Transportation Research Part B 46 (2012) 1202–1217

Contents lists available at SciVerse ScienceDirect

Transportation Research Part B

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

How much can holding and/or limiting boarding improve transit performance?

Felipe Delgado, Juan Carlos Munoz *, Ricardo Giesen

Pontificia Universidad Católica de Chile, Department of Transport Engineering and Logistics, Vicuña Mackenna 4860, Macul, Casilla 306, Correo 22, Santiago, Chile

ARTICLE INFO

Article history:
Received 12 April 2011
Received in revised form 17 April 2012
Accepted 17 April 2012

Keywords:
Bus bunching
Transit operations
Real-time control
Holding
Boarding limits

ABSTRACT

Bus bunching affects transit operations by increasing passenger waiting times and its variability. This work proposes a new mathematical programming model to control vehicles operating on a transit corridor minimizing total delays. The model can handle a heterogeneous fleet of vehicles with different capacities without using binary variables, which make solution times compatible with real-time requirements. Two control policies are studied within a rolling horizon framework: (i) vehicle holding (HRT), which can be applied at any stop and (ii) holding combined with boarding limits (HBLRT), in which the number of boarding passengers at any stop can be limited in order to increase operational speed. Both strategies are evaluated in a simulation environment under different operational conditions. The results show that HBLRT and HRT outperform other benchmark control strategies in all scenarios, with savings of excess waiting time of up to 77% and very low variability in performance. HBLRT shows significant benefits in relation to HRT only under short headway operation and high passenger demand. Moreover, our results suggest implementing boarding limits only when the next arriving vehicle is nearby. Interestingly, in these cases HBLRT not only reduces an extra 6.3% the expected waiting time in comparison with HRT, but also outperforms other control schemes in terms of comfort and reliability to both passengers and operators. To passengers HBLRT provide a more balanced load factor across vehicles yielding a more comfortable experience. To operators the use of boarding limits speed up vehicles reducing the average cycle time and its variability, which is key for a smooth operation at terminals.

1. Introduction

Bus transit services operated without a control system tend to result in vehicle bunching. This phenomenon is produced by two main factors (i) the variability in travel time between stops and (ii) variations in passenger demand. These factors lead to an increase in bus headway variance and a consequent worsening of both the magnitude and variability of average waiting times. Since the user subjective valuation of waiting time is higher than that of any other trip time component (access time, in-vehicle time) (Boardman et al., 2001), the increase in headway variability heavily impacts the level of service perceived by users. This impact gets augmented for highly demanded transit services where vehicle capacities are often exceeded. In these cases passengers waiting at a bus stop might not be able to board the first arriving bus (especially after long

* Corresponding author. Tel.: +56 2 354 7116; fax: +56 2 553 0281.
E-mail addresses: fdb@ing.puc.cl (F. Delgado), jcm@ing.puc.cl (J.C. Munoz), Giesen@ing.puc.cl (R. Giesen).

0191-2615/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.trb.2012.04.005

intervals), and have to wait for the next one. Thus a significant number of passengers suffer long waiting times for a full bus that they cannot board.

Previous research on control strategies (Sun and Hickman, 2008; Eberlein, 1995) have focused mainly on keeping regularity through holding buses at stops if a certain condition is satisfied, which can be very effective, but has the secondary effect of reducing the operational speed. However, an alternative approach to reach more regular headways consists on limiting the number of passengers that can board a delayed bus. This is especially attractive if that bus has a very short headway behind. We have observed this scheme informally operating worldwide when a driver tells the passengers not to board a bus, but board the next one which the passengers can already see approaching, or when passengers preferring not to take a quite loaded bus decide to wait for the next one which they expect to come emptier. This scheme is similar to the ramp metering strategy applied in highways. The main focus of this paper is to understand under which conditions this strategy combined with holding becomes beneficial in comparison with just holding control; how it affects different users in terms not only of the different components of time reductions, but also on regularity, comfort and cycle time reductions. We build on Delgado et al. (2009) in which a combined holding and boarding limits control strategy is proposed, improving the methodology and adapting it in order to handle anew only holding strategy which makes this comparison possible.

The remainder of this paper is divided into six sections. Section 2 discusses the Literature Review. Section 3 introduces the public transit system reflected in the model, describing the characteristics of the bus corridor. Section 4 presents the description of the model, including the state variables of the problem as well as the main assumptions and the notation that will be used for the different variables and parameters. Section 5 sets out the complete formulation of our proposed mathematical programming model, including the objective function and its constraints. Section 6 introduces the simulation experiment, describing the scenarios where the two proposed control strategies (boarding limits combined with holding and only holding) are applied and compares the results obtained with two benchmark strategies of no control and threshold strategy. Finally, Section 7 presents our conclusions, including a summary of the study's main contributions and some final comments on topics for future research.

2. Literature review

Holding control strategies can be grouped into two categories: holding to match a predefined schedules or aiming at regular headways without such a referential framework. A predefined schedule is normally used to serve low demand transit services, which are typical of services with long headways (Ceder, 2001; Furth and Muller, 2007, 2009; Zhao et al., 2006). On the other hand, bus transit systems with high demand and short headways (*e.g.* less than 10 min), such as the one of interest here, are normally operated without predefined schedules. In this case, previous research has proposed simple to implement headway-based threshold rules in which a bus is held at a stop if its preceding headway falls below a given threshold and dispatched immediately otherwise (Barnett, 1974; Turnquist and Blume, 1980; Fu and Yang, 2002).

The appearance of new information and communication technologies, such as GPS and AVL systems, have made possible the development of more complex holding control schemes. The control action of these models are the holding times for each vehicle so as to minimize total passenger waiting time at all stops, or a combination of this factor and in-vehicle delay of passengers due to holding. Table 1 presents a classification of a selection of previous works according to the following characteristics:

(a) Prediction Horizon considered (PH) which can involve a single or multiple events.
(b) Passenger Demand (PD) and vehicle Running Times between stops (RT), which can be deterministic or stochastic in the optimization model.
(c) Overtaking, that can be allowed or forbidden.
(d) The Objective Function to be minimized (OF), that could include waiting time experienced by passengers at stops as they wait for the first bus to arrive (W_{first}), in-vehicle waiting time for passengers aboard a bus being held at a stop

Table 1
Classification of previous work in holding strategies.

Reference	PH	PD and RT	Overtaking	OF	Veh. cap.	Control points	Buses	Sol. method
Ding and Chien (2001)	Multiple	Deterministic	Forbidden	V_h	Ignored	MSC	One	OPT
Eberlein et al. (2001)	Multiple	Deterministic	Forbidden	W_{first}	Ignored	PSS	Multiple	Heuristic
Hickman (2001)	One	Stochastic	Allowed	$W_{first} + W_{in\text{-}veh}$	Ignored	PSS	One	OPT
Zhao et al. (2003)	One	Stochastic	Forbidden	$W_{first} + W_{in\text{-}veh}$	Ignored	MSC	One	Heuristic
Sun and Hickman (2008)	Multiple	Deterministic	Forbidden	$W_{first} + W_{in\text{-}veh}$	Ignored	PMS	Multiple	Heuristic
Zolfaghari et al. (2004)	Multiple	Deterministic	Forbidden	$W_{first} + W_{extra}$	Considered	SSC	Multiple	Meta-heuristic
Puong and Wilson (2008)	Multiple	Deterministic	Forbidden	$W_{first} + W_{in\text{-}veh} + W_{extra}$	Considered	MSC	Multiple	OPT

($W_{in\text{-}veh}$), extra waiting time of passengers that cannot board the first bus (W_{extra}) or the consideration of the minimization of the total variance of headways between buses in all stops (V_h).
(e) Vehicle capacity constraints, which can be considered or ignored.
(f) Control points, which can be a 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).
(g) Number of vehicles considered in the optimization process at an update epoch, which can involve a single or multiple vehicles.
(h) Solution method used, which can be heuristic, metaheuristic, or conventional optimization (OPT).

An adaptive control scheme aiming to provide quasi-regular headways while maintaining as high a commercial speed as possible is proposed by Daganzo (2009) and Daganzo and Pilachowski (2011). In both works stochastic effects in vehicle travel times, stop times and passenger demand are considered. In Daganzo (2009) the control dynamically determines bus holding times at a route's control points based on real-time information about the passage of the previous bus. Although the method proves to be efficient under small disturbances, under large perturbations it reduces performance. To overcome this problem, Daganzo and Pilachowski (2011) continuously adjust bus cruising speed based on a cooperative two way looking strategy based on the spacing of the front and back buses. The cooperative control shows to be effective in preventing bunching. Bartholdi and Eisenstein (2012) abandon the idea of any a priori target headway, allowing the natural headway of the system to spontaneously emerge when disturbances appear by implementing a simple holding rule at a control point. Since the three previous studies aim at maintaining regularity, their proposed approach is not well suited for operations in which vehicles reach their capacities.

As for control strategies to increase vehicle operational speed, Fu et al. (2003) and Sun and Hickman (2005) propose stop-skipping where if a bus skips some stops then the following bus is requested to serve all of them. However, none of them consider vehicle capacity constraints. Cortés et al. (2010) propose a holding and station skipping strategy based on hybrid predictive control. They consider stochastic demand and a short-term horizon in which they aim at minimizing waiting time and include a headway regularity term in their objective function. Holding is implemented over a discrete set of possible values.

In Delgado et al. (2009), we propose a simultaneous control strategy for a transit corridor. The objective of this control strategy is to minimize the total time users devote to their trips, assuming that we have perfect knowledge about the state of the system every time a bus arrives to a stop. We formulated a mathematical programming model to update plans every time a vehicle reaches a stop on a rolling horizon framework. The model considers two simultaneous decisions for all remaining stops on each vehicle route: (i) should the vehicle be held and for how long and (ii) should the number of passengers allowed to board it be limited and by how many. In that work a case study with a high demand scenario is introduced showing the potential of the proposed model. Munoz et al. (2010) provide a performance comparison of Cortés et al. (2010) approach versus Delgado et al. (2009).

The main contribution of that paper, which are preserved in this one, were (i) introducing the boarding limits as a potentially attractive control mechanism that could speed-up the service, (ii) considering holding strategies and boarding limits simultaneously; (iii) incorporating bus capacity constraints without resorting to binary variables that would complicate the solution procedure or increase the computation time; (iv) assigning a continuous variable for the duration of each holding; and (v) including the waiting time experienced by passengers who must await more than one bus due to the capacity constraint in the objective function.

Despite its contributions, Delgado et al. (2009) offered an incomplete approach since it presented the following limitations:

(i) The proposed model could not handle a holding-only strategy if boarding limits was considered unattractive.
(ii) Despite its promising results and its frequent informal operation in high-frequency services worldwide, implementing boarding limits could receive significant resistance from the affected passengers. Unfortunately, no recommendations were given regarding when and how boarding limits should be implemented to guarantee significant benefits and passengers not suffering major delays.
(iii) The impacts reported for the different control strategies neglected comfort and reliability. These are directly affected by the control strategy and have proved extremely relevant for the level of service perceived by the users.
(iv) The model assumes the existence of a dummy bus with unlimited capacity as a boundary condition. This bus was necessary to correctly account for waiting times affecting passengers that arrive at the stop after the last bus departure, but did not analyze correctly the full evolution of the system since a different number of future stops were considered for each bus.
(v) A homogeneous fleet of vehicles is considered.
(vi) Dwell times are governed by boarding only, instead of considering a station-specific combination of boarding and alighting and handle some friction effect between passengers inside the bus and on the platform.

In this paper we overcome these limitations, while we compare an improved version of the strategy Holding and Boarding Limits with Real Time information (HBLRT) introduced in Delgado et al. (2009), with the performance of a new derived model, Holding with Real Time information (HRT), in which the only control policy is vehicle holding at any stop.

3. System characteristics

The system underlying our model is a one-way loop transit corridor with N stops operated by a single high-frequency service consisting of K vehicles each with its own capacity and speed, as shown in Fig. 1. Vehicles start their run at a terminal defined as Stop 1, visiting all stops downstream (2, 3, ..., N) before returning to the same terminal ($N + 1$) where all remaining passengers must alight. The buses are numbered in strict order of advance along the corridor, bus 1 being furthest ahead and K furthest behind.

4. Model description

In this section, we introduce a new formulation for the problem of vehicle Holding and Boarding Limits with Real Time information (HBLRT). This model, also allows us to solve the problem of Holding with Real Time Information (HRT). The HRT differs from the HBLRT in that the former discards the boarding limits strategy, while the latter treats both, holding and boarding limits, simultaneously.

4.1. State variables

Our model is used every time a bus reaches a bus stop to decide how long the arriving bus should be held and how many passengers should be allowed to board it, based on a real-time estimation of the state of the system, i.e. position and number of passengers aboard of each bus, and the number of passengers waiting at the various stops. The state of the system is fully described by the estimation of the following set of state variables:

d_k distance between bus k and its last visited stop (meters). If the bus is still at a stop, then this variable takes a value of zero.
e_k stop immediately upstream from bus k. If bus k is in stop n then $e_k = n - 1$.
$\bar{m}_{ki}$ number of passengers on bus k who boarded at stop i.
c_n number of passengers waiting at stop n.

Passenger demand is estimated based on a given arrival rate and a distribution (proportions) of the destination of these trips for each stop. Note that these two elements are specified separately for each stop originating trips.

4.2. Assumptions

The proposed model considers the following assumptions:

- Buses serve all stops, and overtaking is not permitted.
- Travel times between stops and an estimation of passenger arrival rates to each stop are assumed known and fixed over the period of interest.

Each stop has a dwell time function depending linearly on the number of passengers boarding and alighting, and the number of passengers inside the vehicle that will not alight. However, these stop-specific functions do not change in time.

Note that in contrast with Delgado et al. (2009) no dummy buses are added to represent boundary conditions.

4.3. Notation

The following indices and parameters are used in the model:

k index of buses, $k = 1, \ldots, K$.
n index of stops, $n = 1, \ldots, N + 1$.
t_0 current time, instant when the control decision needs to be made.

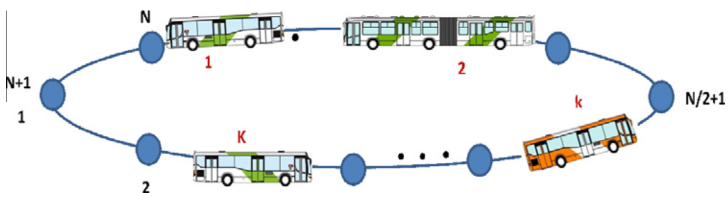

Fig. 1. Transit system model.

θ_i weight factors included in the objective function ($i = 1, 2, 3, 4$).
cap_k passenger capacity for bus k.
λ_n passenger arrival rate at stop n (passengers per minute).
t_b passenger boarding time (minutes per passenger).
t_a passenger alighting time (minutes per passenger).
t_i friction marginal time per passenger inside the bus (minutes per passenger).
r_n distance between stops n and $n + 1$ (meters).
v_{kn} Operating speed for bus k between stops n and $n + 1$ while the bus is moving (meters per minute).
p_{kij} fraction of passengers boarding bus k at stop i whose destination is stop j (for all $i < j$).

The following auxiliary variables regarding the evolution of the system can then be estimated:

mt_{kn} total number of passengers traveling in bus k when it reaches stop n.
m_{kin} number of passengers boarding at stop i and traveling in bus k when it reaches stop n ($\forall i < n$). Notice that $m_{kie_k+1} = \bar{m}_{ki}$.
s_{kn} available capacity in bus k when it reaches stop n (number of passengers).
td_{kn} departure time of bus k from stop n.
a_{kn} number of passengers who alight bus k at stop n.
dp_{kn} number of passengers who want to board bus k at stop n.
b^0_{kn} number of passengers who board bus k at stop n ($n = e_k + 1 \ldots N$).
b_{kn} number of passengers who board bus k at stop n ($n = 1 \ldots e_k$).
f_{kn} dwell time of bus k at stop n (min).

4.4. Decision variables

Two families of decision variables are considered in the HBLRT model:

h_{kn} holding time of bus k at stop n (min).
w_{kn} number of passengers prevented to board bus k at stop n.

For the HRT the w_{kn} are forced to take a zero value if bus capacity is not binding, however when the vehicle capacity is reached it corresponds to the number of passengers that are prevented to board due to capacity constraints.

5. Problem formulation

We now formulate a deterministic mathematical programming problem that simultaneously determines holding times and the number of passengers allowed to board each bus at each following stop assuming that the system will evolve as expected. To take these decisions the model considers the future evolution of the whole system assuming that corrective measures are taken every time a bus reaches a stop. However, since our approach recognizes that the system faces significant operational uncertainties, only the first of these decisions will be implemented; the model will be run again once a new bus reaches a stop.

5.1. Objective Function (OF)

The objective function is to minimize the sum of the individual travel times of all passengers from the moment they arrive at a stop to the moment they reach their destination during the whole planning horizon. In this case the planning horizon consists in each vehicle visiting all stops exactly once. Since vehicle running times (while the vehicles are moving) are assumed to be constant, the objective is equivalent to minimizing waiting times (while vehicles are stopped) both in-vehicle and at-stops. These components can be written in the following way:

(a) At-stop waiting time experienced by passengers as they wait for the first bus to arrive after, t_0, W_{first}, which can be divided in:
 i. Waiting time for passengers at stops where bus k is the first one to arrive,

$$W_{first\_1} = \sum_{k=1}^{K} \sum_{n=e_{k+1}}^{e_{(k-1)}} \left\{ \frac{\lambda_n}{2} \cdot (td_{kn} - t_0)^2 + c_n \cdot (td_{kn} - t_0) \right\}.$$

For the case of bus 1, $e_{(k-1)} = e_K$.
 ii. Waiting time for passengers at a stop n that has already been visited by a bus $(k - 1)$,

$$W_{first\_2} = \sum_{k=2}^{K} \sum_{n=e_{(k-1)}+1}^{e_k} \left\{ \frac{\lambda_n}{2} \cdot (td_{kn} - td_{k-1n})^2 \right\} + \sum_{n=e_K+1}^{e_1} \left\{ \frac{\lambda_n}{2} \cdot (td_{1n} - td_{Kn})^2 \right\}.$$

In this way,

$$W_{first} = W_{first\_1} + W_{first\_2.}$$

(b) In-vehicle waiting time for passengers aboard a bus k being held at stop n,

$$W_{in\text{-}veh} = \sum_{k=1}^{K} \sum_{n=2}^{N} mt_{kn} \cdot h_{kn-1} + \sum_{k=1}^{K} mt_{k1} \cdot h_{kN}.$$

(c) Extra waiting time of passengers who are prevented from boarding bus k because it is at capacity or due to a decision of the controller,

$$W_{extra} = \sum_{k=1}^{K-1} \sum_{n=e_{(k-1)}+1}^{e_k} w_{k-1n} \cdot (td_{kn} - td_{k-1n}) + \sum_{n=e_K+1}^{e_1} w_{Kn} \cdot (td_{1n} - td_{Kn}).$$

(d) Penalty for passengers left behind if there is available capacity,

$$PE = \sum_{k=1}^{K} \sum_{n=2}^{N} w_{kn-1} \cdot s_{kn} + \sum_{k=1}^{K} w_{kN} \cdot s_{k1}.$$

Notice that the arguments of the sums in this penalty take the value zero if for a given bus either the number of passengers prevented to board in station $n - 1$ or the available capacity when it reaches stop n is zero. Thus, the argument is positive only if the bus leaves stop $n - 1$ with available capacity leaving some passengers at the stop.

If we keep the objective function as expressed above (i.e. the total waiting time experienced by all passengers during the planning horizon), we encourage the model to leave more passengers at stops by the end of the planning horizon than those strictly needed to improve performance. This happens because the waiting times experienced by those passengers are not considered in the objective function. We prevent this problem by minimizing instead the average waiting time per passenger in the objective function. Therefore we divide the total waiting time by the total number of passengers involved which is given by:

$$PAX = \sum_{k=1}^{K} \sum_{n=e_k+1}^{e_{(k-1)}} \{\lambda_n \cdot (td_{kn} - t_0) + c_n\} + \sum_{k=2}^{K} \sum_{n=e_{(k-1)}+1}^{e_k} \{\lambda_n \cdot (td_{kn} - td_{k-1n})\} + \sum_{n=e_K+1}^{e_1} \{\lambda_n \cdot (td_{1n} - td_{Kn})\}$$

Thus, we obtain the objective function for the proposed model:

$$\underset{h_{kn}, w_{kn}}{Min} \frac{\theta_1 \cdot W_{first} + \theta_2 \cdot W_{in\text{-}veh} + \theta_3 \cdot W_{extra} + \theta_4 \cdot PE}{PAX} \tag{1}$$

The objective function in (1) is non-linear and not convex. Each of the four waiting-time components are weighted by θ_1, θ_2, θ_3 and θ_4. We observe that a high enough value of θ_4 prevents the system from leaving passengers behind if there is available capacity. Thus, the HRT strategy can be implemented by simply setting a high value for θ_4. On the other hand, if boarding limits is fully accepted by passengers, we set $\theta_4 = \theta$.

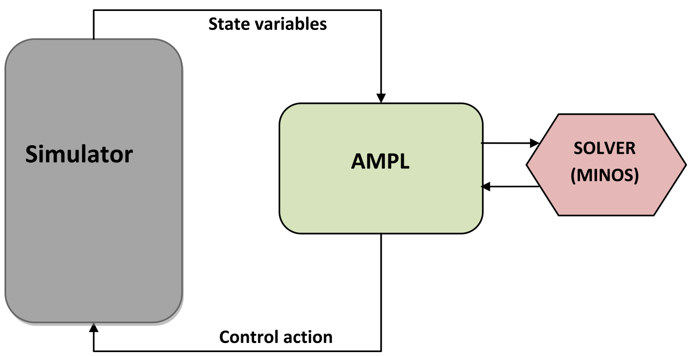

Fig. 2. Interaction between Simulator and optimization solver.

5.2. Constraints

$$td_{kn} = t_0 + \frac{r_{n-1} - d_k}{v_{kn-1}} + f_{kn} + h_{kn} \qquad \forall\ k;\ n = e_k + 1 \tag{2}$$

$$td_{kn} = td_{kn-1} + \frac{r_{n-1}}{v_{kn-1}} + f_{kn} + h_{kn} \qquad \forall\ k, \quad \forall\ n \neq e_k + 1,\ 1 \tag{3}$$

$$td_{k1} = td_{kN} + \frac{r_N}{v_{kN}} + f_{k1} + h_{k1} \qquad \forall\ k \tag{4}$$

$$m_{kin} = b^0_{ki} \cdot \left(1 - \sum_{j=i+1}^{n-1} p_{kij}\right) \qquad \forall\ k;\ n = e_k + 2, \ldots, N; \quad i = 1, 2, \ldots, n-1 \tag{5}$$

$$m_{ki1} = b^0_{ki} \cdot \left(1 - \sum_{j=i+1}^{N} p_{kij}\right) \qquad \forall\ k;\ i = 1, 2, \ldots, N \tag{6}$$

$$m_{kin} = b_{ki} \cdot \left(1 - \sum_{j=i+1}^{n-1} p_{kij}\right) \qquad \forall\ k;\ n = 2, \ldots, e_k; \quad i = 1, 2, \ldots, n-2 \tag{7}$$

$$m_{kn-1n} = b_{kn-1} \qquad \forall\ k;\ n = 2, \ldots, e_k \tag{8}$$

$$mt_{kn} = \sum_{i=1}^{n-1} m_{kin} \qquad \forall\ k;\ n = 2, \ldots, N \tag{9}$$

$$mt_{k1} = \sum_{i=1}^{N} m_{ki1} \qquad \forall\ k \tag{10}$$

$$s_{kn} = cap_k - mt_{kn} \qquad \forall\ k,\ n \tag{11}$$

$$dp_{kn} = c_n + \lambda_n \cdot (td_{kn} - t_0) \qquad \forall\ k;\ n = e_k + 1, \ldots, e_{(k-1)} \tag{12}$$

$$dp_{kn} = w_{k-1n} + \lambda_n \cdot (td_{kn} - td_{k-1n}) \qquad \forall\ k \neq 1;\ n = e_{(k-1)} + 1, \ldots, e_k \tag{13}$$

$$dp_{1n} = w_{Kn} + \lambda_n \cdot (td_{1n} - td_{Kn}) \qquad \forall\ k;\ n = e_K + 1, \ldots, e_1 \tag{14}$$

$$a_{kn} = \sum_{i=1}^{n-1} b^0_{ki} \cdot p_{kin} \qquad \forall\ k;\ n = e_k + 1, \ldots, N \tag{15}$$

$$a_{kn} = \sum_{i=1}^{n-1} b_{ki} \cdot p_{kin} \qquad \forall\ k;\ n = 2, \ldots, e_k \tag{16}$$

$$a_{k1} = \sum_{i=1}^{n-1} b^0_{ki} \cdot p_{kiN+1} \qquad \forall\ k \tag{17}$$

$$w_{kn} \geqslant dp_{kn} - s_{kn} - a_{kn} \qquad \forall\ k,\ n \tag{18}$$

$$w_{kn} \geqslant 0 \qquad \forall\ k,\ n \tag{19}$$

$$b^0_{kn} = dp_{kn} - w_{kn} \qquad \forall\ k;\ n = e_k + 1, \ldots, N \tag{20}$$

$$b_{kn} = dp_{kn} - w_{kn} \qquad \forall\ k;\ n = 1, \ldots, e_k \tag{21}$$

$$f_{kn} = \alpha^b_n \cdot b^0_{kn} \cdot t_b + \alpha^a_n \cdot a_{kn} \cdot t_a + \alpha^{mt}_n \cdot mt_{kn} \cdot t_i \qquad \forall\ k;\ n = e_k + 1, \ldots, N \tag{22}$$

$$f_{kn} = \alpha^b_n \cdot b_{kn} \cdot t_b + \alpha^a_n \cdot a_{kn} \cdot t_a + \alpha^{mt}_n \cdot mt_{kn} \cdot t_i \qquad \forall\ k;\ n = 1, \ldots, e_k \tag{23}$$

$$td_{kn} - td_{k-1n} \geqslant 0 \qquad \forall\ k \neq 1;\ n = e_{(k-1)} + 1, \ldots, e_k \tag{24}$$

$$td_{1n} - td_{Kn} \geqslant 0 \qquad \forall\ n = e_K + 1, \ldots, e_1 \tag{25}$$

$$td_{k-1n} - td_{kn} \geqslant 0 \qquad \forall\ k \neq 1;\ n = e_k + 1, \ldots, e_{(k-1)} \tag{26}$$

$$td_{Kn} - td_{1n} \geqslant 0 \qquad \forall\ n = e_1 + 1, \ldots, e_K \tag{27}$$

Constraints (2)–(4) determine the departure times from downstream stops for each bus, the former associated to the stop immediately downstream from the bus's current location while (3) and (4), to the rest of the stops.

Constraints (5)–(8) establish the number of passengers traveling in bus k when it reaches stop n, that boarded at some previous stop i. In the set of constraints (5)–(7), this quantity is specified as the estimated number of passengers who, having boarded at i, will not alight before stop n. In the case in which i is the stop immediately before n, constraint (8) assign that number to those boarding at this stop. In (9) and (10) the total number of passengers in bus k before it arrives at stop n is stated simply as the sum of those who boarded at previous stops, while (11) relates the available capacity of a bus before arriving at a stop to the total number aboard the bus and its capacity.

Constraints (12)–(14) estimate passenger demand at a given stop for a particular bus. In the case in which a bus is the first to visit a given stop, Constraints (12) indicate that the number of passengers at the stop equals the passengers already waiting there plus the expected number who will arrive before the first bus leaves that stop. Constraints (13) and (14) specifies that for stops already visited by some bus during the planning horizon; this number equals the quantity of passengers who were prevented from boarding the previous bus plus those who arrive at the stop between the previous bus's departure and the arrival of the next one. Constraints (15)–(17) compute the expected number of passengers who will alight from a bus at a certain stop using the estimated probability that each of them will get off at that stop given the stop they boarded.

Constraints (18) and (19) together indicate that the quantity of passengers prevented from boarding at a given stop n must be equal to or greater than the number prevented from boarding bus k because it was at capacity. Thus, passengers can be left behind even if space is available for them on the bus as a way of hastening the vehicle's departure and thereby increasing its operating speed and eventually improve headway regularity. In (20) and (21), the quantity of passengers who are allowed to board bus k at stop n is computed as the difference between potential demand and the number prevented from boarding there. In (22) and (23), dwell time for a bus k at a stop n is a linear combination of boarding, alighting times and the number of passenger inside the bus when it reaches stop n. In this expression α_n^b, α_n^a and α_n^{mt} are stop specific parameters that need to be defined. If at stop n boarding times dominate alighting times, then α_n^b and $\alpha_n^a = 0$; in the case that boarding and alighting are sequential processes then $\alpha_{n=1}^a$ and $\alpha_n^a = 1$. Moreover, if some friction effect between passengers inside the bus and the platform is observed then we choose a value for $\alpha_n^{mt} > 0$. Finally, constraints (24)–(27) establish that buses cannot overtake each other. Notice that the resulting set of constraints (2)–(27) is linear in the variables and all variables are continuous.

6. Simulation experiments

6.1. Simulation environment

The system is simulated using an event based and stochastic simulator. Each event is triggered every time a bus reaches a stop. Bus loads, travel times and position, as well as the number of passengers waiting at each stop are updated in each event. The simulation keeps track of individuals at the stops and in the buses at each point in time.

Even though the optimization model (1)–(27) assumes a deterministic demand and travel times, in the simulation, passengers arrive randomly following a Poisson distribution with mean λ_n at each stop. This distribution is common in service with headways with less than 12 min (Okrent, 1974; Jolliffe and Hutchinson, 1975). The boarding process follows a FIFO discipline. Thus, the earliest arriving passenger will be the first to board a bus if there is available capacity having to wait for the next one if not. For the travel times between two adjacent stops we use a lognormal distribution that has been commonly used in previous studies (Hickman, 2001; Zhao et al., 2003). The mean of this distribution matches the travel time used in the deterministic optimization model.

During each simulation step, a bus could be either moving or dwelling. In the dwelling process, boarding and alighting take place at different doors. In the simulation, each passenger contributes with the same marginal time when boarding and another when alighting. Then, the total dwell time is estimated as the maximum time between these two processes. However, to keep the set of constraints of our optimization model linear, we assume in this set of experiments that dwell times are always governed by boarding (since its marginal effect is larger); i.e. $\alpha_n^b = 1, \alpha_n^a = \alpha_n^{mt} = 0$ for every stop n. This will add some drift to our rolling horizon prediction versus the more realistic simulation.

To determine which action should be taken according to the proposed model, every time a bus reaches a stop, an AMPL file is created containing all the state variables of the system at that time point. This file is the input of the mathematical programming model (1)–(27), which is solved using MINOS. The solution will optimize all decision variables, related to future control actions consisting of holding and boarding limits for all buses, on all future stops until they return to their actual stop in the next cycle. However a single control action affecting only the bus that triggered the event is sent and applied in the Simulator. All other control actions suggested by the optimization model are discarded. This approach is applied in a rolling horizon framework. An illustration of the process is shown in Fig. 2.

6.2. Scenarios

As summarized in Table 2, four scenarios are tested to evaluate and compare the proposed model under different operational conditions: these scenarios differ in two dimensions (i) bus capacities which can be reached or not and (ii) service frequencies, which can be high frequency services, i.e. short headways, or medium frequency services.

Table 2
Simulation Scenarios.

Scenario	Capacity of buses reach	Frequency
1	Yes	High
2	No	High
3	Yes	Medium
4	No	Medium

6.3. Simulated control strategies

We test and compare four different control strategies. The first two (*no control* and *threshold control*) are used for comparison purposes. The third is the control strategy, HBLRT, and the last one, Holding with Real Time Information (HRT) corresponds to the application of HBLRT but using a Lagrangian relaxation approach, in which a high weight on the penalty for passengers left behind is applied whenever there is available capacity, see Eq. (1). Comparing these two last strategies will allow us to understand the impact of implementing boarding limits and the conditions under which it is more beneficial.

Summarizing, we compare the following four control strategies:

- *No control*. That is the spontaneous evolution of the system, where buses are dispatched from the terminal at a designed headway, without taking any control action along the route. Thus the only place where holding can take place is at stop 1.
- *Threshold control*. This is based in a myopic rule of headway regularization between buses, where a bus is held if the headway with the previous bus is less than the schedule headway or is dispatched immediately in other case. As in HBLRT and HRT, we set all the stops along the corridor as control points.
- *HBLRT*, where we do not consider the penalty for passengers left behind ($\theta_4 = 0$).
- *Holding with Real Time Information (HRT)*, in which we set $\theta_4 = 9000$ such that boarding limits are never used, i.e. holding are the only decision variables.

6.4. Simulation analysis

The proposed model is now applied, to an homogenous fleet operating in an imaginary transit corridor of 10 km of length, with 30 bus stops evenly spaced, where the terminal is denoted by stops 1 and 31. Travel times between stops for all buses follow a lognormal distribution with mean 0.77 min. and a coefficient of variation of 0.4. Boarding and alighting time per passenger are set at 2.5 and 1.5 s respectively.. In the medium and high frequency scenarios vehicle's capacity is stated as 150 and 100 passengers respectively. Notice that in scenarios 2 and 4 this capacity is never reached.

For every combination of strategies and scenarios, 30 simulations runs were carried out, each of them representing 2 h of bus operations. We used common random numbers and the same initial conditions, corresponding to buses without passengers aboard and evenly spaced along the corridor. To distinguish the effect of the proposed control strategies under stationary conditions and also under a more chaotic system, for all scenarios we let the system to freely evolve for 15 min before any control strategy is applied. This warm-up period is long enough for some bus bunching to appear.

The weighting factors are set to $\theta_1 = 1, \theta_2 = 0.5$ and $\theta_3 = 2$, to reflect the great annoyance experienced by those passengers that have to wait for more than one bus. In addition waiting time inside the bus is less detrimental than at the stop (Boardman et al., 2001). The value of θ_4 strictly depends in the control actions considered, as established in Section 6.3.

Average waiting times are calculated during the time period when the control strategies take place (minutes 15–120). To isolate the impact of the control strategies we subtract from the average waiting times, the minimum waiting time for the system, which constitutes a fixed cost that cannot be avoided. This lower bound corresponds to half the designed headway multiplied by the number of passengers in the system during the simulation period of interest. Table 3 indicates the average number of passenger per simulation in each scenario as well as the designed headway and the minimum waiting times.

Given that the optimization problem solved for control strategies HBLRT and HRT does not consider the demand uncertainty incorporated in the simulation process we expect that the solver may overreact regarding the control strategies to take. Thus, the control actions proposed may not be optimal. To validate this hypothesis, we compared the performance of the strategy for different reduction levels: instead of applying the control action directly regarding holding (h) and boarding limit (b) we will apply just a fraction of it. Thus, we apply $\alpha \cdot b$ and $\beta \cdot h$ ($0 \leqslant \alpha,\beta \leqslant 1$).

Table 4 shows for Scenario 1 and HBLRT strategy the percentage of reduction in average waiting times compared with no control strategy for different values of α and β. Table 4 indicates that the greater reductions are achieved when $\alpha = 0.5$ and $\beta = 0.5$. While the case when $\alpha = 0.5$ and $\beta = 0.5$ and yields a 61.17% improvement, the case when $\alpha = 1$ and $\beta = 1$ yields only a 32.29%. In addition, the results do not change significantly when α or β takes the values 0.25 or 0.75.

The better results obtained when applying HBLRT with $\alpha = 0.5$ and $\beta = 0.5$ than when $\alpha = 1$ and $\beta = 1$ are consistent for other scenarios as well. This confirms that the optimization model that neglects the stochastic elements tends to overreact. In this case, the overreaction means that the model quite often suggests a holding time greater than needed, forcing the system to speed up the bus later by limiting the boarding once the bus reaches the next stop. In this case, the decision moves from one extreme (holding) to another (boarding limits). Evidence of this can be found in Fig. 3, where we show the percentage of boarding limits applied at each stop to buses that were held in the previous stop. This figure shows that when $\alpha = 1$ and $\beta = 1$ this percentage is much higher than when $\alpha = 0.5$ and $\beta = 0.5$. In this last case, the control strategy provides a more cautious reaction to perturbations, postponing some of the holding to the next stops if needed, taking advantage of updated information. In this manner, the system avoids overreacting to current conditions in case a favorable scenario occurs.

From now on, every time we mention HBLRT or HRT we will be referring to its implementation for $\alpha = 0.5$ and $\beta = 0.5$.

In the next subsections, we report the following performance indicators: (i) the average waiting time and its standard deviations, including its decomposition in $W_{first}, W_{in\text{-}veh}$, and W_{extra}; (ii) bus trajectories; (iii) bus loads; (iv) cycle time distribution; (v) waiting time distributions; (vi) holding time durations; and finally (vii) the computational times required to solve

Table 3
Minimum waiting times for each scenario.

Scenario	Av. number of pax.	Designed headway (min)	Min. waiting time
1	6078	2.00	6078.00
2	3398	1.83	3114.83
3	4011	4.85	9727.56
4	2682	4.25	5699.25

Table 4
Percentage of reduction of HBLRT strategy compared with no control strategy.

Holding (β)	Boarding (α)				
	0	0.25	0.5	0.75	1
0	0	–	–	–	–
0.25	–	−48.09	−59.25	–	–
0.5	−37.57	−60.80	−61.17	−57.92	–
0.75	–	−56.84	−53.80	–	–
1	–	–	–	–	−32.29

the problem. In subsections (ii)–(vii), we focus our analysis in the results obtained for scenario 1 where boarding limits can yield the most benefits.

6.4.1. Average waiting times and standard deviations

Tables 5 and 6 present the results yielded by the four control strategies for scenarios 1 to 4 respectively. In these tables the average value for W_{first}, $W_{in\text{-}veh}$, W_{extra} and the total waiting time with their respective standard deviations are reported. In each case the percentage change with respect to the no-control case are added. We do not report the fourth component of the objective function (penalty for passengers left behind when some capacity is available) since it does not correspond to a waiting time suffered by passengers. The two tables show that both proposed control strategies have very significant impacts over the no control strategy reducing the waiting time due to bunching in around 70%. The tables also show that the strategies present a much more stable performance with standard deviations significantly lower than the no-control and threshold control.

These tables show that boarding limits are almost only implemented in scenarios 1 and 3 in which bus capacity is an active constraint, so in scenarios 2 and 4 both proposed strategies perform quite similarly. These Tables also show that HBLRT only presents a significant reduction over HRT on average waiting times in scenario 1 (high frequency and active capacity constraints) where an extra 6.31% reduction is obtained. However, it also outperforms it in terms of standard deviation with

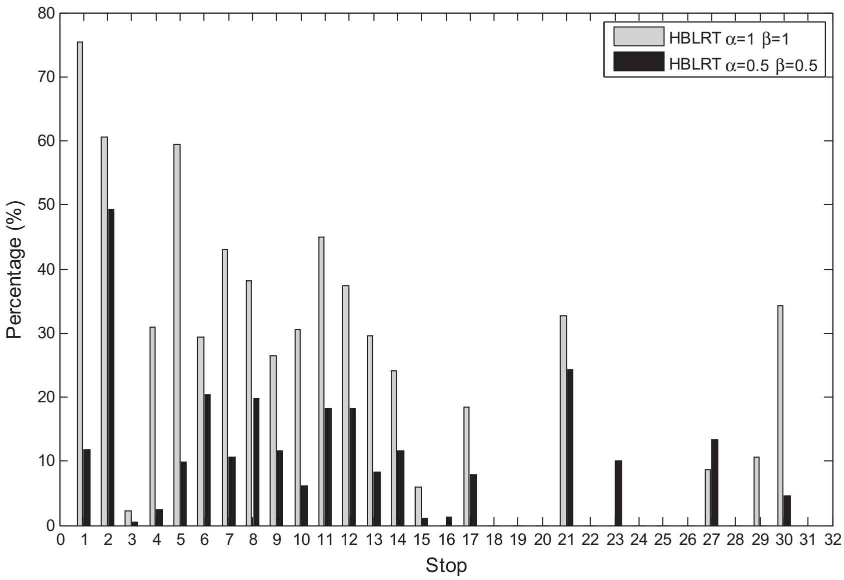

Fig. 3. Percentage of boarding limits applied at each stop to buses that were held in the previous stop.

more significant figures in both scenarios 1 and 3. The tables also indicate that the Threshold strategy performs worst than no control in scenarios 1, 2 (high frequency) and 4 and significantly better in scenarios 3 (but not as good as the proposed strategies). This situation can be explained due to the enormous holding times imposed by threshold strategy in the high frequency services that are more than seven times bigger than in the HBLRT strategy.

It is also interesting to notice that despite boarding limits control action, the HBLRT strategy shows smaller extra waiting times than no control for the capacity binding scenarios 1 and 3 (in which this becomes an issue). The comparison of HRT vs. HBLRT strategies also show that the greater extra waiting times experienced by passengers on the HBLRT strategy are compensated by lower holding times experienced by passengers inside a bus or waiting for their first bus.

6.4.2. Bus trajectories

Fig. 4 shows the bus trajectories for the four different control strategies for a typical simulation run. While in the no control strategy buses bunch up, which leads in turn to long periods of time where no buses pass at a stop, the application of threshold control strategy, as shown in Fig. 4b, avoids some of these bunching, allowing buses to maintain more uniform headways. However, the myopic view of this strategy produces some long holdings that propagate to the following buses at the same stop with significant cost for passengers already in the buses and decreasing the total bus frequency by 19.72% compared with HBLRT. The trajectories for the two proposed control strategies are shown in Fig. 4c and d. As is apparent, both strategies (HRT and HBLRT) produce uniform headway pattern between buses. Notice that in general HBLRT headways are smaller than under HRT, which was expected since in the HRT the only control decision is to hold a bus, while in the HBLRT the possibility of leaving passengers off the bus even when at less than physical capacity, gives more degrees of freedom to the controller who will avoid long holdings specially when bus loads are approaching its capacity. This speeds up some buses, reducing the length of the cycle.

6.4.3. Bus loads

Fig. 5 shows the load of each bus since it departs from stop 1 until it reaches stop 30 for the four different control strategies for a typical simulation run. The horizontal line represents the bus capacity. The figure indicates that HBLRT present the less variability and most uniform pattern in bus loads. In this figure, while under no control, Fig. 5a, loads between buses at a given stop, present a great variability, the application of any of the other three control strategies produce much more uniform loads of buses. In addition, HBLRT and HRT present fewer buses running at capacity and in fewer stops. While under HRT strategy only a couple of buses run at capacity in three stops, under HBLRT strategy, only one bus runs at capacity in only two stops. This is very relevant since discomfort only happens at high load factors so a more balanced load factor across buses yields a more comfortable experience to users. We should also note that a very uncomfortable bus is suffered by much more users than a quite comfortable one. It is interesting that the HBLRT strategy yielded this result even though improving comfort was not considered as a goal. These findings therefore suggest that HBLRT strategy improves comfort compared to the other strategies, allowing buses to travel less crowded and providing a more reliable experience.

6.4.4. Cycle time distribution

Fig. 6 shows the distribution observed for cycle times across all buses for the four different control strategies. The figure shows that HBLRT presents the smallest average cycle time and the lowest variability across the four strategies. While the no control strategy shows an average cycle time of 33 min, the HBLRT strategy drops it to 29.19, a reduction of 11.5%. Similarly, the HRT strategy shows an average cycle time of 30.4, a reduction of 7.9%. The greater reductions in cycle time achieved by HBLRT strategy in comparison with HRT can be explained by the use of boarding limits that speed up buses. In addition, the

Table 5
Objective function value and standard dev. for the four strategies: scenarios 1 and 2.

	Scenario 1				Scenario 2			
	No control	Threshold control	HRT $\beta = 0.5$	HBLRT $\alpha = 0.5;\ \beta = 0.5$	No control	Threshold control	HRT $\beta = 0.5$	HBLRT $\alpha = 0.5;\ \beta = 0.5$
W_{first}	5349.06	1514.73	972.59	513.74	1810.74	399.09	296.21	247.24
Std. dev.	476.53	601.48	255.10	109.51	233.12	60.83	63.86	61.25
% reduction		−71.68	−81.82	−90.40		−77.96	−83.64	−86.35
W_{extra}	1535.20	2147.45	139.25	705.10	21.97	106.70	2.77	43.74
Std. dev.	703.16	3180.44	147.14	60.52	18.92	25.17	3.39	15.81
% reduction		39.88	−90.93	−54.07		385.73	−87.41	99.12
$W_{in\text{-}veh}$	388.40	8127.44	1582.36	1017.08	217.83	3113.45	458.52	420.39
Std. dev.	63.31	1320.70	115.46	68.38	21.50	189.03	44.92	37.50
% reduction		1992.53	307.40	161.86		1329.28	110.49	92.99
Tot	7272.66	11789.62	2694.20	2235.92	2050.54	3619.24	757.50	711.37
Std. dev.	877.59	4906.92	425.08	150.32	229.53	200.05	93.45	89.49
% reduction		62.11	−62.95	−69.26		76.50	−63.06	−65.31

Table 6
Objective function value and standard dev. for the four strategies: scenarios 3 and 4.

	Scenario 3				Scenario 4			
	No control	Threshold control	HRT $\beta = 0.5$	HBLRT $\alpha = 0.5;\ \beta = 0.5$	No control	Threshold control	HRT $\beta = 0.5$	HBLRT $\alpha = 0.5;\ \beta = 0.5$
W_{first}	4237.07	688.78	443.80	52.92	1293.53	290.79	82.00	90.25
Std. dev.	1617.34	165.37	350.87	154.12	632.80	112.51	91.99	72.39
% reduction		−83.74	−89.53	−98.75		−77.52	−93.66	−93.02
W_{extra}	1048.08	223.85	21.82	541.74	17.91	82.21	5.33	28.25
Std. dev.	1071.42	245.71	54.29	149.96	27.63	35.16	6.99	40.79
% reduction		−78.64	−97.92	−48.31		359.08	−70.23	57.73
$W_{in\text{-}veh}$	201.61	2541.78	847.51	667.17	102.58	1308.75	307.21	306.89
Std. dev.	63.71	273.86	150.04	88.96	22.65	122.62	58.27	47.94
% reduction		1160.73	320.37	230.92		1175.90	199.50	199.19
Tot	5486.76	3454.41	1313.13	1261.84	1414.01	1681.75	394.55	425.38
Std. dev.	2668.27	559.56	465.23	284.54	647.84	226.97	140.64	127.23
% reduction		−37.04	−76.07	−77.00		18.93	−72.10	−69.92

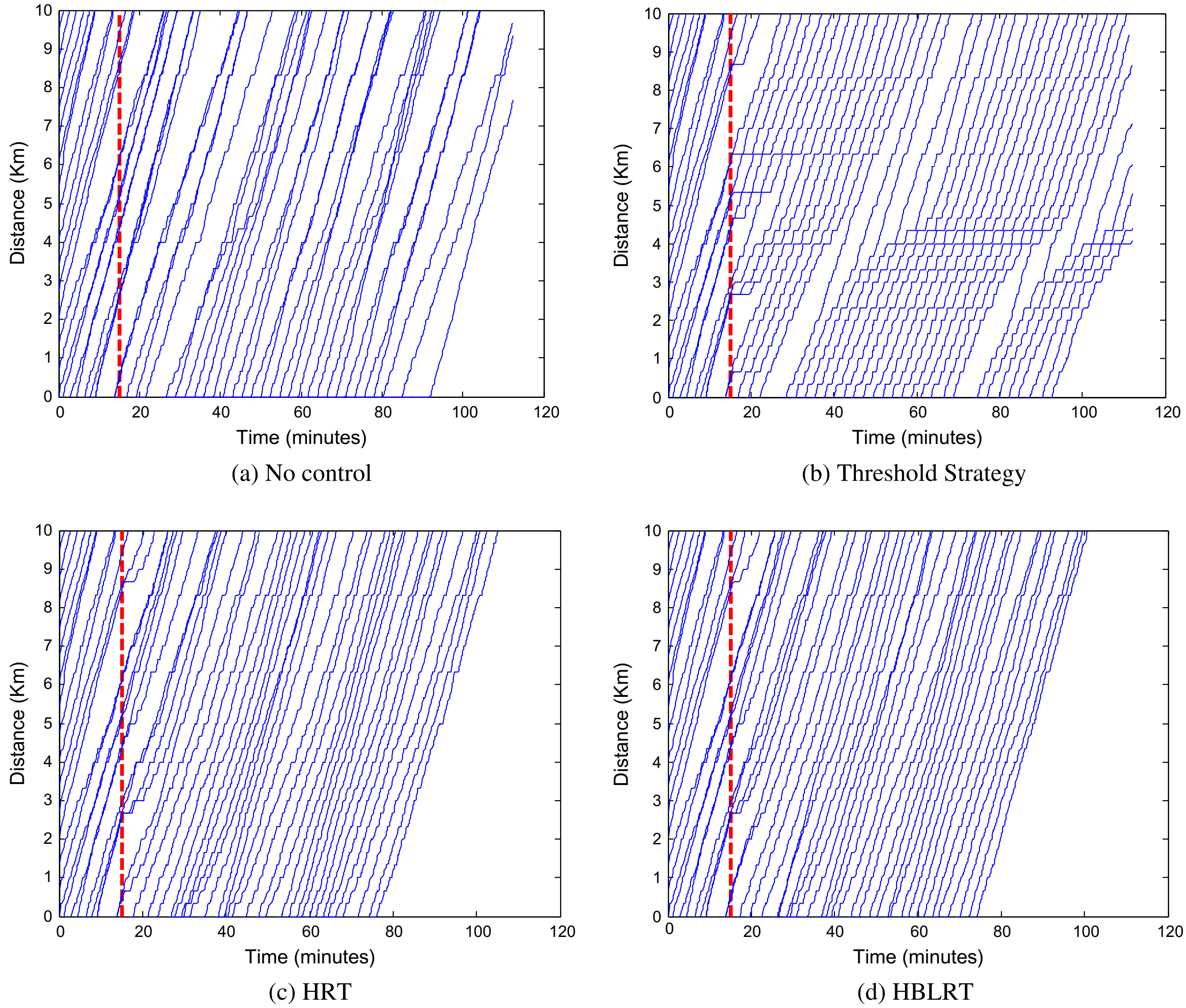

Fig. 4. Trajectories of buses for the different strategies: (a) no control; (b) threshold strategy; (c) HRT; (d) HBLRT.

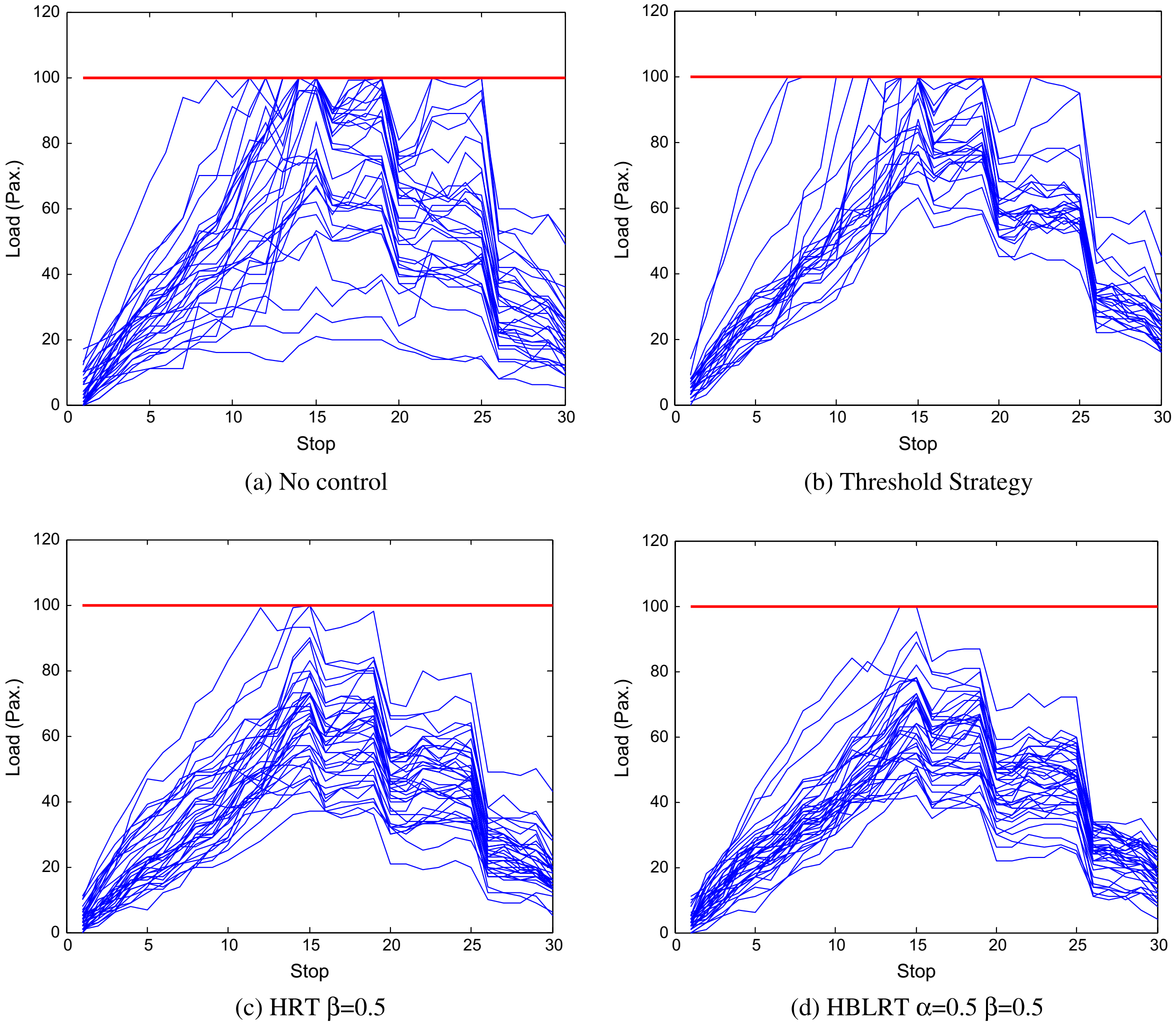

Fig. 5. Bus load at different stops, for different strategies: (a) no control; (b) threshold strategy; (c) HRT; (d) HBLRT.

HBLRT strategy presents the lowest cycle time variability. This data suggests that HBLRT strategy is also the most beneficial strategy from the operator point of view since the low variability allows a smoother and more robust operation and planning at the terminals. Furthermore, the reduction in cycle time also decreases the number of buses needed to provide a given frequency. This reduction in cycle time explains a 11.5% of the reductions of total waiting times; the other 88.5% comes from more regularity.

6.4.5. Waiting times distribution

In order to explore the reliability of the system under the four scenarios, we present a distribution of total waiting times suffered by users at the stops. Table 7 displays the percentage of passengers that wait less than 2 min, between 2 and 4 min and more than 4 min. Since the headway regularity conditions differ significantly between the transient period (15–25) and the stationary period (25–120) both periods are displayed separately in the table.

For the transient period Table 7 shows that applying any sort of control strategy (threshold, HRT, HBLRT) drops the percentage of passengers that have to wait more than 2 min similarly, even though HRT and HBLRT present a significantly better performance for the users suffering a very long waiting. This is particularly noticeable for the HBLRT since its distinguishable control actions consists in leaving passengers off the bus.

For the stationary period, we obtain analogous conclusions but the proposed strategies outperform the threshold even more. The main difference between both periods is that during the stationary state, waiting times are smaller and the percentage of passengers waiting long headways is reduced.

In summary, Table 7 indicates that HBLRT reduces long waiting times for passengers better than the other three strategies, even though boarding limits are applied.

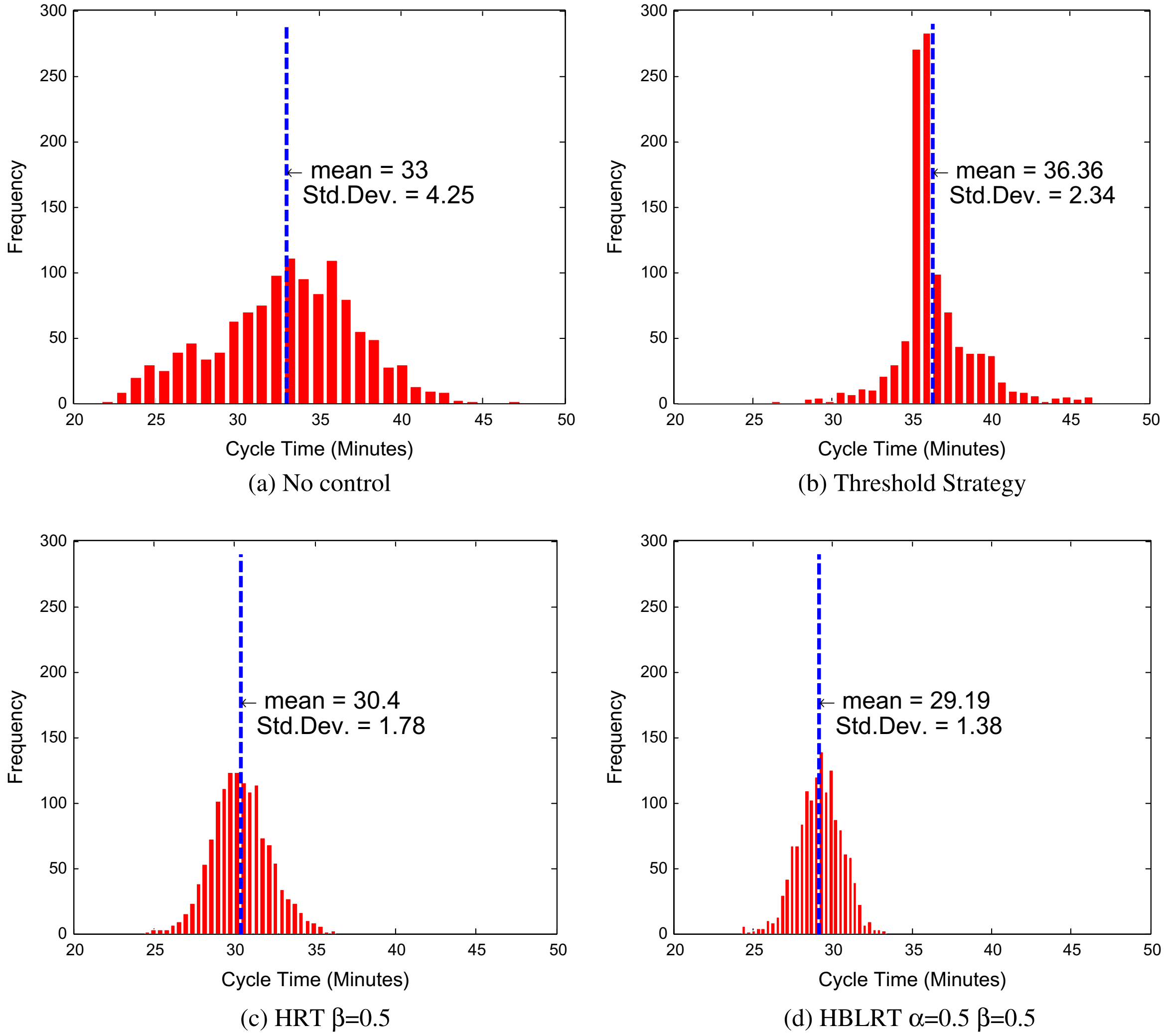

Fig. 6. Cycle time distribution, for different strategies: (a) no control; (b) threshold strategy; (c) HRT; (d) HBLRT.

6.4.6. Frequency of holdings

Finally, we want to understand how long can be the holdings experienced by users. In Fig. 7 we provide a 10-s intervals histogram of holding lengths for the threshold, HRT and HBLRT strategies, for the transient (between minutes 15 and 25) and the stationary (between minutes 25 and 120) periods.

The Figure indicates that, as expected, during the transient period holdings tend to be longer than during the stationary period since control actions during the former are more extreme in order to reach equilibrium. Moreover, HRT requires more holdings than HBLRT.

6.4.7. Computational times

Interestingly, regarding computational requirements, the methodology has the potential to be implemented in real-time applications. We report computational times for scenario one which involves the largest number of variables and constraints

Table 7
Percentage of passengers that have to wait less than 2 min, between 2 and 4 min and more than 4 min, for scenario 1.

	% Of passengers that have to wait between					
	Period 15–25			Period 25–120		
	0–2 min	2–4 min	>4 min	0–2 min	2–4 min	>4 min
No control	50.45	28.11	21.44	61.06	28.00	10.94
Threshold control	72.66	22.81	4.53	79.59	17.96	2.44
HRT $\beta = 0.5$	76.97	21.36	1.67	86.35	13.47	0.18
HBLRT $\alpha = \beta = 0.5$	76.93	22.19	0.88	89.54	10.44	0.02

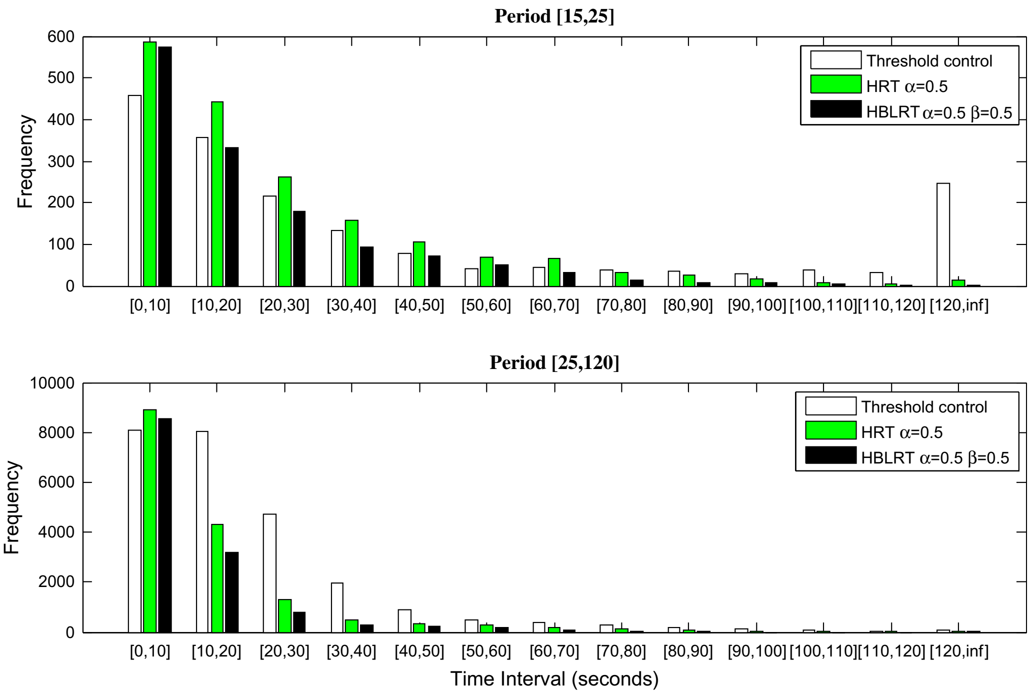

Fig. 7. Distribution of holding time per interval, for periods [15, 25] and]25, 120].

and which was solved using MINOS on a Intel Core2 Duo @ 2.66 GHz. The time taken to solve one instance of the mathematical programming problem lies with 95% of confidence within the following interval:

$$\text{C.I.} = [3.828, 3.864]\ \text{s} \quad (\text{n} = 1403)$$

As we can see, solutions times are in the order of 3.8 s. Furthermore, the maximum solution time observed in this scenario was 5.16 s. The above, highly indicates the very real potential of the model for use in real time operations.

7. Conclusions

The present study was built over a new mathematical programming model of a bus corridor that considers a heterogeneous capacity-limited bus fleet. This new formulation allowed us to consider two different strategies: HRT where holdings are the only control action, and HBLRT where holding and boarding limits are considered simultaneously. In this control scheme, operation plans are updated in real time every time a bus reaches a bus stop.

The proposed model was evaluated in a stochastic simulation environment under different operational conditions that could highlight when boarding limits improve the system's performance the most over conventional bus holding. The results demonstrate that the proposed strategies HBLRT (holding and boarding limits) and HRT (holding) outperform other control schemes in all scenarios. Furthermore, the greatest reductions in the objective function are achieved in those scenarios of high passenger demand where buses reach capacity, with savings reaching 77% and a very low variability in total waiting times in comparison with no control strategy.

HBLRT shows significant benefits in relation to the HRT only in those scenarios of high passenger demand and short bus headway operation. Interestingly, in these cases HBLRT not only reduces an extra 6.31% the expected waiting time in comparison with HRT, but also outperforms other control schemes in terms of comfort and reliability to both passengers and operators. To passengers HBLRT achieve a more balanced load factor across buses yielding a more comfortable experience to users, allowing buses to travel less crowded. To operators the use of boarding limits speed up buses reducing the average cycle time and its variability, which is key for a smooth operation at bus terminals. Furthermore, the reduction in cycle time decreases the number of buses needed to provide a given frequency. Achieving this regularity has usually been considered a major reason to implement schedule-based control by adding some slacks at checking points along the route. This paper shows that we can reach the same goal for high frequency services without slowing down buses.

It is worth noting that our model only suggests implementing boarding limits in cases when the next arriving bus is close. This fact can be observed because only a 0.02% of users suffer waiting times greater than two headways. This percentage is much smaller than the 0.18% of HRT, the 2.44% of Threshold control and the 10.94% of no control. Even more, our results suggest that this strategy should only be implemented for high frequency cases in which these headways are rather short.

We have seen boarding limits happening informally in several cases when (i) a driver asks the passengers in the stop not to board his/her bus, but take instead the next one which is about to arrive or (ii) passengers preferring not to take a heavily loaded bus preferring to wait for the next one which they expect to come emptier. In case that boarding limits is considered infeasible to implement, a more discrete version of the strategy could be tried: if the model suggests limiting the number of passengers boarding then the bus could permit alighting only at the stop or avoid stopping at all if no passenger requests the stop.

Nowadays better infrastructure and communication capabilities could be implemented in services facing these conditions (high frequency and high demand) as a comfortable bus stop, information to users regarding next bus arriving, etc. Under these circumstances implementing boarding limits should be less controversial since passengers would wait less and would be informed about their next bus, and sheltered. This finding also makes us conclude that this control scheme could also be suitable for high-frequency Metros on dense populations. In these systems all information needed for the model is readily available, user-information technologies are usually already implemented and access to platforms is limited to few corridors which makes it feasible to control the number of passengers allowed to reach the platform each time a train leaves a station. The Metro of Sao Paulo has already implemented such a metering strategy to reduce the friction between passengers at the platform and those inside the train by forcing passengers to enter one or sometimes two pre-boarding gates before boarding the train. Even though in this case the boarding rate is defined off-line and is identical for all trains, it appears to be quite effective in reducing dwell times. The control tool introduced in this paper could help them determine the number of passengers allowed to enter each gate by looking in real time not just at a single train and a single station, but instead by looking at the headway regularity effect in the whole line and through a longer time period.

Computational solution time shows to be around 3.8 s in a standard computer for the scenario involving more buses and stops. Thus, from this perspective implementing the methodology in real time operations should be feasible.

Avenues for future research are the inclusion of new control measures such as traffic signal priority mechanisms at intersections where the authors have been already working. Other topics include the study of the robustness of the control strategy under estimation errors.

Acknowledgements

We acknowledge the support of the Chilean Fund for Scientific and Technological Development (FONDECYT) through Project 1120993 and the Bus Rapid Transit Centre of Excellence funded by the Volvo Research and Educational Foundations.

References

Barnett, A., 1974. On controlling randomness in transit operations. Transportation Science 8 (2), 102–116.
Bartholdi III, J.J., Eisenstein, D.D., 2012. A self-coördinating bus route to resist bus bunching. Transportation Research Part B 46 (4), 481–491.
Boardman, A., Greenberg, D., Vining, A., Weimer, D., 2001. Cost-Benefit Analysis. Prentice Hall, E.E.U.U.
Ceder, A., 2001. Bus timetables with even passenger loads as opposed to even headways. Transportation Research Record 1760, 3–9.
Cortés, C.E., Saéz, D., Milla, F., Nuñez, A., Riquelme, M., 2010. Hybrid predictive control for real-time optimization of public transport system's operations based on evolutionary multi-objective optimization. Transportation Research Part C 18 (5), 757–769.
Daganzo, C.F., 2009. A headway-based approach to eliminate bus bunching: systematic analysis and comparisons. Transportation Research Part B 43 (10), 913–921.
Daganzo, C.F., Pilachowski, J., 2011. Reducing bunching with bus-to-bus cooperation. Transportation Research Part B 45 (1), 267–277.
Delgado, F., Munoz, J.C., 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.
Ding, Y., Chien, S., 2001. Improving transit service quality and headway regularity with real-time control. Transportation Research Record 1760, 161–170.
Eberlein, X.-J., 1995. Real Time Control Strategies in Transit Operations: Models and Analysis. PhD thesis. Massachusetts Institute of Technology, Cambridge, MA.
Eberlein, X.-J., Wilson, N.H.M., Bernstein, D., 2001. The holding problem with real-time information available. Transportation Science 35 (1), 1–18.
Fu, L., Yang, X., 2002. Design and implementation of bus-holding control strategies with real-time information. Transportation Research Record 1791, 6–12.
Fu, L., Liu, Q., Calamai, P., 2003. Real-time optimization model for dynamic scheduling of transit operations. Transportation Research Record 1857, 48–55.
Furth, P., Muller, T., 2007. Service reliability and optimal running time schedules. Transportation Research Record 2034, 55–61.
Furth, P., Muller, T., 2009. Optimality conditions for schedules with timepoint holding. Public Transport 1 (2), 87–102.
Hickman, M., 2001. An analytic stochastic model for the transit vehicle holding problem. Transportation Science 35 (3), 215–237.
Jolliffe, J.K., Hutchinson, T.P., 1975. A behavioral explanation of association between bus and passenger arrivals at a bus stop. Transportation Science 9 (3), 248–292.
Munoz, J.C., Giesen, R., Delgado, F., Cipriano, A., Cortés, C., Sáez, D., Valencia F., 2010. Comparison of control strategies for real-time optimization of public transport systems. In: Proceedings of the Triennial Symposium on Transportation Analysis (TRISTAN) VII, Tromsø, Norway.
Okrent, M.M., 1974. Effects of Transit Service Characteristics on Passenger Waiting Time. M.S. thesis. Department of Civil Engineering, Northwestern University, Evanston, Illinois.
Puong, A., Wilson, N.H.M., 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.
Sun, A., Hickman, M., 2005. The real-time stop-skipping problem. Journal of Intelligent Transportation Systems 9 (2), 91–109.
Sun, A., Hickman, M., 2008. The holding problem at multiple holding stations. In: Hickman, M., Mirchandani, P., Voss, S. (Eds.), Computer-Aided Systems in Public Transport, vol. 600. Springer, Berlin/Heidelberg, pp. 339–359.
Turnquist, M., Blume, S., 1980. Evaluating potencial effectiveness of headway control strategies for transit systems. Transportation Research Record 746, 25–29.
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 (1), 43–51.
Zhao, J., Dessouky, M., Bukkapatnam, S., 2006. Optimal slack time for schedule-based transit operations. Transportation Science 40 (4), 529–539.
Zolfaghari, S., Azizi, N., Jaber, M., 2004. A model for holding strategy in public transit systems with real time information. International Journal of Transport Management 2 (2), 99–110.