CASPT 2015

Experimental Evaluation of Real-Time Information Services in Transit Systems from the Perspective of Users

Matías Estrada · Ricardo Giesen · Antonio Mauttone · Emilio Nacelle · Leandro Segura

Abstract We study the influence of real-time information services over the performance of transit systems from the users' perspective. We focus on bus systems and consider services which provide updated arrival time of buses to stops. Six variants of a passenger behaviour model are proposed and implemented, representing different degrees of information availability. To capture the dynamic characteristics of the system, the passenger behaviour model is embedded into a discrete event simulation framework. We perform a comprehensive set of experiments, using a small city with 13 bus lines as a case study. The impact of different assumptions concerning information availability (in particular real-time information) is analysed in terms of user travel time. We test several scenarios, and perform the analysis in terms of both aggregated and non-aggregated measures.

Keywords Real-Time Information · Transit Passenger Behaviour · Discrete Event Simulation

Matías Estrada
Undergraduate Program in Computer Engineering, Universidad de la República, Montevideo, Uruguay
E-mail: tonyhawz@gmail.com

Ricardo Giesen
Department of Transport Engineering and Logistics, Pontificia Universidad Católica de Chile, Santiago, Chile
E-mail: giesen@ing.puc.cl

Antonio Mauttone
Operations Research Department, Universidad de la República, Montevideo, Uruguay
E-mail: mauttone@fing.edu.uy

Emilio Nacelle
Undergraduate Program in Computer Engineering, Universidad de la República, Montevideo, Uruguay
E-mail: enacelle@gmail.com

Leandro Segura
Undergraduate Program in Computer Engineering, Universidad de la República, Montevideo, Uruguay
E-mail: leotuerto@gmail.com

1 Introduction

Nowadays developments on Information and Communication Technologies allow easy access to information about transit operations, which might be changing the way travellers choose services and routes in a network. Several on-line services are currently available for transit users worldwide. These services, which are known as Advanced Public Transportation Information System (APTIS, Coppola and Rosati, 2010) provide information (either static or dynamic) as well as suggested decisions which can be used by transit passengers to improve different aspects of their trips, namely travel time, crowding or a generalized cost. The deployment of APTIS involve large investments, therefore its assessment is a mandatory task for the transit authorities and the government.

We focus on such services which provide real-time information like those implemented in Transantiago[1] and iBus[2], related to transit systems based on buses. These services provide information about the state of the system, usually consisting on updated arrival time of buses to stops, which can be accessed by transit users at any stage of their trip or at specific ones. Currently, different types of users have access to information at different stages previous and/or during their trips. For example, some users might have access to updated information through a computer at the origin, displays at the bus stops, or using mobile devices anywhere. Usually this information is broadcasted from an operation centre, which receives updated positions of the buses and performs estimations based on expected travel times over the network.

Evaluations of the impact of APTIS from observed data can be found in (Breakwood et al., 2014; Watkins et al., 2011). In this work, and with a different approach, we contribute towards the evaluation based on a detailed modelling of the interactions between passengers and buses. Thus, we simulate the transit system based on data related to the services, the demand and hypothesis about the passenger behaviour. The study is focused in the point of view of the users, therefore we left aside the analysis of the implications over the operator’s side, e.g. the fleet management.

1.1 Literature review

Two main approaches can be distinguished in the literature, to evaluate the impact of real-time information over transit users, with goals similar to the one of this study. In one hand, analytical approaches aim to model the system and extract conclusions about it, based on mathematical formulations and properties that can be derived from them. On the other hand, dynamic models representing the evolution of the system as time advances, enable to compute several measures usually based on

[1] http://web.smsbus.cl/web/
[2] http://www.ibus.com.uy/index.html

simulation. In the following, we review the main studies which fall into these categories.

In (Hickman and Wilson, 1995), the authors model a single corridor including issues related to information content (expected departure time and expected running time of buses), place of information (in the terminal and on-board the vehicle) and information accuracy (low and high level of accuracy in projecting travel times). These aspects of the problem are represented by different components of a probabilistic path choice model which is used to simulate a case comprising a network of five nodes, corresponding to part of the transit system of Boston. Different scenarios concerning the issues stated above are simulated, and results are evaluated in terms of several independent replications of the experiment. In general terms, observed gains in travel time are not greater than 3% when real-time information is available, which is considered by the authors as a modest improvement. The results are statistically significant. The authors recognize the difficulty of arriving to general conclusions based on experiments with a single case, and they suggest other scenarios where the usefulness of real-time information should be evaluated, namely, schedule of departure from origin and reduction of passengers' anxiety associated with the uncertainty of transit travel.

The study of Gentile et al. (2005) proposes an analytical framework which generalizes the assignment model under the presence of common lines (Chriqui and Robillard, 1975). The authors develop a formulation of the stop model, assuming that passengers do not take the first bus which lead to destination; instead, when real-time information is available, they can choose a different line that is going to pass by the stop. A numerical example is presented, based in the case of Sioux Falls, which comprises 24 nodes and 76 arcs. The experiments aims to illustrate the effects of real-time information availability and service regularity over the network loading and the passenger's travel time. Improvement in total travel time when real-time information is available is less than 1%. Moreover, the authors found that the impact of service regularity and the availability of real-time information is less relevant for short distance trips.

In (Coppola and Rosati, 2010), the authors perform an evaluation of APTIS based in a simulation framework which comprises three main components: (a) the network performance sub-model, which reproduces the travel time of the buses over the links, (b) the operation control centre, which predicts arrival times and occupancy levels of buses, and (c) the path choice sub-model, which represents passengers' decisions based on a random utility model. The study is done over a case related to the city of Naples, which comprises 11 zones and 9 lines. The resulting network, after exploding the underlying diachronic graph (a specific structure used to represent the temporal characteristics of the transit service) has about 38000 nodes and 75000 links. The experiments simulate 2 hours corresponding to the morning peak hour. Six scenarios are studied, considering different conditions regarding service irregularity, information on waiting time and information about bus occupancy. The results indicate that savings up to 12.5% can be obtained when real-time information is available. In general, waiting time is increased, while on-board travel time contributes to decrease the total time.

In (Cats, 2011), a random utility model is proposed for representing passenger decisions at any stage of the trip. The methodology is applied in (Cats et al., 2011) to a case study about the metro of Stockholm, which comprises 7 lines, 210 platforms and 100 stations. The experiments consider different levels of provision of real-time information (platform, stop, and network) and also different operational conditions (the special case of service disruptions is studied). The simulation consists of 10 independent replications of a period of 3 hours. Conclusions are that path shifting and time savings up to 11% may be obtained by providing real-time information to the transit users. In particular, that information appears to be very useful in cases of service disruptions. The authors recognize the needs for validation with a system-wide case study and real-world data.

More recently, Chen and Nie (2015) developed a model aimed to study the influence of partial information over transit users. The term partial information refers to the fact that real-time information can be available only for a subset of lines from the whole system, thus generalizing the assumptions of Gentile et al. (2005). Since partial information is provided, the best passenger path to reach destination may include several alternative lines, which leads to the need for computing optimal hyperpaths (Nguyen and Pallotino, 1989) under this scenario. Therefore, the authors propose an algorithm for such calculation. Moreover, they provide a sufficient condition to exclude cycles, which may appear under partial availability of information. Numerical results are reported using both small and real sized instances, in particular the bus network of Chicago which has 125 lines. Main conclusions are that real-time information contributes to attract more users to faster lines with lower service frequency and it is more effective in faster lines. The authors also conclude that benefits of real-time information in reducing the total travel time are very small in the particular case study. Nevertheless, they acknowledge that potential gains due to scheduling departures should be investigated, since this feature is not included in their model.

In general terms, analytical models showed to be restrictive to represent several real-time characteristics of the systems. Also, classical assignment models (Spiess and Florian, 1989) are difficult to adapt in order to represent the dynamic nature of passengers' decisions when they face real-time information. Therefore, mesoscopic models have proved to be the most suitable alternative to be used in this context. These models usually combine the discrete event simulation paradigm with behavioural models like those based on random utility theory or optimal strategies.

1.2 Motivation and statement of contribution

The motivation of this study is twofold: (a) evaluation of the effect of real-time information services over transit users in small cities of our region, and (b) using state of the art methodologies concerning transit system modelling and simulation.

The former leads to consider particular characteristics with respect to both transit service and users. We focus in systems which operate a moderate number of lines, not very overlapped, with relatively low frequencies. Congestion is not usual, neither at the road level nor at the bus level (i.e. capacity is sufficient to accommodate the demand). From the passenger point of view, even though the system operates with low frequency, users do not consider timetable information, mainly because it is not published. Therefore, the usual passenger behaviour can be considered as frequency-based (Nuzzolo, 2003) and (due to the size of the city) transfers between lines are not usual. Moreover, we consider the fact that new investments on infrastructure (e.g., provision of real-time information) are under consideration. The potential benefits these investments may introduce, in some cases require that transit users have access to technology, like mobile devices; but different degrees of technology availability is present across the inhabitants of the cities. Therefore, a rational assessment of the gains that can be obtained by providing real-time information under this scenario, should be done.

The modelling of the scenario under consideration should be as realistic as possible. Simulation is the most suitable alternative to model the dynamic characteristics of transit systems with real-time information, since classical assignment models (e.g. Spiess and Florian, 1989) assume steady-state operations. The representation of passenger behaviour is the critical component of the whole model. The existing passenger behaviour models have been successful in terms of realism, at the expense of high computational and calibration requirements. Some of them do not model relevant issues of our scenario, like decisions regarding departure from origin and selection of origin stop. Moreover, due to high computational requirements, several independent runs are difficult to perform, which precludes a statistical analysis of results (Law, 2006). Finally, some authors recognize the need for an analysis at a system-wide level, using real data. Transit systems have a complex structure comprising several trip patterns and several bus lines, which entails complex interactions. Various output measures should be computed and compared in order to obtain a valid assessment

The contribution of this work is a comprehensive experimental evaluation of the influence of real-time information services over the transit system performance from the users' point of view, using state of the art and sound methodologies. We propose six variants of a base passenger behaviour model, which represents different situations of information availability at different stages of travel. The proposed model allows for an efficient implementation and does not require calibration. The experiments are carried out with a test case relative to a small city for which real data is available, particularly those corresponding to transit demand. Conclusions are drawn, paying attention to the magnitude of improvements as a consequence of information availability. Both aggregated and non-aggregated measures are analysed in order to better support the conclusions.

The rest of the article is structured as follows. Section 2 describes the model proposed, including the six variants of passenger behaviour under different conditions of information availability. Section 3 describes the experiments, including the planning, results and specific conclusions. In Section 4 we formulate general conclusions and we state future work. Finally, Appendix A provides a detailed description of the passenger behaviour models proposed.

2 Modelling approach

Our model for evaluating the impact of different degrees of information availability over the performance of the transit system from user's perspective, has three main components which deserve detailed explanation: (a) the transit system representation, that includes the network of lines along with the characteristics of the services and the demand, (b) the passenger behaviour model, which represents the steps followed by passengers to reach their destinations from their origins, and (c) the discrete event simulation model, which performs a dynamic interaction of both sub-models (a) and (b). In the following, we explain in detail each one of these sub-models.

2.1 Transit system representation

We represent the bus network by coding the lines over the real street network, whose segments include street direction and an estimation of the mean travel time. This value is used as parameter for a normal probability distribution, in order to model variations in bus travel time due to different factors like driver characteristics or traffic conditions. This is one important source of randomness of our model, and it is the feature that affects passenger behaviour and system performance under different degrees of information availability. The bus stops are modelled explicitly and they are connected through walk arcs to zone centroids which represent origin and destination places. Each line can have either forward and backward directions (not necessarily using the same streets) or a single circular direction. Moreover, each line has a frequency value (or its inverse, the headway), a timetable according to its frequency, and its sequence of stops. The timetable states the arrival time of buses to every stop of the line within the modelling time horizon, based on a given initial time (first bus departure) and the fixed travel time of each line segment. Another relevant component of the transit system is the demand. In our case, it is represented as an origin-destination (OD) matrix, which expresses rates of individual trips from origins to destination. The demand is represented at the level of centroids and each value is fixed within the whole time horizon. Each element of this matrix is called OD-pair.

2.2 Passenger behaviour model

This sub-model is a critical part of the whole model, since encodes the way in which passengers interact with the transit service. Several figures of system performance (in particular, travel time) are affected by decisions taken by users, therefore a realistic modelling of those decisions is mandatory. As stated in Section 1, static assignment models do not allow for representing dynamic characteristics of the interaction between passengers and buses, like those which are present when real-time information is available. Therefore, dynamic assignment models arise as the most suitable alternative. Among them, we consider the so-called schedule-based transit assignment models, which take as input a detailed representation of the service (the timetable) and the demand, potentially variable in time. A relevant concept in this context is the bus run, which refers to the specific departure of a bus to perform the service of a given line at a given time. According to (Nuzzolo and Crisalli, 2004), to model transit services at the run level, there are three types of system representations. The diachronic graph (Nuzzolo and Russo, 1994) is a highly structured model which comprises a service sub-graph, a demand sub-graph and an access/egress sub-graph. The dual graph representation (Anez et al., 1996) includes dual nodes which represent the runs, and links which model time congruence relative to arrival/departure times of buses at stops. At last, the mixed line-based/database approach (Tong and Richardson, 1984) combines a single topological line representation with temporal information relative to bus movements over the network. Main differences among these models are: (a) the management of the trade-off between complexity of computer implementation and efficiency, and (b) support for representing the passenger behaviour, in particular the path choice. In the context of schedule-based assignment models for transit systems, a path between origin and destination centroids is defined (Nuzzolo and Crisalli, 2004) by the sequence of physical network nodes (including centroids and stops) as well as arrival/departure times from/to them. A standard assumption in the literature is that passengers always seek to minimize a measure (or conversely, maximize a utility) when choosing the way for travelling from origin to destination using a given set of transit lines. In this context, decisions may refer to departure time from origin, initial bus stop, line (or set of lines) to take, and so on. Within the shedule-based approach, a common assumption is that passengers think in terms of single paths. On the other hand, in the frequency-based approach (where passenger do not know the line timetables), a typical assumption is that passengers think in terms of strategies (Spiess and Florian, 1989). In order to choose a single path, passengers are assumed to apply shortest path algorithms over a network which may change its attributes (e.g., cost) as time advances. Moreover, in order to model different perceptions of such attributes, a common accepted methodology is the Random Utility Theory (Domencich and McFadden, 1975).

In this work, we propose a simple model for both service representation and passenger behaviour, mainly related to the line-based/database approach, and more specifically based on all-or-nothing assignment with dynamic rescheduling. This means that passengers always take decisions about a single path and such decisions are updated as the time advances, which entails that travel stages are performed (walk, wait, board, etc.) and information about the state of the system is updated. Under this general model we instantiate six variants which represent different assumptions, most of them conditioned by different degrees of information availability concerning the transit system. In the context of this study and in the light of its goals, the proposed model and its variants (also called models in the remaining part of the paper) have the following advantages:

- Simplicity, which entails they are easy to understand and to validate.
- Suitable for modelling the characteristics of our interest, where the effective timetables do not differ significantly from the nominal ones. Also, the transit system is assumed to operate without congestion in terms of bus capacity.
- Consistent with the passenger behaviour under consideration, where users do not use naturally information about timetables, even under the presence of medium-to-low frequencies.
- Since the model is simple, it is also easy to implement and to extend in order to include additional variants concerning passenger behaviour.
- Because the model allows for an efficient implementation, this enables to make several experiments, which is needed in order to have statistically significant results.

In the following, we explain the six proposed models in general terms; Appendix A gives a more detailed description.

1. Real-time information at any stage (**RTI-allways**): Users with real-time information about all the lines of the system at any stage of their trip. This represents users that have a mobile device, which can be used anywhere.
2. Real-time information only at the origin (**RTI@origin**): Users that can access real-time information about the system only at the origin. This represents users that access the information exclusively through the computer at their home/office.
3. Real-time information of a single line (**RTI-1Line**): Users with real-time information of a single line at any stage of their trip. This represents some services where users have to pay for obtaining real-time information about a given line.
4. Static timetable (**STT**): No real-time information is available. Users choose lines using a static timetable and no rescheduling is performed.
5. Real-time information only at the bus stop (**RTI@stops**): Users with real-time information only at the bus stops, through screens provided by the transit system infrastructure.
6. Frequency-based approach (**FBA**): No timetables are available. Users take decisions using a frequency-based approach.

Moreover, in models 1 to 4 the users schedule their departure from origin using pre-trip information, while in models 5 and 6 users begin to walk to the bus stop when they appear in the origin centroid. Note that model 1 represents a somehow sophisticated variant of passenger behaviour, while model 6 represents the most uninformed passenger. It is worth noting that none of these models consider transfers, i.e. users travel from origin to destination using a single line; this greatly simplifies the models and the analysis of results.

2.3 Discrete event simulation model

In order to model the dynamic characteristics of the system, the lifecycles of buses and passengers are embedded into a discrete event simulation framework (Tocher, 1963). The simulation schedules and executes the events according to the times stated by the problem parameters (line network, demand) and associated probability distributions. To do that, several events are designed, either bounded (its time of occurrence can be predicted) or conditioned (the execution depends on a particular condition). In the following, we describe roughly those lifecycles:

- **Bus**: For each line given by its stop sequence and timetable, the model schedules the starting of each bus run (event StartBusRun). We are not concerned with fleet management issues in this work. Buses appear at the initial node of the route, at the time indicated by the corresponding run; when the bus finishes its route, simply disappears in the final node. The event BusArrivalToStop executes the corresponding interaction with passengers and schedules the arrival to the next stop according to a random value determined by the mean travel time of the corresponding link and the normal distribution.
- **Passenger**: For each OD-pair, initially the first passenger arrival to the origin is scheduled by means of the event PassengerArrivalToOrigin. When this event is executed for each OD-pair, the next arrival is scheduled using a negative exponential distribution which rate is the value stated by the OD-matrix; this is other source of randomness of our model. Once at the origin, the passenger plan her/his trip and begin to walk to the origin stop (event PassengerDepartsFromOrigin) or waits some time period, depending on the specific behaviour model. Once at the stop (event PassengerArrivalToStop), the passenger executes her/his plan which may entail a change of decision for each bus that passes by the stop. For passengers already on-board, if the current stop is destination, they alight and immediately begin to walk till destination centroid.

2.4 Computer implementation

The model was coded in C++, using the EOSimulator library[3]. A relevant aspect of the model is the representation of the dissemination of the real-time information. In real systems, typically the location of each bus is reported to a central planning unit, which uses the information to predict arrival times to subsequent stops. These predictions are then broadcasted to users. In this context, the accuracy of predicted arrival times depends (among other factors) on the frequency in which the information is reported by the buses. Analogous, as the predicted information is more recent, the more accurate will be the data available to users. This aspect of the real system is modelled in (Coppola and Rosati, 2010) by an Operation Control

[3] http://www.fing.edu.uy/inco/cursos/simulacion/eosim_html/index.html

Center, which represents the central unit that gathers information, performs estimations and broadcasts to users. In our model we use a simpler approach: once a value of travel time along a network link is sampled, the corresponding dynamic timetables are immediately updated. These timetables are available to all users, in the models where real-time information is considered to take decisions. This greatly simplifies the model and its implementation. Its main consequence is that numerical results corresponding to users' travel time are underestimated in the models which consider real-time information.

The travel time is recorded by using the Histogram feature of EOSimulator. Several histograms accumulate data from every passenger generated by the simulation, discriminating the travel time by its components walk from origin, wait at the stop, travel on-board and walk to destination. Different independent executions of the simulation can be performed by changing the random seed. The parameters which are affected by the seed are the bus travel time along the street and the inter-arrival time of passengers to the origin centroids.

3 Simulation experiments and results

3.1 Methodology and goals

We test the simulation model over a case related to Rivera, Uruguay, a small city with 65,000 inhabitants approximately. Its public transportation system has 13 lines, some of them with overlapping segments (see Fig. 1). The model comprises 84 zone centroids and 378 OD-pairs, which represent the transit demand within a time horizon of 12 hours. Line headways range from 20 to 60 minutes. The complete model has 522 nodes and 1528 arcs. A single execution of the model (6 hours of simulation time) takes 18 seconds (average) in a Core i7 2.4 GHz computer.

The goals of the experiments are the following:

– Evaluation of the transit system's performance under the hypothesis of the six passenger behaviour models explained in Section 2.2. We analyse the results in terms of travel time, using both aggregated and non-aggregated measures. The former provides a framework for comparison among different models in terms of a single and specific value, namely, mean travel time. The latter allows for a detailed analysis, enabling to discover potential unseen facts, and therefore it complements the aggregated approach. The detailed analysis of results is done in two directions: (a) different components of the travel time, namely, walk to origin, wait at the stop, on-board the bus and walk to destination, and (b) different OD-pairs.
– Sensitivity of the model to changes in service characteristics and system conditions. In particular, we run the model using the same case, but the frequencies and the service irregularity are increased. The goal of these specific experiments is to study whether results and conclusions of the previous one also hold under these conditions.

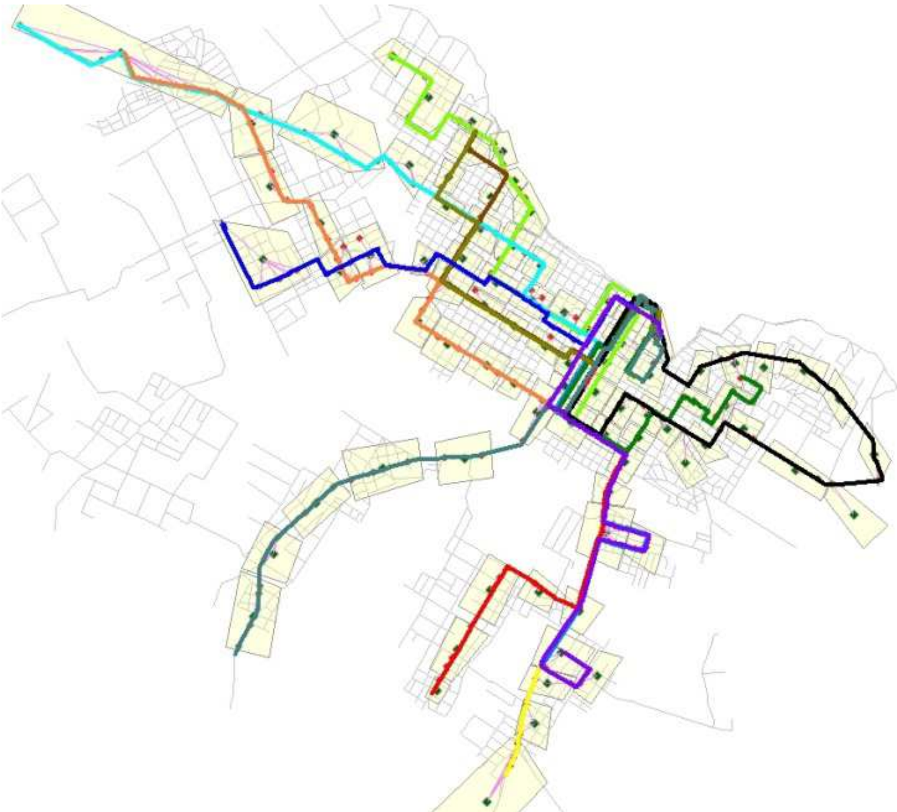

Fig. 1 Case of Rivera

Given that our model has a stochastic nature (given by the inter-arrival times of passengers to origins and the travel times of buses along the street segments), the outputs are samples of (unknown) probability distributions. Therefore, for each experiment we run 100 independent executions, which produce results that are summarized by their mean and standard deviation. These values are used to build confidence intervals of 95% level which are finally used for comparisons. Note that more elaborated methodologies for comparisons have been proposed in the simulation area, either for pairwise as for multiple configurations (Law, 2006).

3.2 Results from current system

Table 1 shows mean values corresponding to travel time averaged over the 100 independent executions. The third column shows the size of the half confidence interval built as explained in Section 3.1. Note that travel time of each independent execution is averaged over all the passengers of the simulation. Therefore, it is a significantly aggregate measure because it sums up very different values corresponding to different OD-pairs. Small confidence intervals may be also explained by this fact and by the high number of independent replications. We can observe that travel times range from 43 to 63 minutes, which is a reasonable value, taking into account the size of the city, the walking distances given the zonal division, and the headways of the transit system.

Table 1 Mean total travel time and confidence interval for the six models

Model	Mean (secs.)	Half conf. interval
1. RTI-Allways	2589.03	3.03
2. RTI@origin	2612.59	2.99
3. RTI-1line	2625.36	3.20
4. STT	2693.29	3.30
5. RTI@stops	2960.34	3.35
6. FBA	3778.66	5.11

As we may expect under this low frequency scenario, saving in travel time can be obtained using real-time information. In particular, models RTI-allways, RTI@origin and RTI-1Line exhibit very similar results. Model STT causes a small travel time increase, and finally, models RTI@stops and FBA show an even higher increase, particularly the last one. Note that even though model RTI@stops considers real-time information, since users do not schedule their departure, the result is worse than the one of model STT.

Figure 2 compares average results disaggregated by stage of travel. We can observe that again, results from models RTI-allways, RTI@origin and RTI-1Line are very similar, even at the different stages. The main saving with respect to other models is in terms of waiting time, since in these models users schedule their arrival based on real-time pre-trip information. Model STT presents a slight increase in waiting time, since users schedule their departure based on the static timetable. Model RTI@stops increases even more the waiting time, since users do not schedule their departure. Real-time information only at the stop does not seem to be useful in this case; this is probably due to the low frequencies. Finally, users that plan their trips using a frequency-based approach, experience a similar waiting time as in model RTI@stops, but they have a higher on-board travel time since they board the first bus that reaches their destination.

In order to compare the models from another non-aggregated point of view, we selected five OD-pairs with different characteristics, namely, geographic distance between origin and destination, and service availability (lines, frequencies). Figure 3 plots for each of these pairs (identified by their origin and destination) mean total travel time of the six passenger behaviour models. We can observe that the tendency already observed in Table 1, also holds for different OD-pairs. As we may expect, larger differences in the graphic shapes can be observed between the first and the last OD-pair, which in fact are the ones having the minimum and maximum distance between origin and destination respectively.

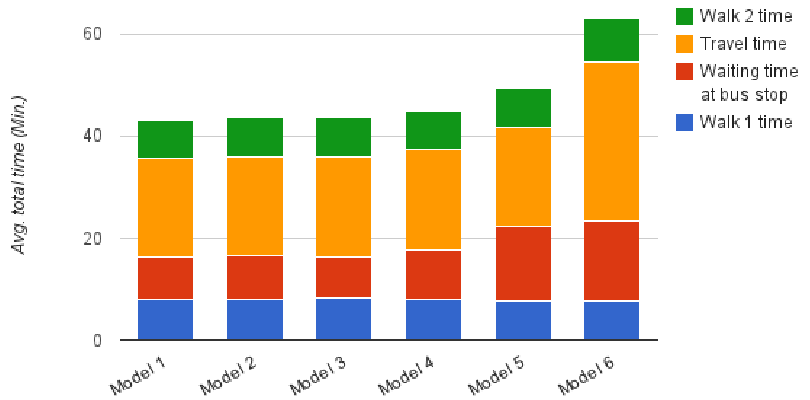

Fig. 2 Average total travel time by stage of travel

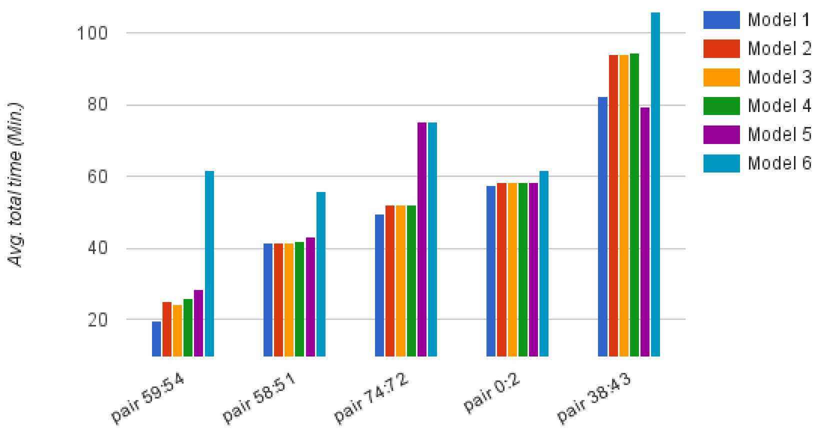

Fig. 3 Average total travel time for different OD-pairs

Despite the results observed above and the conclusions which may be drawn from them, we also investigate the waiting time histogram of the “extreme” models RTI-allways and FBA and the “intermediate” model STT (see Fig. 4). The waiting at the stop is usually considered as the most onerous component of the total travel cost and it is also the most influenced by real-time information. Based on these results, we can verify that users with better information on average experience lower waiting times than less informed users, i.e. smaller waiting time values are experienced by much less users (in the order of one third approximately). Moreover, the waiting time experienced by users which have static timetable information is very similar to the best results obtained by model RTI-allways. This somehow suggests that by using the static timetable information only, the improvements with respect to the frequency-based behaviour is greatly improved. Also, the magnitude of that improvement is larger than the one that can be obtained by using real-time information with respect to the model of static timetable. It is worth mentioning that this observation holds for a scenario of high regularity and low frequencies.

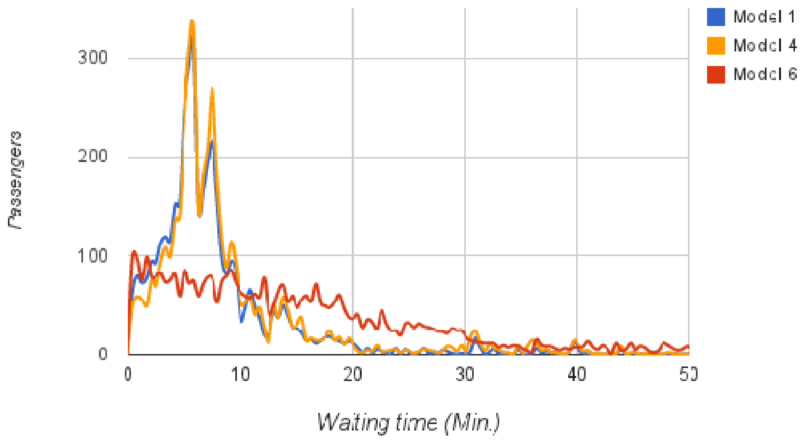

Fig. 4 Waiting time histograms for different passenger behaviour models

3.3 Higher frequencies

With the aim of investigating the differences among the six passenger behaviour models in a scenario of higher frequencies, we multiplied by four the original ones of the case of Rivera, thus obtaining values ranging from 5 to 15 minutes. As we can observe in Table 2, results from all models but FBA are very similar, in fact, confidence intervals are overlapped. Models with real-time information (RTI) do not improve user's travel time, when compared with the model based on static timetable (STT). Model FBA obtains significantly worse results.

Figure 5 shows that very similar results are obtained even at the level of stage of travel. This strongly suggests that under a scenario of high regularity and high frequencies, the main factor which influences the improvement in travel time is the availability of timetables, either dynamic (updated in real-time) or static.

Table 2 Mean total travel time and confidence interval for higher frequencies

Model	Mean (secs.)	Half conf. interval
1. RTI-Allways	2389.26	2.54
2. RTI@origin	2393.26	2.46
3. RTI-1line	2399.17	2.65
4. STT	2439.69	2.51
5. RTI@stops	2423.92	2.41
6. FBA	3135.37	3.62

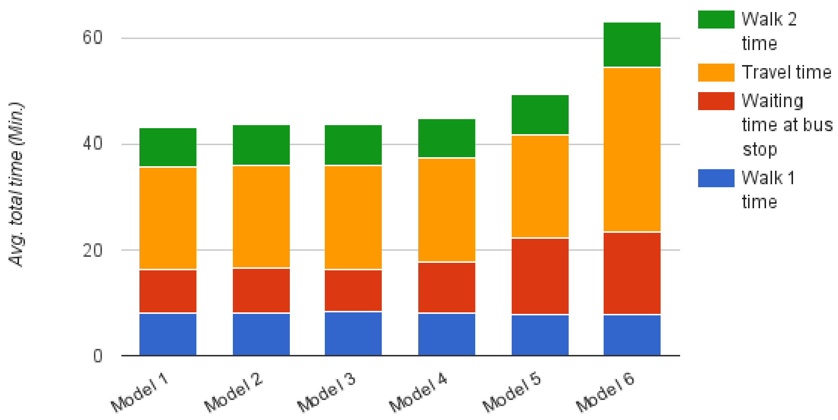

Fig. 5 Average total travel time for higher frequencies

3.4 Higher irregularity

In this experiment we simulate a scenario of higher service irregularity, meaning that effective timetables differ from the static ones to a higher extent. This may be due mainly to heavy traffic conditions and it is simulated in our model by increasing the standard deviation parameter of the normal probability distribution which represents the bus travel time along network links.

As we can observe from Table 3, mean travel time increased 14% in average, with respect to the current system. The main cause of this fact is that the system appears to be less reliable; this can be observed in Fig. 6, which clearly shows that waiting time is the main component which contributes to the overall time increase. Moreover, the tendency already observed in the experiment of Section 3.2 does not hold here. Passengers that do not update their decisions based on real-time information (models RTI@origin and STT) experience the worst increase with respect to the more regular scenario, which is a reasonable expectation.

Table 3 Mean total travel time and confidence interval for high irregularity

Model	Mean (secs.)	Half conf. Interval	% increase w.r.t. current system
1. RTI-Allways	2972.40	4.05	15
2. RTI@origin	3037.10	3.75	16
3. RTI-1line	2979.42	4.07	13
4. STT	3159.30	3.92	17
5. RTI@stops	3219.14	4.07	9
6. FBA	4189.05	6.07	11

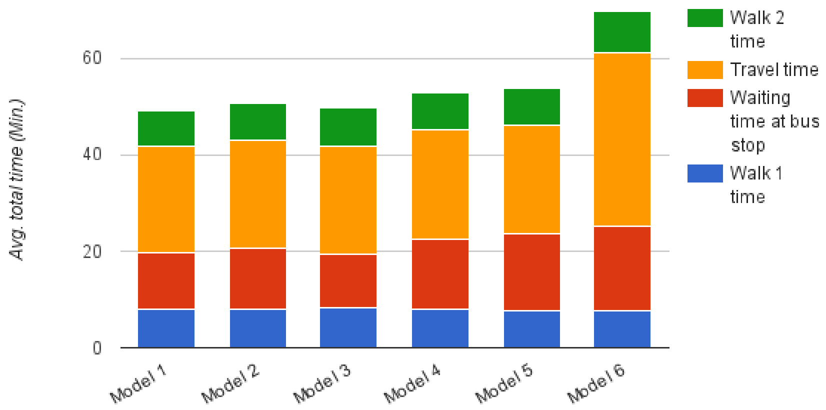

Fig. 6 Average total travel time for high irregularity

4 Conclusions and future work

We have presented an experimental evaluation of the impacts of real-time information in transit systems from the perspective of users. The study is focused on scenarios concerning small cities, where transit systems have low frequencies, are not congested and operate with high regularity. Six variants (models) of passenger behaviour concerning information availability were proposed an implemented, based in a common framework of discrete event simulation. The models are simple an efficient, which contributes to facilitate the validation and experimentation. The experiments aim to compare results of the six models in terms of average total travel time and respective confidence intervals. Also, non-aggregated measures are analysed (different stages of travel and different OD-pairs) to support the conclusions.

Main findings are that in regular systems with low frequencies, simply by publishing the static timetables and by encouraging to use them, significant saving in terms of total travel time (about 29%) can be obtained, with respect to current practices where users plan their trips by adopting a frequency-based approach. The most sophisticated scenario attains an even greater saving (45%) with respect to current practices. This scenario assumes that users have access to real-time information at any stage of their trips, which requires having mobile phones or displays at the stops, and also assumes that an accurate estimation of arrival times is computed and broadcasted to users. It is worth noting that the frequency-based approach used in our model assumes a behaviour which impacts negatively in total travel time: users take the first bus which leads to destination, because they consider that waiting time is the most onerous component. In general terms, we can say that even though real-time information enables to save travel time, an intermediate alternative which publishes timetables and encourage to use them seems to be very effective and doubtlessly cheaper in terms of infrastructure requirements. This statement is reinforced for transit systems with high frequencies. Finally, real-time information turns itself more relevant under the presence of high irregularity of the transit services, where users that adapt their decisions based on updated information have less chances of experiencing an increase in total travel time due to irregularity.

As future work, we identify several lines. Tests with other cases are needed in order to investigate whether conclusions formulated in this work also hold for similar scenarios. Applying other state-of-the-art methods for modelling passenger behavior to the case of our study, would give additional validation of the conclusions; also, complexities not considered by the models proposed in this work (e.g. congestion or different cost perceptions) could be included. Experiments with larger cases is a pending task, in fact, all published studies deal with small or medium sized cases. This poses at least two challenges: the computational cost of running a large model (more than 100 lines) and the processing of results, which requires a more detailed analysis since different OD-pairs may have very different trip patterns. Finally, a graphical validation and experimenting tool would be very useful. Transit systems have an intrinsic geographic characteristic, therefore a GIS[4]-based tool which shows paths of different OD-pairs contributes to increase the trust in the developed models. Our group is currently working on some of these research lines.

Acknowledgements: This research was partially supported by the Latin American and Caribbean Collaborative ICT Research (LACCIR, project RFP-1212LAC002). Ricardo Giesen also gratefully acknowledges the support provided by CEDEUS, CONICYT/FONDAP 15110020 and FONDECYT 1120993.

[4] Geographic Information System

Appendix A Passenger behaviour Models: detailed description

Model 1, RTI-allways:

1. At the origin centroid, the passenger checks real-time information about the transit system.
2. The stops connected to the origin centroid which are unreachable due to temporal reasons are discarded. Among the resulting stops, the passenger checks for the lines that lead to destination and selects the one that minimizes the total travel time.
3. The passenger holds at the origin centroid a reasonable time in order to coordinate its arrival to the stop with the arrival of the bus. This time is not computed as waiting.
4. Walk to the selected stop.
5. While waiting at the stop, for each bus that passes from there and leads to destination (independent of the original decision), the passenger compares the arrival time to destination using that line, against arrival times of the other buses which are going to arrive. If a better alternative (in terms of total travel time and considering real-time information) is approaching the stop, the passenger skips the current bus and repeats this reasoning for the next bus which arrives to the stop and leads to destination. If there are not better upcoming alternatives, the passenger takes the current bus.
6. The bus stops at destination, the passenger alights and walks to its destination centroid.

Model 2, RTI@origin: Identical to model 1, except for the behaviour at the stop.

5. While waiting at the stop, the passenger only takes the bus of the line that was selected at origin.

Model 3, RTI-1line:

1. At the origin centroid, the passenger selects the line which minimizes total travel time to destination, based on static information. For each line, the waiting time is considered as the headway (worst case).
2. The passenger checks real-time information about the selected line and holds at the origin centroid a reasonable time in order to coordinate the arrival to the stop with the arrival of the bus. This time is not computed as waiting.
3. Walk to the selected stop.
4. Once at the stop, the passenger checks again the real-time information about the selected line and remembers the corresponding arrival time of the bus for future comparisons.

5. Whenever a bus arrives to the stop, if it belongs to the line selected, the passenger boards. If not, but it is a line that leads to destination, the passenger compares the remaining travel time using that line, against the travel time of the selected line. If it is better, she/he takes the alternative bus; otherwise, continues waiting for the one selected originally.
6. The bus stops at destination, the passenger alights and walks to destination centroid.

Note that information about the static timetable is not used in this model.

Model 4, STT: Almost identical to model 2, but using static information.

Model 5, RTI@stops:

1. Identical to model 3.
2. Walk to the selected stop (do not hold at origin).
3. While waiting at the stop, the passenger behaves as in step 5 of model 1.
4. The bus stops at destination, the passenger alights and walks to destination centroid.

Note that information about the static timetable is not used in this model.

Model 6, FBA: Almost identical to model 5, except for the behaviour at the stop.

3. Whenever a bus arrives to the stop and belongs to a line that leads to destination, the passenger boards, independent of the travel time.

References

Anez, J., de la Barra, T. and Perez, B. (1996) Dual graph representation of transport networks. Transportation Research Part B, 30, 209-216.

Brakewood, C., Barbeau, S. and Watkins K. (2014) An experiment evaluating the impacts of real-time transit information on bus riders in Tampa, Florida. Transportation Research Part A, 69, 409–422.

Cats, O. (2011) Dynamic modelling of transit operations and passenger decisions. Doctoral Dissertation, KTH Royal Institute of Technology.

Cats, O., Koutsopoulos, H., Burghout, W. and Toledo, T. (2011) Effect of real-time transit information on dynamic path choice of passengers. Transportation Research Record 2217, 46–54.

Chen, P. and Nie, Y. (2015) Optimal transit routing with partial online information. Transportation Research Part B, 72, 40–58.

Chriqui, C. and Robillard, P. (1975) Common bus lines. Transportation Science, 9(2), 115-121.

Coppola, P. and Rosati, L. (2010) Simulation-based evaluation of advanced public transportation information systems (APTIS). In: Wilson, N. and Nuzzolo, A.

(Eds.), Schedule-based modeling of transportation networks: theory and applications. Springer, pages 195-215.

Domencich, T. and McFadden, D. (1975) Urban travel demand. North Holland.

Gentile, G., Nguyen, S. and Pallottino, S. (2005) Route choice on transit networks with online information at stops. Transportation Science, 39(3), 289–297.

Hickman, M. and Wilson, N. (1995) Passenger travel time and path choice implications of real-time transit information. Transportation Research Part C, 3(4), 211-226.

Law, A. (2006) Simulation modeling and analysis. McGraw-Hill.

Nguyen, S. and Pallottino, S (1988) Equilibrium traffic assignment for large scale transit networks. European Journal of Operational Research, 37(2), 176–186.

Nuzzolo, A. (2003) Transit path choice and assignment model approaches. In: Lam, W. and Bell, M. (Eds.), Advanced Modeling for Transit Operations and Service Planning. Elsevier, pages 93–124 .

Nuzzolo, A. and Crisalli, U. (2004) The schedule-based approach in dynamic transit modelling: a general overview. In: Wilson, N. and Nuzzolo, A. (Eds.), Schedule-based dynamic transit modeling: theory and applications. Kluwer, pages 1-24.

Nuzzollo, A. and Russo, F. (1994) An equilibrium assignment model for intercity transit networks. Proceedings of TRISTAN II Conference, Capri, Italy.

Spiess, H. and Florian, M. (1989) Optimal strategies: a new assignment model for transit networks. Transportation Research Part B, 23(2), 83–102.

Tocher, K. (1963) The art of simulation. English Universities Press.

Tong, C. and Richardson, A. (1984) Estimation of time-dependent origin-destination matrices for transit networks. Journal of Advanced Transportation, 18, 145-161.

Watkins, K., Ferris, B., Borning, A., Rutherford, G. and Layton D. (2011) Where is my bus? Impact of mobile real-time information on the perceived and actual wait time of transit riders. Transportation Research Part A, 45, 839–848.