yang, H., Wirasinghe, S.C., 2003. Simultaneous optimization of transit line configuration and passenger line assignment. Transportation Research Part B 40, 885–902.
Guihaire, V., Hao, J.K., 2008a. Transit network design and scheduling: a global review. Transportation Research Part A 42, 1251–1273.
Guihaire, V., Hao, J.K., 2008b. Transit network re-timetabling and vehicle scheduling. In: Le Thi, H.A., Bouvry, P., Pham Dinh, T. (Eds.), Modelling, Computation and Optimization in Information Systems and Management Sciences, Communications in Computer and Information Science, vol. 14. Springer, Berlin Heidelberg, pp. 135–144.
Guihaire, V., Hao, J.K., 2010a. Improving timetable quality in scheduled transit networks. In: Proceedings of the 23rd international conference on Industrial engineering and other applications of applied intelligent systems, IEA/AIE '10, pp. 21–30.
Guihaire, V., Hao, J.K., 2010b. Transit network timetabling and vehicle assignment for regulating authorities. Computers and Industrial Engineering 59, 16–23.
Haase, K., Desaulniers, G., Desrosiers, J., 2001. Simultaneous vehicle and crew scheduling in urban mass transit systems. Transportation Science 35, 286–303.
Hadas, Y., Shnaiderman, M., 2012. Public-transit frequency setting using minimum-cost approach with stochastic demand and travel time. Transportation Research Part B 46, 1068–1084.
Hadjar, A., Marcotte, O., Soumis, F., 2006. A branch-and-cut algorithm for the multiple depot vehicle scheduling problem. Operations Research 54, 130–149.
Haghani, A., Banihashemi, M., 2002. Heuristic approaches for solving large-scale bus transit vehicle scheduling problem with route time constraints. Transportation Research Part A 36, 309–333.
Hall, R., Dessouky, M., Lu, Q., 2001. Optimal holding times at transfer stations. Computers & Industrial Engineering 40, 379–397.
Han, A., Wilson, N., 1982. The allocation of buses in heavily utilized networks with overlapping routes. Transportation Research Part B 16, 221–232.
Han, X., Li, P., Sikder, R., Qiu, Z., Kim, A., 2014. Development and evaluation of an adaptative transit signal priority control with updated transit delay model. In: Proceedings of the Transportation Research Board 93rd Annual Meeting.
Hartog, A., Huisman, D., Abbink, E., Kroon, L., 2009. Decision support for crew rostering at NS. Public Transport 1, 121–133.
Hassold, S., Ceder, A., 2012. Multiobjective approach to creating bus timetables with multiple vehicle types. Transportation Research Record 2276, 56–62.
Hassold, S., Ceder, A., 2014. Public transport vehicle scheduling featuring multiple vehicle types. Transportation Research Part B 67, 129–143.
He, Q., Head, K., Ding, J., 2014. Multi-modal traffic signal control with priority, signal actuation and coordination. Transportation Research Part C 46, 65–82.
van den Heuvel, A., van den Akker, J., van Kooten, M., 2008. Integrating timetabling and vehicle scheduling in public bus transportation. Technical Report UU-CS-2008-003. Department of Information and Computing Sciences, Utrecht University, Utrecht, The Netherlands.
Hickman, M., 2001. An analytic stochastic model for the transit vehicle holding problem. Transportation Science 35, 215–237.
Holryod, E., 1967. Optimum bus service: a theoretical model for a large uniform urban area. In: Edie, L. (Ed.), Proceedings of the Third International Symposium on the Theory of Traffic Flow. Elsevier, pp. 308–328.
Huang, Z., Ren, G., Liu, H., 2013. Optimizing bus frequencies under uncertain demand: case study of the transit network in a developing city. Mathematical problems in Engineering 2013.
Huisman, D., Freling, R., Wagelmans, A., 2004. A robust solution approach to the dynamic vehicle scheduling problem. Transportation Science 38, 447–458.
Huisman, D., Freling, R., Wagelmans, A.P.M., 2005. Multiple-depot integrated vehicle and crew scheduling. Transportation Science 39, 491–502.
Hurdle, V., 1973. Minimum cost locations for parallel public transit lines. Transportation Science 7, 340–350.
Ibarra-Rojas, O., Giesen, R., Rios-Solis, Y., 2014. An integrated approach for timetabling and vehicle scheduling problems to analyze the trade-off between level of service and operating costs of transit networks. Transportation Research Part B 70, 35–46.
Ibarra-Rojas, O., López-Irarragorri, F., Rios-Solis, Y., 2015. Multiperiod synchronization bus timetabling. Transportation Science. http://dx.doi.org/10.1287/trsc.2014.0578, in press.
Ibarra-Rojas, O., Rios-Solis, Y., 2012. Synchronization of bus timetabling. Transportation Research Part B 46, 599–614.
Jordan, W., Turnquist, M., 1979. Zone scheduling of bus routes to improve service reliability. Transportation Science 13, 242–268.
Karaboga, D., Gorkemli, B., Ozturk, C., Karaboga, N., 2012. A comprehensive survey: artificial bee colony (abc) algorithm and applications. Artificial Intelligence Review, 1–37.
Kecskeméti, B., Bilics, A., 2013. Bus driver duty optimization using an integer programming and evolutionary hybrid algorithm. Central European Journal of Operations Research 21, 745–755.
Kepaptsoglou, K., Karlaftis, M., 2009. Transit route network design problem: review. Journal of Transportation Engineering 135, 491–505.
Kéri, A., Haase, K., 2007a. Simultaneous vehicle and crew scheduling with trip shifting. Operation Research Proceedings 2007, 467–472.
Kéri, A., Haase, K., 2007b. Vehicle and crew scheduling with flexible timetable. Operation Research Proceedings 2006, 339–342.
Kim, M., Schonfeld, P., 2014. Integration of conventional and flexible bus services with timed transfers. Transportation Research Part B 68, 76–97.
Klemt, W.D., Stemme, W., 1988. Schedule synchronization for public transit networks. In: Proceedings of the 4th International Workshop on Computer-Aided Scheduling of Public Transport. Springer Verlag, Hamburg, Germany, pp. 327–335.
Kliewer, N., Amberg, B., Amberg, B., 2012. Multiple depot vehicle and crew scheduling with time windows for scheduled trips. Public Transport 3, 213–244.
Kliewer, N., Mellouli, T., Suhl, L., 2006. A time-space network based exact optimization model for multi-depot bus scheduling. European Journal of Operational Research 175, 1616–1627.
Kramkowski, S., Kliewer, N., Meier, C., 2009. Heuristic methods for increasing delay-tolerance of vehicle schedules in public bus transport. In: Proceedings of the Metaheuristic International Conference VIII (MIC 2009), Hamburg, Germany.
Kraus, W., de Souza, F., Papageorgiou, M., Dantas, L., Camponogara, E., Kosmatopoulos, E., Aboudolas, K., 2010. Cost effective real-time traffic signal control using the TUC strategy. IEEE Intelligent Transportation Systems 2, 6–17.
Kwan, A., Kwan, R., Wren, A., 1999. Driver scheduling using genetic algorithms with embedded combinatorial traits. In: Wilson, N. (Ed.), Computer-Aided Transit Scheduling, Lecture Notes in Economics and Mathematical Systems, vol. 471. Springer, Verlag, Heidelberg, pp. 81–102.
Kwan, C.M., Chang, C.S., 2008. Timetable synchronization of mass rapid transit system using multiobjective evolutionary approach. IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews 38, 636–648.
Larrain, H., Giesen, R., Muñoz, J., 2010. Choosing the right express services for bus corridor with capacity restrictions. Transportation Research Record 2197, 63–70.
Laurent, B., Hao, J., 2009. Iterated local search for the multiple depot vehicle scheduling problem. Computers and Industrial Engineering 57, 277–286.
Laurent, B., Hao, J.K., 2008. Simultaneous vehicle and crew scheduling for extra urban transports. In: Proceedings of the 21st international conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems: New Frontiers in Applied Artificial Intelligence. Springer-Verlag, Berlin, Heidelberg, pp. 466–475.
Leiva, C., Muñoz, J., Giesen, R., Larrain, H., 2010. Design of limited-stop services for an urban bus corridor with capacity constraints. Transportation Research Part B 44, 1186–1201.

Leone, R., Festa, P., Marchitto, E., 2011. A bus driver scheduling problem: a new mathematical model and a GRASP approximate solution. Journal of Heuristics 17, 441–466.
Li, J., Kwan, R., 2003. A fuzzy genetic algorithm for driver scheduling. European Journal of Operaitonal Research 147, 334–344.
Li, J., Mirchandani, P., Borenstein, D., 2007. The vehicle rescheduling problem: model and algorithms. Networks 50, 211–229.
Li, J.Q., Mirchandani, P., Borenstein, D., 2009. A lagrangian heuristic for the real-time vehicle rescheduling problem. Transportation Research Part E 45, 419–433.
Li, Y., Xu, W., He, S., 2013. Expected value model for optimizing the multiple bus headways. Applied Mathematics and Computation 219, 5849–5861.
Li, Z.C., Lam, W., Wong, S., Sumalee, A., 2010. An activity-based approach for scheduling multimodal transit services. Transportation 37, 751–774.
Liebchen, C., Stiller, S., 2012. Delay resistant timetabling. Public Transport 4, 55–72.
Liu, H., Skabardonis, A., Zhang, W., 2003. A dynamic model for adaptative bus signal priority. In: Proceedings of the Transportation Research Board 82nd Annual Meeting, p. 20.
Liu, Z., Shen, J., 2007. Regional bus operation bi-level programming model integrating timetabling and vehicle scheduling. Systems Engineering - Theory & Practice 27, 135–141.
Liu, Z., Shen, J., Wang, H., Yang, W., 2007. Regional bus timetabling model with synchronization. Journal of Transportation Systems Engineering and Information Technology 7, 109–112.
Löbel, A., 1998. Vehicle scheduling in public transit and lagrangean pricing. Management Science 44, 1637–1650.
Lodwick, W., Untiedt, E., 2010. Introduction to fuzzy and possibilistic optimization. In: Lodwick, W., Kacprzyk, J. (Eds.), Fuzzy Optimization, Studies in Fuzziness and Soft Computing, vol. 254. Springer, Berlin Heidelberg, pp. 33–62.
Lourenço, H., Paixão, J., Portugal, R., 2001. Multiobjective metaheuristics for the driver scheduling problem. Transportation Science 35, 331–343.
Lownes, N., Machemehl, R., 2010. Exact and heuristic methods for public transit circulator design. Transportation Research Part B 44, 309–318.
Mandl, C., 1979. Evaluation and optimization of urban public transportation network. European Journal of Operational Research 5, 396–404.
Martínez, H., Mauttone, A., Urquhart, M., 2014. Frequency optimization in public transportation systems: formulation and metaheuristic approach. European Journal of Operational Research 236, 27–36.
Marwah, B., Umrigar, F., Patnaik, S., 1984. Optimal design of bus routes and frequencies for Ahmedabad. Transportation Research Record 994, 41–47.
Mauttone, A., Urquhart, M., 2009. A multiobjective metaheuristic approach for the transit network design problem. Public Transport 1, 253–273.
Medina, M., Giesen, R., Muñoz, J., 2013. Model for the optimal location of bus stops and its application to a public transport corridor in Santiago, Chile. Transportation Research Record 2352, 84–93.
Mesa, J., Ortega, F., Pozo, M., 2013. Locating optimal timetables and vehicle schedules in a transit line. Annals of Operations Research 222, 439–455.
Mesquita, M., Moz, M., Paias, A., Paixão, J., Pato, M., Respício, A., 2011. A new model for the integrated vehicle-crew-rostering problem and a computational study on rosters. Journal of Scheduling 14, 319–334.
Mesquita, M., Moz, M., Paias, A., Pato, M., 2013. A decomposition approach for the integrated vehicle-crew-roster problem with days-off pattern. European Journal of Operational Research 229, 318–331.
Mesquita, M., Paias, A., 2008. Set partitioning/covering-based approaches for the integrated vehicle and crew scheduling problem. Computers & Operations Research 35, 1562–1575.
Mingozzi, A., Boschetti, M., Ricciardelli, S., Bianco, L., 1999. A set partitioning approach to the crew scheduling problem. Operations Research 47, 873–888.
Miranda, F., Muñoz, J., Ortúzar, J., 2008. Identifying transit driver preferences for work shift structures: an econometric analysis. Transportation Science 42, 70–86.
Mohaymany, A., Amiripour, S., 2009. Creating bus timetables under stochastic demand. International Journal of Industrial Engineering and Production Research 20, 83–91.
Montalva, S., Muñoz, J., Paredes, R., 2012. Assignment of work shifts to public transit drivers based on stated preferences. Public Transport 2, 199–218.
Moz, M., Respício, A., Pato, M.V., 2009. Bi-objective evolutionary heuristic for bus driver rostering. Public Transport 1, 189–210.
Muñoz, J., Cortés, C., Giesen, R., Sáez, D., Delgado, F., Valencia, F., Cipriano, A., 2013. Comparison of dynamic control strategies for transit operations. Transportation Research Part C 28, 101–113.
Naumann, M., Suhl, L., Kramkowski, S., 2011. A stochastic programming approach for robust vehicle scheduling in public bus transport. Procedia Social and Behavioral Sciences 20, 826–835.
Nayeem, M., Rahman, M., Rahman, M., 2014. Transit network design by genetic algorithm with elitism. Transportation Research Part C 46, 30–45.
van Nes, R., Hamerslag, R., Immers, B., 1988. Design of public transport networks. Transportation Research Record 1202, 74–83.
Newell, G., 1971. Dispatching policies for a transportation route. Transportation Science 5, 91–105.
Nikolić, M., Teodorović, D., 2013. Transit network design by bee colony optimization. Expert Systems with Applications 40, 5945–5955.
Nikolić, M., Teodorović, D., 2014. A simultaneous transit network design and frequency setting: computing with bees. Expert Systems with Applications 41, 7200–7209.
Nurmi, K., Kyngäs, J., Post, G., 2011. Driver rostering for bus transit companies. Engineering Letters 19, 125–132.
van Oort, N., Boterman, J., van Nes, R., 2012. The impact of scheduling on service reliability: trip-time determination and holding points in long-headway services. Public Transport 4, 39–56.
Ortúzar, J., Willumsen, L., 2001. Modelling Transport. John Wiley and Sons Ltd., Chichester, England.
Ouyang, Y., Nourbakhsh, S., Cassidy, M., 2014. Continuum approximation approach to bus network design under spatially heterogeneous demand. Transportation Research Part B 68, 333–344.
Palma, A.D., Lindsey, P., 2001. Optimal timetables for public transportation. Transportation Research Part B 35, 789–813.
Patrikalakis, I., Xerocostas, D., 1992. A new decomposition scheme of the urban public transport scheduling problem. In: Desrochers, M., Rousseau, J.M. (Eds.), Proceedings of the Fifth Inter-national Workshop on Computer-aided Scheduling of Public Transport, Lecture Notes in Economics and Mathematical Systems, vol. 386. Springer, Berlin, pp. 407–425.
Pepin, A., Desaulniers, G., Hertz, A., Huisman, D., 2009. A comparison of five heuristics for the multiple depot vehicle scheduling problem. Journal of Scheduling 12, 17–30.
Petersen, H., Larsen, A., Madsen, O., Petersen, B., Ropke, S., 2013. The simultaneous vehicle scheduling and passenger service problem. Transportation Science 47, 603–616.
Portugal, R., Lourenço, H., Paixão, J., 2009. Driver scheduling problem modelling. Public Transport 1, 103–120.
Puong, A., Wilson, N., 2004. A train holding model for urban transit systems. In: 9th International Conference on Computer-Aided Scheduling of Public Transport (CASPT), San Diego, California.
Puong, A., Wilson, N., 2008. A train holding model for urban rail transit. In: Hickman, M., Mirchandani, P., Voß, S. (Eds.), Computer-aided Systems in Public Transport, vol. 600. Springer, Berlin Heidelberg, pp. 319–337.
Raveau, S., Muñoz, J., de Grange, L., 2011. A topological route choice model for metro. Transportation Research Part A 45, 138–147.
Respício, A., Moz, M., Pato, M., 2013. Enhanced genetic algorithms for a bi-objective bus driver rostering problem: a computational study. International Transactions in Operations Research 20, 443–470.
Roca-Riu, M., Estrada, M., Trapote, C., 2012. The design of interurban bus networks in city centers. Transportation Research Part A 46, 1153–1165.
Sáez, D., Cortés, C., Milla, F., Nuñez, A., Tirachini, A., Riquelme, M., 2012. Hybrid predictive control strategy for a public transport system with uncertain demand. Transportmetrica 8, 61–86.
Salicrú, M., Fleurent, C., Armengol, J., 2011. Timetable-based operation in urban transport: run-time optimisation and improvements in the operating process. Transportation Research Part A 45, 721–740.
Salzborn, F., 1972. Optimum bus scheduling. Transportation Science 6, 137–148.

Scheele, S., 1980. A supply model for public transit services. Transportation Research Part B 14, 133–146.
Schmöcker, J., Fonzone, A., Shimamoto, H., Kurauchi, F., Bell, M., 2011. Frequency-based transit assignment considering seat capacities. Transportation Research Part B 45, 392–408.
Schröder, M., Solchenbach, I., 2006. Optimization of Transfer Quality in Regional Public Transit. Technical Report 84. Berichte des Fraunhofer Instituts for Techno-und Wirtschafts.
Scott, D., 1985. A large scale linear programming approach to the public transport scheduling and costing problem. In: Rousseau, J.M. (Ed.), Computer-Aided Scheduling of Public Transport. Elsevier, Amsterdam.
Shen, Y., Kwan, R., 2001. Tabu search for driver scheduling. In: Voß, S., Daduna, J. (Eds.), Computer-Aided Scheduling of Public Transport, Lecture Notes in Economics and Mathematical Systems, vol. 505. Springer-Verlag, Heidelberg, pp. 121–135.
Shen, Y., Xia, J., 2009. Integrated bus transit scheduling for the Beijing bus group based on a unified mode of operation. International Transactions in Operational Research 16, 227–242.
Shrivastava, P., Dhingra, S., 2002. Development of coordinated schedules using Genetic Agoritms. Journal of Transportation Engineering 128, 89–96.
Shrivastava, P., Dhingra, S., Gundaliya, P., 2002. Application of Genetic Algorithm for scheduling and schedule coordination problems. Journal of Advanced Transportation 36, 23–41.
Shrivastava, P., O'Mahony, M., 2006. A model for development of optimized feeder routes and coordinated schedules – a genetic algorithms approach. Transport Policy 13, 413–425.
Shrivastava, P., O'Mahony, M., 2009. Modeling an integrated public transportation system – a case study in Dublin, Ireland. European Transport 41, 1–19.
Sidi, M., Hammadi, S., Hayat, S., Borne, P., 2008. Urban transport network regulation and evaluation: a fuzzy evolutionary approach. IEEE Transactions on Systems, Man and Cybernetics Part A 38, 309–318.
Sivakumaran, K., Li, Y., Cassidy, M., Madanat, S., 2012. Cost-savings properties of schedule coordination in a simple trunk-and-feeder transit system. Transportation Research Part A 46, 131–139.
Smith, B., Wren, A., 1988. A bus crew scheduling system using a set covering formulation. Transportation Research Part A 22, 97–108.
Steinzen, I., Becker, M., Suhl, L., 2007. A hybrid evolutionary algorithm for the vehicle and crew scheduling problem in public transit. In: Proceedings of the IEEE Congress on Evolutionary Computation, Singapore. pp. 25–28.
Steinzen, I., Gintner, V., Suhl, L., Kliewer, N., 2010. A time-space network approach for the integrated vehicle and crew scheduling problem with multiple depots. Transportation Science 44, 367–382.
Steinzen, I., Suhl, L., Kliewer, N., 2009. Branching strategies to improve regularity of crew scheduling in ex-urban public transit. OR Spectrum 31, 727–743.
Suh, W., Chon, K., Rhee, S., 2002. Effect of skip-stop policy on a korean subway system. Transportation Research Record 1793, 33–39.
Sun, A., Hickman, M., 2004. Scheduling considerations for a branching transit route. Journal of Advanced Transportation 38, 243–290.
Sun, A., Hickman, M., 2005. The real-time stop-skipping problem. Journal of Intelligent Transportation Systems 9, 91–109.
Sun, A., Hickman, M., 2008. The holding problem at multiple holding stations. In: Proceedings of the 9th International Conference on Computer-Aided Scheduling of Public Transport (CASPT), San Diego, California.
Szeto, W., Jiang, Y., 2012. Hybrid artificial bee colony algorithm for transit network design. Transportation Research Record 2284, 47–56.
Szeto, W., Jiang, Y., 2014. Transit route and frequency design: Bi-level modeling and hybrid artificial bee colony algorithm approach. Transportation Research Part B 67, 235–263.
Szeto, W., Wu, Y., 2011. A simultaneous bus route desing and frequency setting problem for Tin Shui Wai, Honk Kong. European Journal of Operational Research 209, 141–155.
Takahashi, S., Kataoka, K., Kojima, T., Asami, M., 2008. An algorithm for automatically modifying train crew schedule. IEEJ Transactions on Industry Applications 128, 1291–1297.
Tétreault, P., El-Geneidy, A., 2010. Estimating bus run times for new limited-stop service using archived AVL and APC data. Transportation Research Part A 44, 390–402.
Tilahun, S., Ong, H., 2012. Bus timetabling as a fuzzy multiobjective optimization problem using preference-based genetic algorithm. PROMET - Traffic & Transportation 24, 183–191.
Tosini, E., Vercellis, C., 1988. An interactive system for extra-urban vehicle and crew scheduling problems. In: Daduna, J., Wren, A. (Eds.), Computer-Aided Transit Scheduling, Lecture Note in Economics Mathematical Systems, vol. 308. Springer, Berlin Heidelberg, pp. 41–53.
Tóth, A., Krész, M., 2013. An efficient solution approach for real-world driver scheduling problems in urban bus transportation. Central European Journal of Operations Research 21, 75–94.
Turnquist, M., Blume, S., 1980. Evaluating potencial effectiveness of headway control strategies for transit systems. Transportation Research Record 746, 25–29.
Ulusoy, Y., Chien, S., Wei, C.H., 2010. Optimal all-stop, short-turn, and express transit services under heterogeneous demand. Transportation Research Record 2197, 8–18.
Valouxis, C., Housos, E., 2002. Combined bus and driver scheduling. Computers and Operations Research 29, 243–259.
Verbas, I., Frei, C., Mahmassani, H., Chan, R., 2015. Stretching resources: sensitivity of optimal bus frequency allocation to stop-level demand elasticities. Public Transport 7, 1–20.
Verbas, I., Mahmassani, H., 2013. Optimal allocation of service frequencies over transit network routes and time periods: formulation, solution and implementation using bus route patterns. Transportation Research Record 2334, 50–59.
Verbas, I., Mahmassani, H., 2015. Integrated frequency allocation and user assignment in multi-modal transit networks: methodology and application to large-scale urban systems. Transportation Research Record (accepted for publication).
Verma, A., Dhingra, S., 2006. Developing integrated schedules for urban rail and feeder bus operation. Journal of Urban Planning and Development 132, 138–146.
Visentini, M., Borenstein, D., Li, J.Q., Mirchandani, P., 2013. Review of real-time vehicle schedule recovery methods in transportation services. Journal of Scheduling.
Wan, Q., Lo, H., 2003. A mixed integer formulation for multiple-route transit network design. Journal of Mathematical Modelling and Algorithms 2, 299–308.
Wang, H., Shen, J., 2007. Heuristic approaches for solving transit vehicle scheduling problem with route and fueling time constraints. Applied Mathematics and Computation 190, 1237–1249.
Wei, M., Sun, B., Jin, W., 2012. Model and algorithm of regional bus scheduling with grey travel time. Journal of Transportation Systems Engineering and Information Technology 12, 106–112.
Wei, M., Sun, B., Jin, W., 2013. A bi-level programming model for uncertain regional bus scheduling problems. Journal of Transportation Systems Engineering and Information Technology 13, 106–112.
Welding, P., 1964. The instability of a close-interval service. Operational Research Quarterly 8, 133–148.
Wong, R., Yuen, T., Fung, K., Leung, J., 2008. Optimizing timetable synchronization for rail mass transit. Transportation Science 42, 57–69.
Wren, A., Rousseau, J., 1995. Bus driver scheduling – an overview. In: Computer-Aided Transit Scheduling. Lecture Notes in Economics and Mathematical Systems, vol. 430. Springer, Berlin Heidelberg, pp. 173–187.
Wu, J., Song, R., Wang, Y., Chen, F., Li, S., 2014. Modelling the coordinated operation between bus rapid transit and bus. Mathematical Problems in Engineering. Article ID 709389, in press.
Xie, L., Kliewer, N., Suhl, L., 2012. Integrated driver rostering problem in public bus transit. Procedia – Social and Behavioral Sciences 54, 656–665.
Xie, L., Suhl, L., 2014. Cyclic and non-cyclic crew rostering problems in public bus transit. OR Spectrum, 1–38.
Xuan, Y., Argote, J., Daganzo, C., 2011. Dynamic bus holding strategies for schedule reliability: optimal linear control and performance analysis. Transportation Research Part B 45, 1831–1845.

Yan, Y., Liu, Z., Meng, Q., Jiang, Y., 2013. Robust optimization model of bus transit network design with stochastic travel time. Journal of Transportation Engineering 139, 625–634.
Yan, Y., Meng, Q., Wang, S., Guo, X., 2012. Robust optimization model of schedule design for a fixed bus route. Transportation Research Part C 25, 113–121.
Yang, H., Lou, D., 2013. Acyclic real-time traffic signal control based on a genetic algorithm. Cybernetics and Information Technologies 13, 111–123.
Yang, Z., Yu, B., Cheng, C., 2007. A parallel ant colony algorithm for bus network optimization. Computer-Aided Civil and Infrastructure Engineering 22, 44–55.
Yen, J., Birge, J., 2006. A stochastic programming approach to the airline crew scheduling problem. Transportation Science 40, 3–14.
Yoo, G., Kim, D., Chon, K., 2010. Frequency design in urban transit networks with variable demand: model and algorithm. Journal of Civil Engineering 14, 403–411.
Yu, B., Yang, Z., Yao, J., 2010. Genetic algorithm for bus frequency optimization. Journal of Transportation Engineering 136, 576–583.
Yu, B., Yang, Z.Z., Jin, P.H., Wu, S.H., Yao, B.Z., 2012. Transit route network design-maximizing direct and transfer demand density. Transportation Research Part C 22, 58–75.
Zhao, F., Zeng, X., 2008. Optimization of transit route network, vehicle headways, and timetables for large-scale transit networks. European Journal of Operational Research 186, 841–855.
Zhao, J., Bukkapatnam, S., Dessouky, M., 2003. Distributed architecture for real-time coordination of bus holding in transit networks. IEEE Transactions on Intelligent Transportation Systems 4, 43–51.
Zhao, L., 2006. Heuristic method for analyzing driver scheduling problem. IEEE Transactions on Systems, Man, and Cybernetics A 36, 521–531.
Zhoucong, X., Peijia, H., Jing, T., Liping, L., 2013. Transit vehicles intelligent scheduling optimization based on the division of characteristic periods. Procedia - Social and Behavioral Sciences 96, 1502–1512.
Zitzler, E., Laumanns, M., Thiele, L., 2002. SPEA2: improving the strength pareto evolutionary algorithm for multiobjective optimization. In: Giannakoglou, K., Tsahalis, D., Periaux, J., Papailiou, K., Fogarty, T. (Eds.), Evolutionary methods for design, optimisation and control. CIMNE, Barcelona, pp. 95–100.
Zolfaghari, S., Azizi, N., Jabner, M., 2004. A model for holding strategy in public transit systems with real time information. International Journal of Transportation Management 2, 99–110.