Travel time prediction for~~in~~ selective bus lines

Abstract

The prediction of the progress ~~evolution of the trip~~ of a bus trip is a critical issue in the operational planning of urban transportation. It is critical not only because it allows the planners to anticipate ~~the~~ problems that arise during ~~the~~ daily operation, but also because it allows users to be informed about the remaining time of the trip. Numerical results ~~found~~ for such predictions show ~~the~~ greater potential of ~~in the~~ us~~e~~ing ~~of~~ neural networks ~~for~~ over other techniques. However, the studies mostly ~~essentially~~ examine ~~consider~~ bus lines with intermediate bus stops, and pay ~~given~~ less attention to the time series effect that exists in the historical data. A relevant case occurs when planning selective bus lines that board or drop ~~leave~~ passengers upon request~~on demand~~. These are bus ~~the~~ lines without intermediate stops, so ~~that~~ passengers embark or disembark at any point along the route. Recurrent neural networks ~~are~~ specialize~~d~~ in~~to~~ predict~~ing~~ time series, such as those that occur in a sequence of GPS points transmitted by a bus. Although there is some evidence of ~~about~~ their efficiency in different machine learning problems, less attention has been focused on ~~have received in terms of~~ predicting bus travel times. This manuscript address the prediction of travel time for ~~in~~ selective bus lines using ~~by means of~~ recurrent neural networks. The network predicts the bus location ~~off in~~ at a future point in time~~instant~~ depending on~~of~~ the start time of the bus trip, the current bus location (~~measured in~~ GPS coordinates) at the current time, and the distance variation for ~~in~~ the last period. Specifically, a network that predicts the next minute of ~~the~~ bus travel has a prediction error of up to 30 seconds, whereas the network predicts the next 10 minutes with an error of up to one~~1~~ minute.

Keywords: travel time prediction, bus travel time, recurrent neural networks.

1. Introduction

Predicting the travel time of a bus in a public transportation system has a high~~great~~ impact on the quality of passenger service.~~ to passengers.~~ In~~This~~ fact, operational management deals with fleet and scheduling definitions that define~~s~~ the accessibility of passengers to the service. When an unexpected situation arises affecting ~~the~~bus trips~~ of the buses~~, the appropriate contingency measures must be adopted as soon as possible, forcing a reprogramming of the service. A typical situation arises when ~~the~~ buses gradually lose separation ~~the headway~~ between each other ~~them~~ during a ~~the~~ trip, causing them ~~giving rise~~ to bunch ~~ing of buses~~together, a phenomenon that directly affects ~~the~~ service quality ~~of the service~~ and that can be mitigated by knowing the dynamics of the trip in the near future. ~~Then,~~By knowing in advance the trip duration ~~of the trip~~ or the time remaining to reach any point on the route, ~~allows~~ management is able to make~~ing~~ ~~the~~ appropriate decisions to maintain the planned~~scheduled~~ service quality~~ of the service~~. The remaining travel time is also useful information for the passengers of a particular bus, especially in the face of unexpected events (Bin, Zhongzhen, & Baozhen, 2006; H.-K. Chen & Wu, 2012).

The prediction of bus travel time is complex because it depends on several factors. One of them is whether a bus trip occurs during ~~the moment in which the bus is developing its trip either in~~ a peak~~,~~ or ~~in an~~ off-peak hour, and during a weekday or ~~on a day of the week, or on a~~ holiday.~~ day.~~ In addition, the prediction depends on the demand for trips distributed along the route and the number of passengers traveling. Other relevant factors are the driving style of each driver, the types of buses available, ~~the~~ weather conditions, and any special ~~the~~ events scheduled ~~or not, that take place~~ in the city. Also, two different situations arise when a bus line has ~~is~~ defined ~~with~~ stops (~~, that is~~ the standard case), or~~,~~ when the line is selective - ~~,~~that is, buses stop upon request~~on demand~~ (Vuchic, 2005). As such, ~~Then,~~ a theoretical model that represents this complex problem has not been presented ~~is not known~~ until now, and the literature on the subject~~,~~ primarily consists of ~~has focused essentially on~~ empirical studies based on historical data.

~~The b~~Bus travel time prediction has been studied for some specific situations, and two of them stand out: the prediction of ~~the~~ bus stop arrival time ~~to a bus stop~~ and the total travel time (Bin, Zhongzhen, & Baozhen, 2006, H.-K. Chen & Wu, 2012). For such cases, non-linear regression models have been considered. These models are based on Kalman filters, support vector machines and standard neural networks (Karlaftis & Vlahogianni, 2011; Mori, Mendiburu, Álvarez, & Lozano, 2015). The solution methods predict the travel time based on a set of independent variables. Authors suggest that standard neural networks predict bus travel time more accurately than other models (Mazloumi, Moridpour, Currie, & Rose, 2012; Mazloumi, Rose, Currie, & Moridpour, 2011). However, such networks do not consider the fact that GPS data is recorded sequentially in time, producing ~~giving rise to~~ time series data. The evolution of~~n~~ the field of neural networks, largely arising ~~that comes essentially~~ from ~~the~~ advances in machine learning, has resulted in ~~given rise to~~ a type ~~kind~~ of network that captures ~~from the data~~ the dynamics of a time series from the data. ~~; it is t~~These recurrent neural networks (RNN~~R~~s) have ~~that has~~ been minimally ~~least~~ considered in the field of urban transportation (Goodfellow, Bengio, & Courville, 2016; Lewis, 2017). The use of an RNN~~R~~ would enable~~allow~~ a more accurate prediction than ~~the prediction~~ with a standard ~~RN~~neural network. In fact, the large amount of information ~~that is~~ stored daily ~~every day~~ in ~~product of~~ GPS devices~~,~~ is the appropriate dataset for the training and evaluation~~ng~~ phases of an RNN~~R~~, for both ~~a~~ standard bus lines and the selective bus lines that have received less attention in the literature.

Bus travel time prediction can be used in many ways. Data may be transmitted to portable device allowing passengers to know the remaining duration of their trip. The same information may be available on~~by a~~ passenger information display systems in a bus or in a bus stop (2004). In addition, operators could identify when buses are ~~in~~ bunching in order to take ~~the~~ appropriate early steps to mitigate ~~the~~ potential problems.

In this manuscript, ~~the~~ bus travel time prediction for a given time window is examined for ~~taken into account, considering~~ a selective bus line. The prediction is performed with an RNN that is trained with one year of operations data. The performance of the network is studied using ~~by~~ the mean squared error (MSE) and the quality of the prediction is measured in a time window. Likewise, the robustness of the network is studied in order to predict ~~the~~ travel time during ~~in~~ peak ~~time~~ and off-peak hours~~time~~.

2. Literature review

One of the p~~P~~ioneering works in this field was one that ~~the~~ propos~~al~~ed ~~of~~ two models of neural networks to predict bus travel times in a computer-~~-~~simulated transportation line (Chien, Ding, & Wei, 2002). These authors used software to simulate a bus line in New Jersey, USA, and two neural networks were proposed. The first one considered information between each pair of intersections such as the volume of traffic, bus speed and travel time. The second network considered aggregated information such as average vehicular traffic, average bus speed and the standard deviation of the speed. ~~From all input variables, d~~Different scenarios were analyzed from the input variables, generating ~~originating~~ 10 different networks for training. The largest of these networks ha~~s~~d seven~~7~~ nodes in a single intermediate layer that was reflected ~~is considered~~ in all the networks. A total of 380 training examples were ~~–~~studied~~considered~~. The low prediction errors obtained ~~allowed~~ established~~ing~~ the potential of ~~the~~ using~~e~~ ~~of~~ neural networks for ~~in the~~ time predictions for urban transportation lines.

Chen et al. (2004) proposed a neural network that predicts the bus travel time between two points of the route, considering as input variables the day of the week, the time of ~~the~~ day, rainfall amounts and the segment under study.~~ as input variables.~~ Some ~~Part~~ of the data were~~are~~ obtained by means of an *automatic passenger counting~~er~~ system* from a 29.5-mile route that runs through three counties in New Jersey and has ~~, which has a length of 29.5 miles and with~~ 17 bus stops. Both the number of nodes and the number of layers were empirically determined by trial and error. The best performance was~~is~~ obtained with a network that had ~~of~~ one layer and six nodes. In order to dynamically incorporate the information that is received in real time, the authors also included a Kalman filter to adjust the prediction of the network. The evaluation MSE for the five tested variables is 0.009.

In one of the field's classic~~al~~ works~~ of this field~~, Jeong & Rillet (2005) show that when using a neural network with a single intermediate layer of up to 15 nodes, significantly better results can be obtained than those using ~~found by~~ a regression model and those based on average historical data. In order to reach this conclusion, a neural network with 13 learning patterns and two~~2~~ different activation functions were considered. Neural n~~N~~etworks predicted the time required by a bus ~~that is~~ at ~~in~~ a current stop to reach a future ~~bus~~ stop. A relation was~~is~~ established between that time and: the arrival time at~~to~~ the current bus stop, the time used to ~~of~~ board~~ing~~ and drop ~~leaving~~ passengers at~~in~~ that bus stop, and the interval between the current time and the ~~time~~ originally scheduled time. To determine the number of nodes of the network's internal layer~~ of the network~~, different runs were performed, demonstrating ~~observing~~ that the best performance was~~is~~ achieved with the highest number of nodes. These results were evaluated with a dataset from the city of Houston, Texas.

The use of a standard neural network with a single intermediate layer and GPS data efficiently predicted the remaining travel times in inter-municipality buses. Specifically, Gurmu & Fan (2014) predicted the time interval from a~~the~~ current location and a future bus stop. The independent variables were~~are~~: the period of the day, the ID~~id~~ of the current bus, and an ID ~~id~~ of the future bus stop. The study case used ~~considers~~ GPS data from the bus line section (with 35 bus stops) between Macae and Rio de Janeiro, two cities in Brazil ~~between which there are 35 bus stops~~. It was observed that the network produced~~s~~ better results than~~t~~ the analysis based on average historical data. In addition, Amita et al. (2016) used ~~, with~~ a similar model with ~~that considers~~ a single intermediate layer to determine ~~the~~ arrival times for real-time ~~to inform both~~ passenger~~s~~ information~~in real time~~, and for traffic agencies to apply service improvement strategies ~~that improves the service~~. The standard NN ha~~s~~d a hidden layer with up to 15 nodes and was~~is~~ trained with data corresponding to two bus lines (33 and 35 bus stops, respectively) in New Delhi, India~~, with 33 and 35 bus stops, respectively~~. The problem was~~is~~ also approached using ~~by~~ a multivariate, linear regression model. The most important result was that ~~is~~ the best performance was produced ~~generated~~ by the standard NN, as compared ~~related~~ to the regression model.

The prediction of ~~the~~ bus arrival time has been also studied using ~~, taking together~~ GPS, traffic and weather data. This phenomenon was studied in order to identify the uncertainty produced by a neural network with a single intermediate layer with up to 15 nodes. Specifically, Mazloumi et al. (2011) evaluated a set of ~~consider a~~ 1,800 bus trips ~~set, carried out~~ during one ~~a~~ day in Melbourne, Australia; the ~~, by a~~ bus line was ~~subd~~divided into four subsections with intermediate bus stops. Five independent periods were identified in a day to perform the prediction. The uncertainty is divided in two parts: the first part ~~that~~ comes from the noise of the original data, and the ~~a~~ second part ~~that~~ comes from the model structure. It can be concluded that most of the uncertainty of the predictions is due to the noise of input data and a smaller part, though also important, is due to the model structure.

Models of deep neural networks have been ~~recently~~ used recently to predict the travel time of buses from a point at~~in~~ which a control sensor has been ~~is~~ installed, to ~~the~~ arrival at ~~until~~ the next traffic light. The aim is to adjust the ~~cycle of the~~ traffic light cycle according to the size of the string ~~row~~ of buses in order to speed up ~~the~~ traffic in the studied sector (Bie, Wang, & Qi, 2012). The problem was~~is~~ approached using ~~through~~ a neural network with several intermediate layers that are gradually trained by using the auto-encoder concept. The authors defined the auto-encoder as a standard feed-forward NN containing a single intermediate layer with the same number of inputs and outputs. The advantage of using the auto-encoder is that ~~the~~ training can be performed gradually and sequentially for each of the auto-encoders by using the ~~output of the~~ training network output as the training input for ~~to train~~ the next auto-encoder. Thus, the networks reflected ~~considered~~ up to five intermediate layers with up to 20 nodes each. To evaluate the network performance~~ of the network~~, an intersection was simulated with ~~by~~ software that ~~,~~ generated~~ing~~ sufficient ~~enough~~ input data to the network. This procedure allowed the authors to identify ~~found~~ a four~~4~~-second error in the prediction.

The potential~~ity~~ of ~~the~~ using~~e~~ ~~of~~ recurrent neural networks to ~~in the~~ predict~~ion of~~ city travel~~s in a city~~ was evident in a recent international competition to discover the locations ~~place~~ of taxi destinations in a city, based on ~~the~~ information from the beginning of their routes. The winn~~er~~ing team among the 381 participating groups presented a recurrent neural network and a bidirectional network to predict the taxi destinations ~~of the taxis~~ (de Brébisson, Simon, Auvolat, Vincent, & Bengio, 2015).

In the literature on travel time prediction, m~~M~~odels based on standard neural networks ~~in the literature for travel time prediction~~ have some common characteristics. In all cases, the prediction of an ~~the~~ arrival time at ~~to~~ a future bus stop is addressed using ~~based on~~ information from the current stop. To the best of our knowledge, there is no study that addresses selective bus lines. Although different sets of input variables are considered, the models are structurally very similar. Typically, a network is used that has ~~with~~ one or two intermediate layers with up to 10–~~ or 1~~15 nodes in each stage~~ is used~~. In particular, ~~the~~ networks working ~~that have worked~~ with two layers have used more ~~greater number of~~ nodes in the first layer than in the second one. In addition, the ~~considered~~ standard NN considered has no ~~information~~ feedback information to take into account the dynamics of a time series produced by ~~that comes from~~ GPS data.

3. Data preparation

Data used for the computational experiment are selected from a selective bus line. The buses make ~~perform~~ their trip ~~joining~~ between two neighborhoods in Porto Alegre, Brazil. The ~~travelled~~ distance travelled is 41.7 k~~m~~ilometers (km), and ~~in average,~~ the line transports an average of ~~around~~ 3,207 passengers per day. The bus company's 18 buses are ~~of the company,~~ equipped with a GPS device that transmits their locations to the servers through GPRS (general packet radio services) technology, ~~which is~~ a service offered by the cellular telephone companies. Time, latitude and longitude d~~D~~ata are recorded ~~registered~~ at ~~non-~~irregular time intervals~~ consist of the triplet (time,~~

~~latitude, longitude)~~. In addition, data are distributed according to the time of ~~the~~ day, and follow a similar pattern ~~trend for~~ Monday through Friday. Data from the selective bus line for ~~corresponding to~~ the months of March- ~~to~~ December, ~~of the year~~ 2015, were selected. It is observed that some points are far from the main cloud of points that are approximated from the average travel time between origin and destination. In some cases, it is observed that the trip time is practically double~~s~~ the time taken for the bus at the same time and during ~~at~~ the same period. Generally, these data points correspond to errors produced during ~~when~~ data ~~were~~ collect~~ed~~ion ~~–~~or~~,~~ to an error in recording the ~~resulting from the~~ GPS data.~~registering.~~ Thus, the trips considered for the numerical experimentation have a travel distance between 16-~~.0 and~~ 29~~.0~~ km and a duration of ~~taking between~~ 36 minutes to three ~~and 3~~ hours. Since ~~As~~ our models predict~~s~~ at most the following 30 minutes, we only considered trips with travel times greater or equal than 40 minutes. In addition, because departure and arrival GPS points are known, we only considered trips with no more than a 0.0025 km difference between the known coordinates of ~~with a difference in~~ these two ~~such~~ points. ~~that is at most 0.0025 related to the known coordinates.~~ Figure 1 shows the data dispersion ~~of data already~~ filtered for each day of the ~~all~~ week~~s~~ in the measurement period, as well as aggregated ~~agglomerate~~ data for the five days. It is observed that ~~in~~ the morning hours have ~~are~~ the shortest travel times, and that these increase ~~that go up~~ quickly, ~~to~~ reaching a peak ~~hour~~ around 7:00 AM. ~~o'clock in the morning.~~

Insert~~ar~~ Figure~~a~~ 1
here~~aquí~~

The time-space points that reflect the progress of the bus were calculated and normalized. To do that, the GPS coordinates were transformed into a percentage of progress ~~advance~~ for a bus in terms of the distance travelled for a given time of ~~the~~ day. As the departure ~~time~~ of the first bus is ~~takes place~~ around 6:00 AM~~a.m.~~, and the last buses arrive at their ~~to~~ destination by midnight, the time of the day was normalized on a ~~in the~~ scale of 6- ~~to~~ 24 hours. Thus, data corresponding to each hour are in the interval [0.0 - 1.0]. The percentage of progress along ~~advance of~~ the travel distance is also normalized. To accomplish this, ~~According to this,~~ GPS points are divided by the travel distance recorded ~~registered~~ for that trip, producing ~~given rise to~~ a percentage that represents the progress of the trip. In addition, a second transformation of the data by lineal interpolation is performed to have points for every 60 seconds.

4. Neural network models

The first model is a standard NN (SN) that allows predicting the progress of a bus one minute after the current time. The model ~~It is~~ considers~~ed~~ that this ~~such~~ progress depends on both time elapsed and distance travelled from the departure time to the ~~until the~~ current time~~moment~~. Let $\tau_{0,v}$ be the departure time of trip v, and let $\tau_{t,v}$, be the current time. In addition, $\delta_{t,v}$ is the percentage of progress~~advance~~ of trip v from the departure at the terminal, and $\delta_{t+1,v}$ is the percentage of progress ~~advance~~ one minute after the current time. The~~n~~refore, the model is as presented in equation (1), where W is the matrix of parameters to be evaluated from historical data, and f represents the relationship between δ_{t+1} and the independent variables. This network has a layer with 20 nodes, each o~~i~~f which is activated by a sigmoid function. It is the standard model used in the literature, and is ~~considers~~ the largest ~~greater size~~ of those that have been studied.

$$\delta_{t+1,v} = f(\tau_{0,v}, \tau_{t,v}, \delta_{t,v}, W) \tag{1}$$

The second model (RN) predicts the next time-space point from a sequence of points belonging to the same trip. The model uses ~~considers~~ a RNN that uses the sequence of n points and $\tau_{0,v}$ as independent variables, as shown in equation (2). A layer with 20 LSTM nodes is implemented in this network.

$$\delta_{t+1,v} = f'(\tau_{0,v}, (\tau_{i,v}, \delta_{i,v}),\ldots,(\tau_{i+k,v}, \delta_{i+k,v}), W') \tag{2}$$

A combination of both models give rise to the third model that receives the time series information by ~~a~~ recurrent layers. The output is passed to a standard one-layer network. Both layers have 20 nodes, with ~~that an~~ LSTM nodes~~os~~ in the first layer and sigmoidal nodes in the second layer, following the same dependencies of model (2).

Input sequences to the network

The SN is trained with data of the bus's progress ~~of the bus~~ occurr~~ed~~ing in the ~~following~~ minute following;~~,~~ i.e., with data of the form ($\tau_{0,v}$, $\tau_{t,v}$, $\delta_{t,v}$, $\delta_{t+1,v}$). In addition, input to the RN and the SN-RN are sequences of points that correspond to parts of a given trip. Such sequences are composed by a set of sequential points of the trip, and the following steps allow~~s~~ the ~~to~~ preparation of~~e~~ 19,000 sequences, 3,800 of which are validation sequences.

- Randomly select a trip *v* and ~~the~~ departure point *i* from ~~of~~ the sequence ($\tau_{i,v}$, $\delta_{i,v}$).
- Randomly define ~~the~~ length *k* of a sequence $k \in \{1, 2, \ldots, 10\}$.

Construct the sequence of *k* points of the trip $\{(\tau_{i,v}, \delta_{i,v}), (\tau_{i+1,v}, \delta_{i+1,v}), \ldots, (\tau_{i+k,v}, \delta_{i+k,v})\}$

-
- Prepare the input to the network of the form $\{\tau_{0,v}, (\tau_{i,v}, \delta_{i,v}), (\tau_{i+1,v}, \delta_{i+1,v}), \ldots, (\tau_{i+k,v}, \delta_{i+k,v}), \delta_{i+k+1,v}\}$.

Network training

The training of the neural network is carried out with a number of epochs, so that in each epoch, ~~of them~~ the entire training dataset is used by the optimization algorithm. The training is performed with 80% of the data, and the validation of the constructed model is carried out with the remaining 20%. The validation data is~~was~~ obtained from 20% of the ~~from~~ data for ~~of~~ each month in order to obtain a representative sample of what occurs during the year. The training process is carried out through cross-validation, by dividing the training dataset into small batches of 10% each. Thus, the network is trained with 90% of the training data;~~,~~ that is, the optimization algorithm ~~considers~~ sequentially considers ~~the~~ 90% of the training data, and once that phase is completed, it is evaluated with the remaining 10%. The process is repeated 10 times to assure that all data are used as evaluation data. At the end of an epoch, the order of data is randomly altered, and the training and evaluation process is repeated. Both the training and evaluation processes are implemented in Torch 7 ~~is used~~ ("Torch | Scientific computing for LuaJIT.," n.d.). This language has a library that contains the codes to implement the recurrent neural network. The experiment was performed on a desktop computer with Ubuntu 16 and an Intel ® Core ™ i7-3770, 3.4 GHz processor and 12GB of RAM.

5. Results

The RN-SN ~~presents better~~ performs~~ance~~ better than the RS; the r~~. R~~esults are presented in Table 1. The first column indicates the name of the network, the second ~~one~~ explains the network characteristics, and the third and fourth columns show the number of layers and ~~number of~~ nodes. Columns five and six~~fifth and sixth~~ shows the values for the mean squared error (MSE) and the mean absolute error (MAE), respectively. In the last two columns, an estimate of the error of each network measured in meters and seconds is presented. It is observed that both MSE and MAE have better results for RN-SN and RN than for SN. ~~Then t~~The recurrent parameters capture the ~~de~~ time series better than the ~~de~~ standard NN. This is an interesting result, because it shows the benefit of using the recurrent information of the time and distance series. It also allows establishing the comparison with the state-of-the-art SN network~~ SN~~. The approximate difference in prediction between the RN-SN and the SN is 29.41 meters and 3.60 seconds. Such values have practical relevance because they mean that, o~~i~~n average, the remaining time for a passenger to get off ~~leave~~ the bus has an error of that~~such~~ magnitude.

Insert~~ar~~ Tabl~~a~~e 2~~1~~
here~~aquí~~

The RN-SN is effective for ~~to~~ performing the ~~bus~~ prediction in a 20-minute ~~half an hour~~ time window. To study the efficiency of the trained neural network, we used it to predict the position of some randomly selected buses o~~i~~n different days and times ~~periods of the day~~ by selecting different areas of the validation dataset. Specifically, data were selected from the peak ~~time~~ and off-peak time intervals. Figure 2 shows 10 instances starting at three different~~s~~ times and corresponding to different months. The distance in meters travelled by a bus ~~in meters~~ is presented on the y-axis and the time is on the abscissa (x) axis. The c~~C~~urves shows real and projected progress. In all the situations presented, the prediction follows practically the same progress, regardless of the time range, the day of the week, ~~or~~ the month, or ~~of~~ the year. This highlights the robustness of the network, at least for the selected examples. Furthermore, Figure 3 depicts the predictions corresponding to 10 randomly selected trips that start in the off-peak interval. As in the case of peak hours~~ time~~, the projections for the next 30 ~~next~~ minutes are practically superimposed with the real progress. Either way, in these selected cases, the prediction error is no more than a few seconds. Thus, the projected progress follows the same trend as for the ~~on~~ off-peak intervals.

Insert~~ar~~ Figure 2 here~~aquí~~

The evaluation of the trips performed during ~~in~~ peak hours produces an average error equivalent to 140 meters in the bus location. From the validation set, all travel during~~s~~ ~~corresponding to~~ peak hours ~~time~~ were selected and the prediction was performed sequentially, minute by minute, in the 30-~~-~~minute time window. The overall average ~~of the~~ deviation of the bus's position ~~of the bus~~ in relation to its real position is a little more than one block. By considering an average speed of 30.3 km/h, this distance is equivalent to a difference of 17.2 seconds. This difference is consistent with the performance found in the literature on ~~in~~ the use of recurrent neural networks in problems ~~of a~~ different from the one~~nature to the problem~~ studied here.

Insert~~ar~~ Figure 3 here~~aquí~~

The average error in the difference of the curves for~~on~~ this set of 20 trips is 157 meters in the prediction of the next minute. Thus,~~Then~~ the prediction has an effect similar to that produced by the model proposed by Chen et al. (2004)~~i~~, who performed predictions with a conventional neural network and corrected the prediction with a Kalman filter. In this case, the trained neural network offers an alternative way to perform this adjustment.

6. Conclusions

This article considers the problem of predicting the travel time of~~in~~ urban buses. Specifically, ~~In particular,~~ the prediction of travel time in buses that belong to a selective bus line with no intermediate stops. In order to conduct~~carry out~~ the study, ~~there is~~ a dataset was used that corresponded~~ing~~ to a year of bus line operation, ~~of the line~~ and that ~~they~~ contained both the time and location of each bus~~,~~ transmitted by a GPS device. The data corresponded to a bus line that ~~performs its daily~~ operates~~ions~~ daily in Porto Alegre, Brazil. A recurrent neural network model is proposed to predict the travel time of each bus. Networks are trained with ~~data~~ available data that were~~are~~ separated into two groups; the first group was~~is~~ used to perform the ~~training of the~~ network training through cross-validation, and the second was used ~~one~~ to evaluate the ~~already~~ trained network. Through a trial and error procedure, several network topologies were explored. The prediction accuracy ~~in the prediction~~ of the models was tested by projecting the predictions in a 30-~~-~~minute time window.

The best prediction was~~is~~ obtained by ~~with a~~ combining~~ation~~ ~~of~~ a standard NN with a recurrent NN. The resulting network had~~s~~ two layers of ~~with~~ 20 nodes each. In addition, the trained model was able to ~~is capable of~~ decode the progress ~~evolution~~ of the bus ~~during its trajectory in~~ at different times ~~periods~~ of ~~the~~ day, ~~in~~ on different days of the week, and in different months of the year. This results suggests that the use of the network model for managing integrated information in the training phase enables the ~~is capable to~~ learning of the ~~proper~~ irregularities pertaining to ~~of~~ the system.

References

Amita, J., Jain, S. S., & Garg, P. K. (2016). Prediction of Bus Travel Time Using ANN: A Case Study in Delhi. *Transportation Research Procedia*, *17*, 263–272. https://doi.org/10.1016/j.trpro.2016.11.091

Bie, Y., Wang, D., & Qi, H. (2012). Prediction Model of Bus Arrival Time at Signalized Intersection Using GPS Data. *Journal of Transportation Engineering-Asce*, *138*(1), 12–20. https://doi.org/10.1061/(ASCE)TE.1943-5436.0000310

Bin, Y., Zhongzhen, Y., & Baozhen, Y. (2006). Bus arrival time prediction using support vector machines. *Journal of Intelligent Transportation Systems*, *10*(4), 151–158. https://doi.org/10.1080/15472450600981009

Chen, H.-K., & Wu, C.-J. (2012). Travel Time Prediction Using Empirical Mode Decomposition and Gray Theory Example of National Central University Bus in Taiwan. *Transportation Research Record*, (2324), 11–19. https://doi.org/10.3141/2324-02

Chen, M., Liu, X. B., Xia, J. X., & Chien, S. I. (2004). A dynamic bus-arrival time prediction model based on APC data. *Computer-Aided Civil and Infrastructure Engineering*, *19*(5), 364–376. https://doi.org/10.1111/j.1467-8667.2004.00363.x

Chien, S. I. J., Ding, Y. Q., & Wei, C. H. (2002). Dynamic bus arrival time prediction with artificial neural networks. *Journal of Transportation Engineering-Asce*, *128*(5), 429–438. https://doi.org/10.1061/(ASCE)0733-947X(2002)128:5(429)

de Brébisson, A., Simon, É., Auvolat, A., Vincent, P., & Bengio, Y. (2015). Artificial Neural Networks Applied to Taxi Destination Prediction. Retrieved from https://arxiv.org/abs/1508.00021

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. (F. Bach, Ed.). Cambridge, Massachusetts: The MIT Press.

Gurmu, Z. K., & Fan, W. (2014). Artificial Neural Network Travel Time Prediction Model for Buses Using Only GPS Data. *Journal of Public Transportation*, *17*(2), 45–65.

Jeong, R. H., & Rilett, L. R. (2005). Prediction model of bus arrival time for real-time applications. In *Transit: Planning, Management and Maintenance, Technology, Marketing and Fare Policy, and Capacity and Qualtiy of Sevice* (Vol. 1927, pp. 195–204). Washington: Transportation Research Board Natl Research Council.

Karlaftis, M. G., & Vlahogianni, E. I. (2011). Statistical methods versus neural networks in transportation research: Differences, similarities and some insights. *Transportation Research Part C-Emerging Technologies*, *19*(3), 387–399. https://doi.org/10.1016/j.trc.2010.10.004

Lewis, N. D. (2017). *Neural Networks for Time Series Forecasting with R: An Intuitive Step by Step Blueprint for Beginners*. CreateSpace Independent Publishing Platform.

Mazloumi, E., Moridpour, S., Currie, G., & Rose, G. (2012). Exploring the Value of Traffic Flow Data in Bus Travel Time Prediction. *Journal of Transportation Engineering-Asce*, *138*(4), 436–446. https://doi.org/10.1061/(ASCE)TE.1943-5436.0000329

Mazloumi, E., Rose, G., Currie, G., & Moridpour, S. (2011). Prediction intervals to account for uncertainties in neural network predictions: Methodology and application in bus travel time prediction. *Engineering Applications of Artificial Intelligence*, *24*(3), 534–542. https://doi.org/10.1016/j.engappai.2010.11.004

Mori, U., Mendiburu, A., Álvarez, M., & Lozano, J. A. (2015). A review of travel time estimation and forecasting for Advanced Traveller Information Systems. *Transportmetrica A: Transport Science*, *11*(2), 119–157. https://doi.org/10.1080/23249935.2014.932469

Torch | Scientific computing for LuaJIT. (n.d.). Retrieved October 23, 2018, from http://torch.ch/

Vuchic, V. R. (2005). *Urban Transit : Operations, Planning and Economics* (1 edition). Hoboken, N.J: Wiley.

Table 1: Comparison of three neural networks

Name	Type of network	Number of layers	Number of nodes	MSE	MAE	Distance error [m]*	Time error [s]**
SN	Standard	1	20	0.025	0.38	186.27	22.80
RN	Recurrent	1	20	0.022	0.37	181.37	22.20
SN-RN	Recurrent	2	20	0.019	0.32	156.86	19.20

* Considering an average travel distance ~~of advance~~ of 490.18 m/min.

** Considering an average speed of 29.41km/h

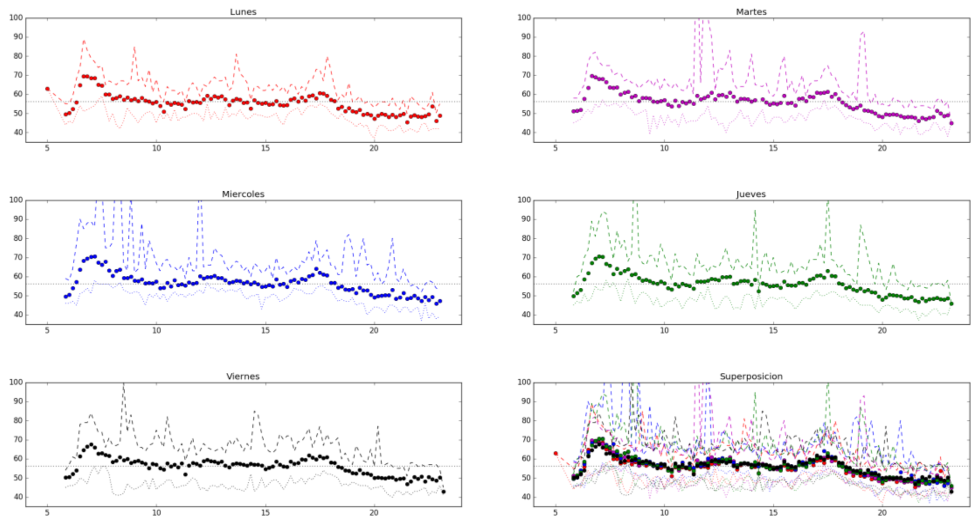

Figure 1: Display of filtered travel times

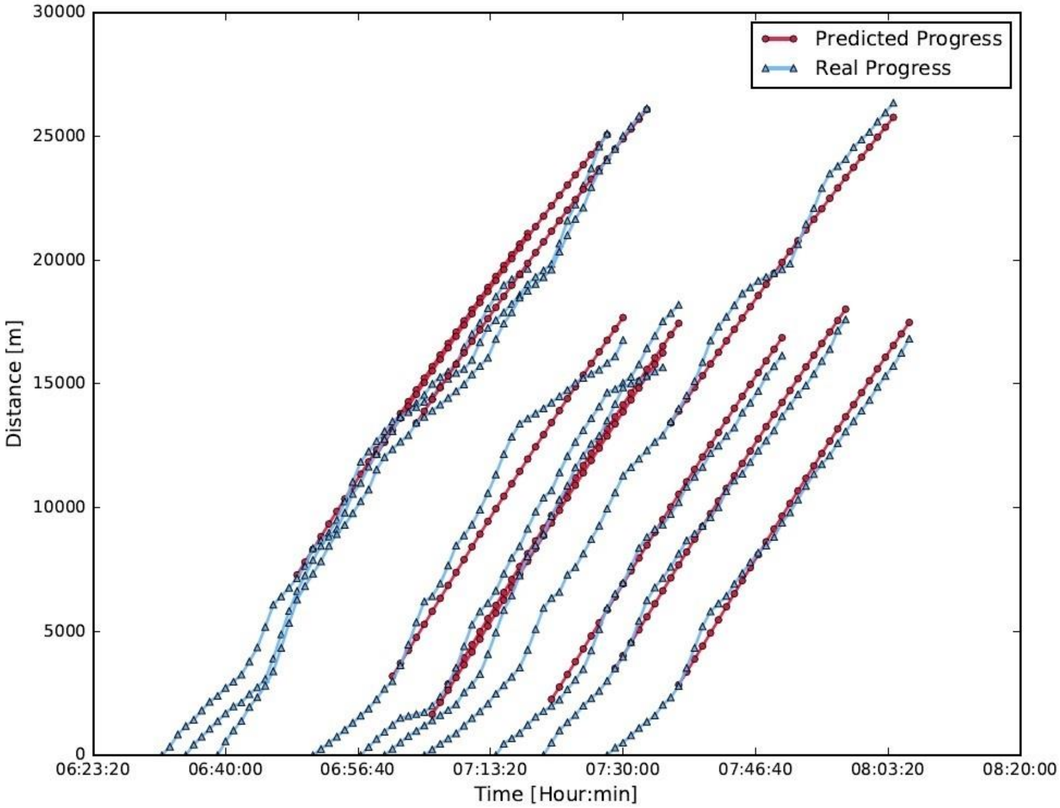

Figure 2: Predictions ~~about predictions~~ of the next minute of a bus during ~~in~~ peak hours~~time.~~

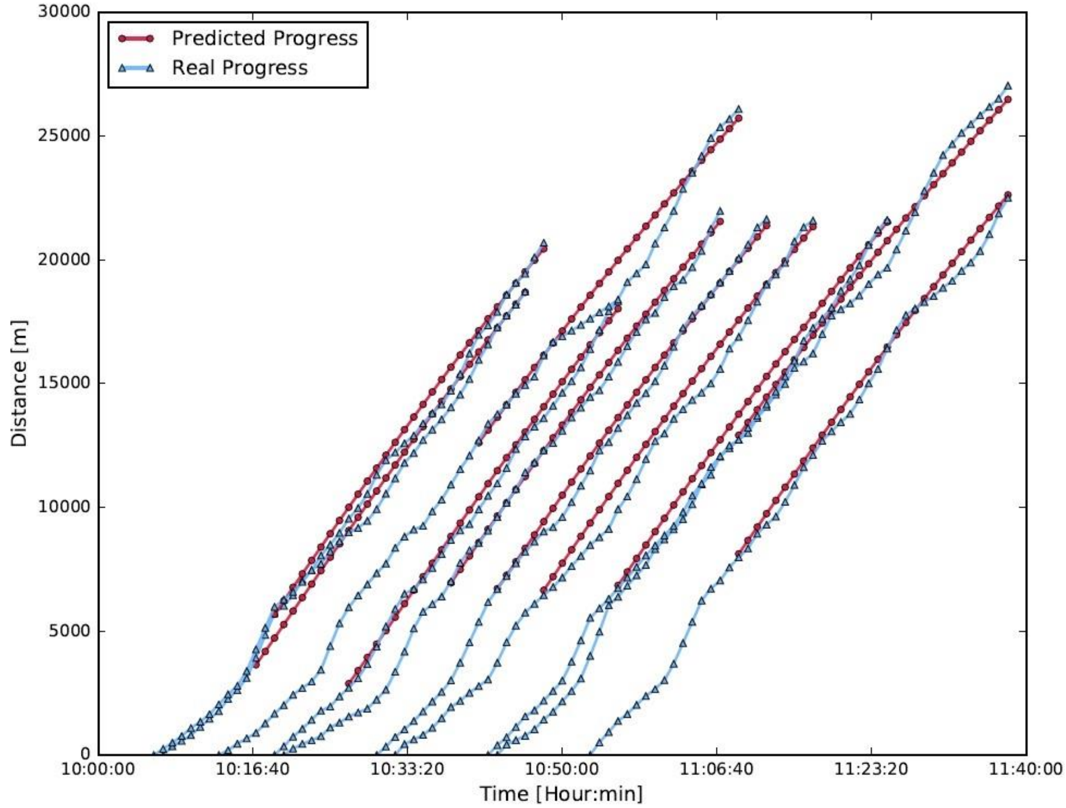

Figure 3: ~~Prediction about p~~Predictions of the next minute of a bus during ~~in~~ off-peak hours~~time.~~