Australasian Transport Research Forum 2011 Proceedings
26 - 28 September 2012, Perth, Australia
Publication website: http://www.patrec.org/atrf.aspx

Benefits and issues of bus travel time estimation and prediction

Le Minh Kieu[1], Ashish Bhaskar[1], Edward Chung[1]

[1]Smart Transport Research Centre, Faculty of Built Environment and Engineering, Queensland University of Technology, Brisbane, Australia

Email for correspondence: *le.kieu@qut.edu.au*

Abstract

Bus travel time estimation and prediction are two important modelling approaches which could facilitate transit users in using and transit providers in managing the public transport network. Bus travel time estimation could assist transit operators in understanding and improving the reliability of their systems and attracting more public transport users. On the other hand, bus travel time prediction is an important component of a traveller information system which could reduce the anxiety and stress for the travellers. This paper provides an insight into the characteristic of bus in traffic and the factors that influence bus travel time. A critical overview of the state-of-the-art in bus travel time estimation and prediction is provided and the needs for research in this important area are highlighted. The possibility of using Vehicle Identification Data (VID) for studying the relationship between bus and cars travel time is also explored.

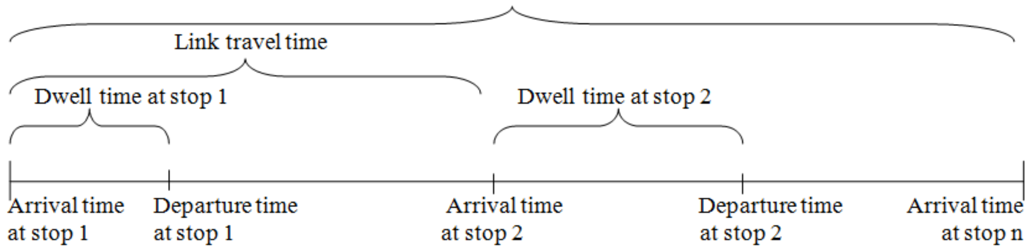
1. Introduction

Encouraging passengers to switch from using a car to a public transport is believed to be an effective solution for the emission and congestion problems of modern transport (Wilkie & Van Ewijk, 2009). Many policies from transit agencies have been issued in order to increase the attractiveness of public transportation to passengers. On the other hand, approaches for modelling, predicting and evaluating the network performance before and after the implementations of these policies have been proposed by the scientists around the world. Estimating the bus travel time and predicting the arrival time of the bus are the two promising modelling approaches which help the transit provider in improving their service reliability and attracting more passengers to use public transport.

2. Travel time estimation and prediction definition

The travel time, as defined in the passenger view's in Transit Capacity and Quality of Service Manual (TRB, 2003) is the total time it takes to travel from the origin to the destination. In this study, we define the travel time is the total time it takes to travel from one bus stop to another i.e., the time difference of the arrival time of a bus at upstream and downstream bus stops. The travel time definition is illustrated in the Figure 1.

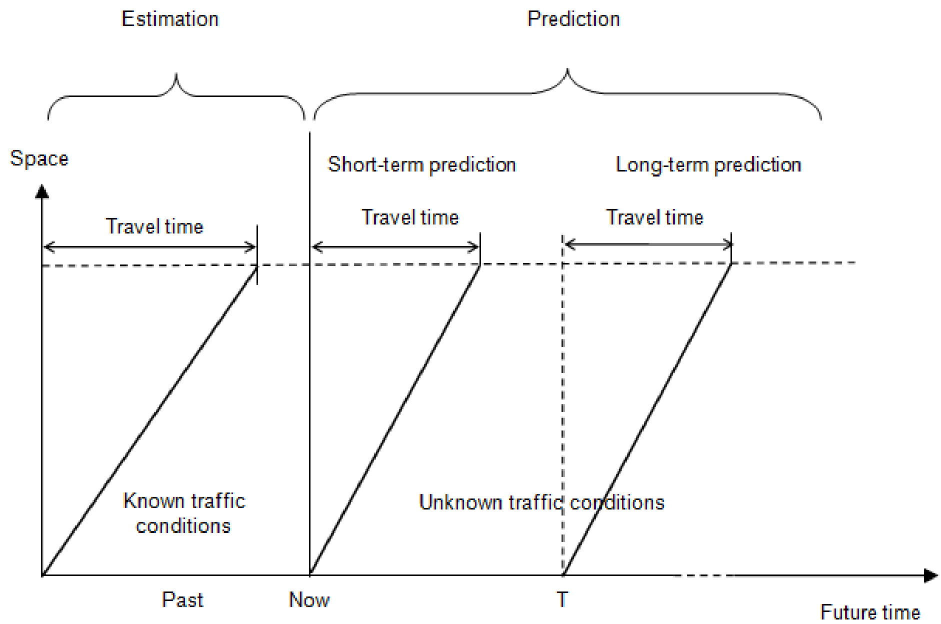
Figure 1: Travel time and related glossary



Travel time estimation is defined here as an offline application. The historical data of traffic and bus trips in the past are known and are used to estimate the travel time of the bus, i.e. we transform other traffic quantities to the bus travel time. Travel time estimation is needed if observations of arrival/departure times are not available. In contrast to travel time estimation, travel time prediction is forecasting the travel time from unknown traffic conditions. Most of the models in bus travel time focus on prediction, since travellers need to know the arrival time of the bus in advance.

Here we classify the travel time prediction into two main types: short-term prediction and long-term prediction of travel time. The difference between the two types is mainly their prediction horizons. Short-term travel time prediction aims to find the travel time at the prediction horizon of zero or smaller than a T value of time, in which T could be defined case by case, e.g. 3 hour. Long-term travel time prediction aims to forecast the travel time at prediction horizon longer than T, which could be next day, week or year (Liu, 2008; Van Lint, 2004). The long-term prediction only uses the historical average data of traffic conditions to predict the future state of traffic and forecast the travel time. Figure 2 illustrates the definition of travel time estimation and prediction.

Figure 2: Systematic representation of travel time estimation and prediction definition



3. Benefits of bus travel time estimation and prediction

Bus travel time estimation has the potential to significantly contribute to the operations of the transit providers and traffic managers. Bus travel time estimation could facilitate transit operators in offline management by optimizing their transit schedule and understanding the reliability of their systems (Higatani et al., 2009). The estimated travel time of the bus could be utilised in evaluating the performance of the transit system by several criteria such as average bus travel time, passenger's waiting time or Level of service of the public transport system. By understanding the reliability of the transit system, the bus operators could revise their transit schedule to fit with the actual bus running time and reduce the waiting time of transit riders. Finally, the bus travel time could also be used in the offline evaluation of other transport system such as Public Transport Priority Systems (PTPS) and Signal Controller System at signalised intersections. By investigating the changes in bus travel time, traffic managers could evaluate their systems and provide necessary measures to optimise them.

Bus travel time prediction, on the other hand, could be an important component of a real time traveller information system, and could also reduce anxiety and stress for travellers by helping them to select bus routes with minimum waiting time (Bates, Polak, Jones, & Cook, 2001). Bus travel time prediction models could provide forecast of the running and arrival time of the transit vehicles. This information assists travellers in transport mode choice and route choice, especially in congestion, by providing them the different travelling alternatives in real time. Moreover, prediction of bus arrival time to intersections could facilitate the operation of a PTPS, helping it to trigger the green phase for transit vehicle at the right time.

4. Characteristics of bus in traffic

When trying to estimate and predict the bus travel time, it is important to understand its characteristic in traffic, and how it is different to cars. In this section, the factors affecting both bus and car travel time and the factors affecting only bus will be studied.

4.1 Factors influencing both bus and car travel time

The delay experienced at the signals is a major component of the travel time on signalised urban networks for all the vehicles. It accounts for of 10 to 20% of the total travel time of buses (Sunkari, Beasley, Urbanik, & Fambro, 1995). The traffic signal delay factor is highly correlated with the traffic demand factor in many cases, since the signal delay is larger when the traffic demand is higher and vice versa. Although PTPS strategies are designed to reduce delay experienced by the public transport vehicle at the signals, the accuracy and practical complexity of their implementation introduces significant stochasticity in the delay experienced at the signalised intersection.

4.1.1 Schedule adherence

The main difference between bus and car is the fact that buses have to stop at bus stops. While cars travel by the principle of reducing their travel time, buses have to stay with their schedule. Hence, schedule adherence is one of the factors which we should consider in estimating or predicting the bus travel time. The schedule adherence is the amount of difference in time between the scheduled arrival time and the actual arrival time. It affects the bus travel time on a stop-to-stop basis. If the bus arrives at a bus stop behind the predefined schedule, the driver would probably speed up to adhere to the schedule at the next bus stop. On the other hand, if the bus arrives ahead of the schedule on a bus stop, the driver would slow down or stop longer at stops to stay with the schedule. The schedule adherence factor has been considered as one of the variables of some bus travel time estimation and prediction models (Abkowitz & Engelstein, 1983; Dueker, Kimpel, Strathman, & Callas, 2004; Jeong & Rilett, 2005; Lin & Zeng, 1999).

4.1.2 Dwell time

Dwell time is another factor that only influences the bus travel time and related to the stopping characteristic of the bus. It is defined by TCQSM (TRB, 2003) as the total time the bus has to stop for passenger boarding and alighting. Dwell time has been extensively studied in the literature. It has been considered as a major factor that influencing the bus travel time in some estimation models (Bertini & El-Geneidy, 2004; Bertini & Tantiyanugulchai, 2004; Chakroborty & Kikuchi, 2004; Levinson, 1983; Seneviratne, 1988) and prediction models (Chen, Teng, Zhang, Yang, & Yang, 2012; Jeong & Rilett, 2005; Patnaik, Chien, & Bladikas, 2004; Shalaby & Farhan, 2004).

Dwell time is believed to accumulate up to 26% of the total travel time of buses (Levinson, 1983). According to TCQSM (TRB, 2003), the main factors that influence dwell time includes: passenger demand and loading, bus stop spacing, fare payment procedures, vehicle types, in-vehicle circulation and platform crowding. Each of these factors themselves could have significant influence to the bus dwell time. Levinson (1983) modelled the dwell time as 2.75s per boarding-alighting passenger plus 5s for opening and closing the door. However, when the fare payment procedures are considered, TCQSM defines that pre-payment boarding takes 2.5s, exact change boarding takes 4s, while using a smartcard takes 3.5s for each passenger to get in the bus (TRB, 2003). Analysing bus dwell time further, Milkovits (2008) found that the advantage in dwell time of fare-cards over tickets was dismissed when the bus was full. On the other hand, the crowded platform effects have also been studied by Fernández (2008) and Jaiswal (2009).

4.1.3 Transit operations and strategies

In comparison with cars, buses might be given priority such as dedicated right-of-way, signal priority and some other transit priority strategies such as queue jump, boarding islands, curb extension, yield-to-bus law and parking restriction.

- *Busways and arterial street bus lanes* provide dedicated right-of-way to buses. These facilities are believed to save significantly amount of travel time for buses. Analysing the observed data, Levinson (2003) concluded that busways saved 2 to 3 minutes per mile and arterial street bus lanes saved 1 to 2 minutes per mile compared to normal mixed traffic lanes
- *Transit signal priority (TSP)* is an important part of any PTPS. TSP modifies the regular signal timings at signalised intersections to give priority to transit vehicles. The green phase of the signal is extended or early started prior to the approaching of a transit vehicle. In general, this system provides better accommodation for buses, but still maintains coordinated operation of the signals and overall cycle length (Danaher, 2010). The impacts of TSP to the bus travel time have been proved as positive by several case studies (Baker et al., 2002; Currie & Shalaby, 2008) and a simulation study (Collura, Vlachou, & Mermelstein, 2008)
- *Queue jump* (or queue bypass) is generally a short lane that allows transit vehicles to bypass the traffic queue at intersections (TRB, 2003). The queue jump lane could be used in combination with the signal priority, where the bus is given green phase before the traffic, or without the signal priority. It has the potential to save up to 17% of bus delay (Zhou, Gan, Lue, & Shen, 2008)
- *Boarding islands* as defined by the TCQSM (TRB, 2003) is the bus priority method that locates the bus stop between traffic lanes, so that buses could stay in faster lane without having to change to the left lane before each stop. *Curb extensions* (bus bulbs) has the same purpose with the boarding islands. They are built by extending the curb accommodating the parking lane, so that the bus can still stay at the online travel lane and the clearance time (the re-merge time of the bus to the traffic) is eliminated (TRB, 2003)
- *Yield-to-Bus Law* is the regulation that requires motorists to yield to buses that are re-entering the traffic from the bus stop. Similar to the boarding islands and curb extensions method, these laws also aim to reduce the clearance time of the bus
- *Parking restriction* is the method to allow some space for the above mentioned bus priority operations (TRB, 2003). Parking restriction is mainly applied near a curb side bus stop where the transit agency has to free up some space for the bus to pull out of the traffic and up to the curb for dwelling.

4.1.4 Acceleration/Deceleration time

The bus usually has to decelerate to stop for boarding/alighting of passengers and then accelerate to join in the traffic. This could also explain why buses always tend to use the leftmost lane of arterial roads. Levinson (1983) proposed a linear algorithm to calculate the acceleration/deceleration time based on the number of stop per mile. However, as this amount of time is relatively small, only ranged from 11 to 23s (Levinson, 1983), it is often included in the link running time of the bus when estimating/predicting the bus travel time.

4.1.5 Bus queuing time

Bus queuing time is the amount of time the bus has to wait prior to entering the boarding/alighting position of the bus stop. If the bus frequency is low, the bus stop would be free most of the time and the bus queuing time would be small on average. On the other hand, if the bus frequency is high there would be more chance the bus would have to queue at the bus stop (Fernandez & Tyler, 2005). Bus queuing delay has not been received much attendance from the researchers. In TCQSM (TRB, 2003), the interference between buses at the bus stop was modelled by the factor Failure rate, which is the probability that the bus will have to wait for another bus finish it service time to occupy the boarding/alighting area. Bąk (2010) proposed an analytical statistic method to estimate the bus queuing time. His algorithm based on the ratio of arrival bus intensity and service time at stop. Chen et al. (2012) is the first study that introduced bus queuing phenomenon as a factor in bus dwell time model, according to the authors.

4.1.6 Bus bunching

Bus bunching is possible problem on any high frequency urban transit network. It happens when the bus in front experiences some problems and runs slower than its schedule. As it comes later than the schedule, there are more passengers boarding it and therefore, it tends to get later and later due to the longer dwell times. As most of the passengers boarded the in front bus, the bus following it tends to get earlier and earlier. Bus bunching is the phenomenon when the two buses form a pair of buses of the same line and travel together in a platoon. Bus bunching negatively influences the travel times of all the buses involved in the phenomenon. The bus in front is delayed further by the increase in number of boarding and alighting passenger, while the following bus could also be slowed down if it is not allowed to overtake the leading bus. The bus bunching problem is complicated and has not yet been studied intensively in the literature. Some authors explored the transit system to find a possible strategy for reducing the bus bunching problem (Daganzo 2009; Pilachowski, 2009). To the best of our knowledge, so far no bus travel time prediction and estimation approach in the literature has considered the bus bunching phenomenon as a factor affecting the bus travel time.

4.1.7 Bus stop location and design

As bus stop is the location where the passengers are allowed to board or leave the bus, the bus travel time is also influenced by the design of the bus stop and its location.

Online bus stops are located adjacent to the street curb in one of the three locations: near-side, far-side and mid-block (TRB, 2003). Near-side stop is the bus stop located directly prior to an intersection. Far-side stop is the bus stop located immediately after an intersection. Mid-block stop is located middle of the block between intersections. Different locations of online bus stops could lead to variations in bus lost time due to traffic signal at intersections. Fitzpatrick et al. (1996) studied the difference in bus delay at intersections because of online bus stop locations. The authors believed that in a far-side stop, the bus would have to stop twice: one for the signal and one for the bus stop. In a near-side stop, it could be possible for the bus to stop only once: loading and unloading the passenger at the same time of waiting for the green phase of traffic signal. However, Furth and SanClemente (2006) pointed out that the near-side stop could be suffered from triple-stopping problem, where the bus stopped first at the rear of a passenger car queue that blocked the stop, stopped the second time at the bus stop and third time at the intersection.

Off-line bus stops are located separately from the flow of traffic. They provide higher capacity compared to online stops when there are four or more loading areas and lower capacity when there are less than three loading areas (TRB, 2003). Although at off-line stop, the bus will not block or be blocked by the traffic, the delay when the bus tries to join the traffic after leaving a bus stop needs to be examined. Fitzpatrick and Nowlin (1997) found out that the advantage of a bus bay design (off-line bus stop) to a curbside design (on-line stop) in average vehicle speed ranged from about 0 to 19 km/h in mid-block and far-side bus stop.

5. Critical overview of the literature

The objectives of this section are to provide critical overview of existing bus travel time estimation and prediction approaches in the state-of-the-art. The benefits of them are also studied.

5.1 Bus travel time estimation

Based on our definition, not many studies in the literature could be classified as bus travel time estimation. Bus travel time estimation models are only proposed in cases when the direct measurement is not available, or when the authors estimate the bus travel time for another purpose, e.g. determining the factors that influence the bus running time. Bus travel time has been estimated in the literature either base on bus information itself, or in relationship with the private transport modes. In this study, we classify these two approaches as the two categories for bus travel time estimation.

5.1.1 Bus travel time estimation based on bus information

When estimating the bus travel time based on bus information itself, most of the studies in literature used regression analysis (Abkowitz & Engelstein, 1983; Bertini & El-Geneidy, 2004; Tétreault & El-Geneidy, 2010), statistical method (Gao & Liang, 2011) or simulation method (Seneviratne, 1988) to estimate the bus travel time. The bus information used in these models is the factors that affecting the bus travel time. For instance, they could be the number of boarding and alighting passengers, average running time, the number of times the bus stops (dwells), etc. Many studies aimed to find the impacts of these different factors to the bus travel time (Abkowitz & Engelstein, 1983; Bertini & El-Geneidy, 2004; Tétreault & El-Geneidy, 2010).

5.1.2 Bus travel time estimation in the relationship with cars

Most of the authors when trying to estimate the bus travel time from the data of other modes of transportation studied the differences between bus trips and car trips. They tried to exclude the dwell time and acceleration/deceleration time from the bus travel time to formulate a ratio between actually running time of the bus to the car travel time (Levinson, 1983; McKnight, Paaswell, Ali, Kamga, & Cruz, 1997). As most of the authors aimed to briefly estimate the bus travel time from cars travel time (or the other way around such as in (Bertini & Tantiyanugulchai, 2004) and (Chakroborty & Kikuchi, 2004)), the linear regression method was chosen because of the simplicity and the ability to find the formulation between these two values.

5.2 Bus travel time prediction

In this section, six main approaches of bus travel time prediction in the literature is critically reviewed. The existing methods in bus travel time prediction are classified into six types of models: Historical Average, Time Series, Regression, Kalman Filter, Artificial Neutral Network and other Pattern Recognition methods.

5.2.1 Historical average model

Historical average is a conventional statistical forecasting method that uses the average of link travel time and sometimes also dwell time at transit stop from observed historical data to predict the travel time. Historical average models are based on the assumption that the travel time patterns remain stable over time. Historical average method is suitable for real-time dynamic information systems in providing forecasting travel time data since the algorithms are usually simple and require relatively small computation time. However, the performances of the models are not impressive, especially when compared to some others method in bus travel time prediction (Jeong & Rilett, 2005). Bus travels are affected by the fluctuations in traffic demand, road capacity, driver behaviours and unexpected service interruptions. Hence, the assumption that the travel time patterns of bus trips remain the same over time is a strong assumption over the reality.

5.2.2 Time series model

Time series models are based on the assumption that current and future travel time patterns depend only to the observed historical data (Jeong, 2004). The aim of them is to find out the mechanism of the series of data and forecast the upcoming values (Billings & Jiann-Shiou, 2006). An example of time series model in bus travel time prediction is Suwardo (2010). The authors developed an Autoregressive Integrated Moving Average (ARIMA) model using historical time series data for predicting the bus travel time. The study was carried out in an 82.6 km length bus route in Peninsular, Malaysia. Rajbhandari (2005) added the bus delay propagation (from a Markov chain) to his time series model for predicting bus arrival time.

The strength of time series based models are high computation speed due to simple formulation of the algorithm and the models do not need large number of bus operation variables, only time related data are needed. The models could be built with only historical data, without real-time observations. However, the main disadvantage of this type of model is the averaging of input data over time. The predictions of travel time tend to concentrate on the trend of the historical travel time data and miss the extremes, e.g. the short-term fluctuations due to signalized control (Bhaskar, 2009) Variations in the historical data set itself, and also variations in the relationship between it and the current traffic patterns could dramatically affect the prediction in a negative way (Abdelfattah & Khan, 1998). Moreover, the performance of these models is highly dependent on the quality of the historical data set, which is not always available (Chien, Ding, & Wei, 2002).

5.2.3 Regression model

Regression-based method has been used by many authors in bus travel time prediction (Abdelfattah & Khan, 1998; Jeong & Rilett, 2005; Lin & Zeng, 1999; Patnaik, et al., 2004). Besides the ability of accurately predicting the travel time of buses, this type of model can also estimate the impact of each parameter to the bus travel time. Regression models require the mathematical function between independent variables and dependent variables. This requirement limits the application of regression models in many cases where the variables are correlated (Chang, Park, Lee, Lee, & Baek, 2010; Kalaputapu & Demetsky, 1995). Regression models are suitable for forecasting of travel time with their relatively high accuracy and large variance of inputs. Moreover, they can also facilitate the transit managers to explore the influence of independent variables to the bus travel time.

5.2.4 Kalman Filter-based algorithm

Kalman Filter is an algorithm which could be used to predict the future state of the dependent variables. Unlike the historical average, time series and regression models, the Kalman Filter-based algorithms only estimate state from the previous time steps. The current measurements are also needed. Thus, these algorithms are not as dependent on historical data as in the aforementioned models. Reinhougt and Velastin (1997) were the first authors who successfully developed dynamic Kalman Filter algorithms to predict bus travel time. Shalaby & Farhan (2004) improved the algorithm from Reinhougt and Velastin (1997) by replacing the previous predicted value p(k) with the actual running time.

Kalman Filter is a promising approach for predicting the bus travel time. The estimation could be updated when the new observations are ready. The deviation of prediction could be reduced by apply the algorithm, using the new observation data (Chang, et al., 2010). The functions also do not require high computation time and large historical data, which is feasible for applying in a dynamic predicting system (Chen, Liu, Xia, & Chien, 2004). In predicting one or two time steps ahead, the Kalman Filter performs better than the historical profile, real-time profile, neutral network, exponential smoothing method (Park & Rilett, 1999). However, in order to update the estimation, the Kalman Filter-based algorithms require real time observations at each time interval, as the main input (Chen & Chien, 2001). Therefore, all the bus in the system need to be equipped with location and time measurement equipment, which could be too expensive to some certain cities. Moreover, this type of model requires real-time data of travel time observations. The Kalman Filter algorithm could face some difficulties if the time series data changes dramatically. In the same Park and Rilett study (1999), the performance of the Kalman Filter dropped lower than the historical profile and neutral network when predicting more than 4 time steps ahead.

5.2.5 Artificial Neutral Network

Artificial Neutral Network (ANN) was demonstrated as a potential method for predicting the traffic conditions (Chen, et al., 2004; Chien, et al., 2002; Kalaputapu & Demetsky, 1995; Shalaby, Lee, Greenough, Hung, & Bowie, 2006; Smith & Demetsky, 1994). Chien et al.(2002) developed two ANNs models for predicting the bus arrival times with the case at New Jersey, U.S. Chen et al. (2004) introduced an ANN model based on historical APC data, bus operation and weather data. Considering the influence of non-recurrent situations which could affect the bus trip, the authors also developed a dynamic algorithm based on the Kalman filter for adjusting the prediction from the ANN model. Mazloumi (2011) developed the first models that included real-time traffic flow data in their two ANN models. The traffic flow data was obtained from the Sydney Coordinated Adaptive Traffic Systems (SCATS) loop detectors in Melbourne, Australia.

ANN models are believed by some authors as a more accurate prediction method over other techniques such as Kalman Filter, historical data based regression models (Chang, et al., 2010; Jeong & Rilett, 2005; Park & Rilett, 1999). ANN models could also be incorporated with other method such as time series models, statistical models (Chien, et al., 2002) or Kalman Filter models (Chen, et al., 2004). ANN models are suitable to find complex nonlinear relationship between the dependent variable bus travel time and the independent variables that influence the travel time. ANN can be built without the need of specifying the exact formulation, unlike the other models which have been described here. However, the input-output function could not be found by using ANN (Bhaskar, 2009; Chien, et al., 2002), which could be a hindrance for the understanding of the transit system. The overfitting problem of ANNs could also be a weakness of the method (Chen, et al., 2004). To avoid the problem and get optimal prediction results, training an ANN requires extensive knowledge of the technique in selecting input variables, hidden layers, learning rate and momentum (Yu, Yang, & Yao, 2006). Many authors also discussed that ANN models requires a large historical database for training and a lengthy training procedure (Mazloumi, et al., 2011; Yu, et al., 2006).

5.2.6 Pattern recognition

This section explores some other approaches for predicting the bus travel time based on other pattern recognition technique rather than the ANN. These methods utilise the large quantities of data to study the pattern of the historical data and they are promising techniques in bus travel time prediction. Yu (2006) proposed a support vector machines (SVM) model for predicting the bus arrival time in Dalian, China. SVM is alternative types of neutral networks, which is based on the statistical learning theory, introduced in 1995 and further described by some recent researches (Burges, 1998; Cortes & Vapnik, 1995; Horváth, 2003; Vapnik, 1999). Chang et al. (2010) proposed a k-NN model for predicting the bus travel time in Seoul, Korea. The study has proved that the k-NN method is suitable for real-time prediction of the bus travel time. Vu and Khan (2010) developed a locally weighted scatter smoothing (LOWESS) model with data collected from AVL and APC systems in Ottawa, Canada. This type of model was first introduced by Cleveland (1979) it is amended from the extreme point problem in locally fitted polynomial smoothing by focusing more on the local points.

Pattern recognition is a promising method for predicting bus travel time, because it usually does not require a mathematic pre-defined function. Instead of relying on a set of historical data, pattern recognition techniques also use historical data, but search through it in operation, find the closest data to the current input and provide the prediction. The disadvantage of them may be the time and computation required for the training of the model. They are also relatively more difficult to share to the academic than the regression, Kalman Filter or historical average methods because there is no exact function of the model.

6. Discussion

The previous section critically reviews the factors influencing bus travel time and existing literature in bus travel time estimation and prediction. The following issues have been derived from the understanding of the bus operations and bus travel time estimation and prediction models.

Estimation and prediction of bus travel time is often sophisticated due to the randomness of various factors affecting the bus journey. Apart from the factors influencing both bus and automobiles travel time, buses also have some exclusive characteristics (as buses have to stop, they prefer to use the left most lane, they are heavy vehicles and they are more adhered to the speed limits) and be affected by some different factors than the cars. Although these factors are extensively studied in the state-of-the-art, there is limited understanding on some factors as follows:

- Traffic signal delay: Delay at the signalised intersection is stochastic, and depends on the traffic demand, transit priority measure, and the location of a bus stop close to the intersection. Traffic signal delay is rarely considered in bus travel time prediction and estimation models because the arrival of the bus to the intersection and the green phase of the signal are both stochastic
- Bus queuing time: The increase in bus demand and reduction in bus on-time performance results in congestion at the bus stops, mainly at the main stops where different lines compete for the loading area. The queuing of the buses at the stop not only increases the delay of the stopping buses but also for non-stopping buses. The bus queuing time due to the congestion of bus approaching a bus stop has also not been thoroughly studied in the literature, but it is becoming a serious problem for large transit system with many buses using the same stop
- Bus bunching: Bus bunching has a direct impact on the dwell time and schedule adherence of the buses. The leading bus in bunched buses will service most of the passengers at stops and the following bus could be slowed down after the leading one, or take the chance to overtake and swap the schedule with the first one. Bus bunching clearly indicates that at most only one of the buses in the bunch could stay on schedule. Bus bunching phenomenon has not yet been considered in bus estimation and prediction models.

The problem of bus travel time prediction has been extensively studied in the literature. Several methods have been proposed, including Historical Average (Chen, Yang, Zhang, & Teng, 2011; Jeong & Rilett, 2005), Time Series (Rajbhandari, 2005; Suwardo, et al., 2010), Regression (Abdelfattah & Khan, 1998; Lin & Zeng, 1999; Patnaik, et al., 2004), Kalman Filtering (Reinhoudt & Velastin, 1997; Shalaby & Farhan, 2004; Wall & Dailey, 1999), Artificial Neutral Network (Chen, et al., 2004; Chien, et al., 2002; Kalaputapu & Demetsky, 1995; Mazloumi, et al., 2011) and Pattern Recognition (Chang, et al., 2010; Vu & Khan, 2010; Yu, et al., 2006). The performances of these approaches have been reported as relatively positive by their authors. However, most of the models assumed that all the bus stops and bus links in their models were the same in designs and locations by considering the same factors for all the links, even though there could be different types of bus stops and bus links within a bus route in reality. The Brisbane transit network is an example of these transit systems, where different types of bus stop designs and locations are combined. Within a bus route, there could be links in which buses are operating in shared lanes, median bus lanes and busways. Therefore, **there is a need for an integrated bus travel time prediction methods that could consider each link between bus stops separately in different models. Each model considers different factors influencing the bus travel time.** As each bus link needs a dedicated model for prediction, a method which requires low computation time but still provides high accuracy forecasts of bus travel time is needed. Secondly, **there is limited knowledge on the relationship between multimodal (bus and car) travel times on the urban networks.** A comprehensive understanding of the multimodal travel time is an important performance measure for managing the traffic on the networks. Bus travel time could be used to estimate car travel time and vice versa.

Many metropolitan transit networks are monitored so that traffic managers could maintain and improve the service quality. The data from bus detectors, on-board smart card equipment, AVL/APC data and beacon detectors at bus stops are the main sources. The Table 1 shows the generally available data of bus travels.

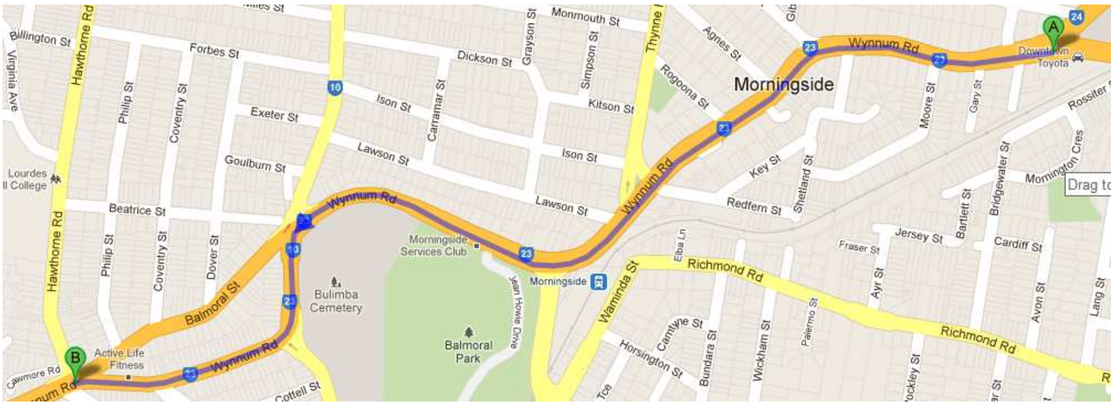
Table 1: Available bus data sources

Source	Characteristic	Arrival/ departure time at each bus stop	Bus ID	Timestamp at detector location (generally intersection)	Comments
Bus detector	Bus ID can be captured	✕	✔	✔	Bus travel time between VID locations could be calculated, but not travel time between stops
	Bus ID cannot be captured	✕	✕	✔	These detectors only provide a timestamp when a bus is observed, and are mainly for PTPS purpose
Smart card	With bus stop location data	✔	✔	✕	Arrival/departure time to each bus stop could be estimated from the first and last touch to the smartcard equipment
	No bus stop location data	✕	✔	✕	Need data clustering and data mining to group the samples at each bus stop
AVL/ APC Data	Exact location of the bus (GPS)	✔	✔	✔	Full information on bus location but limited in relationship with cars
Beacon detector	Located at bus stop	✔	✔	✕	Timestamp of bus arrival/departure time at each stop, the data is only updated at bus stops

Most of the studies utilise the data sources related to the GPS technology (e.g. AVL or APC systems) or beacon detectors at bus stops. This type of data provides very precise location and timestamp of the bus, but limited in multimodal understanding of the traffic systems. **There is a need for further exploration on other data sources such as the Vehicle Identification (VID) and utilise the data of not-in-service bus as probe vehicle for better understanding of the relationship between bus and cars travel time.** The VID scanners are installed at intersections, where other detectors such as Bluetooth, loop detectors could be used for collecting cars arrivals. The comparison between bus travel time and cars travel time could gain some understanding on the relationship between them.

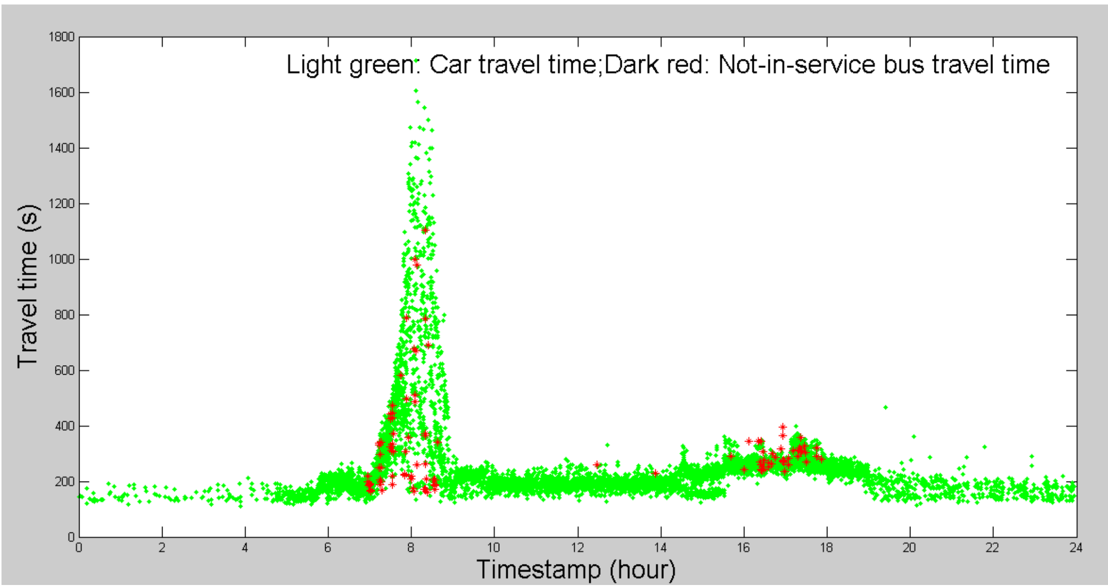
As an example, the relationship between bus and car travel time in an urban corridor in Brisbane has been explored. The Figure 3 shows our study site: the arterial between Hawthorn Rd/Wynnum Rd and Junction Rd/Wynnum Rd in Brisbane.

Figure 3: The study site between two intersections: Junction Rd/Wynnum Rd and Hawthorn Rd/Wynnum Rd, Brisbane, Queensland.



Intersection A and B have Bluetooth scanners for collecting the MAC address and VID scanners for collecting the bus number, route, service number of each bus pass by the intersections. The analysis has been carried out for all the weekdays of July 2011 on inbound traffic. Because both MAC and VID address is unique of each vehicle, a simple algorithm has been used for matching a vehicle upstream with the same vehicle downstream for calculating the travel time between A and B. Figure 4 illustrates the travel time of car (light green dots) and not-in-service bus (dark blue dots) on our studied corridor.

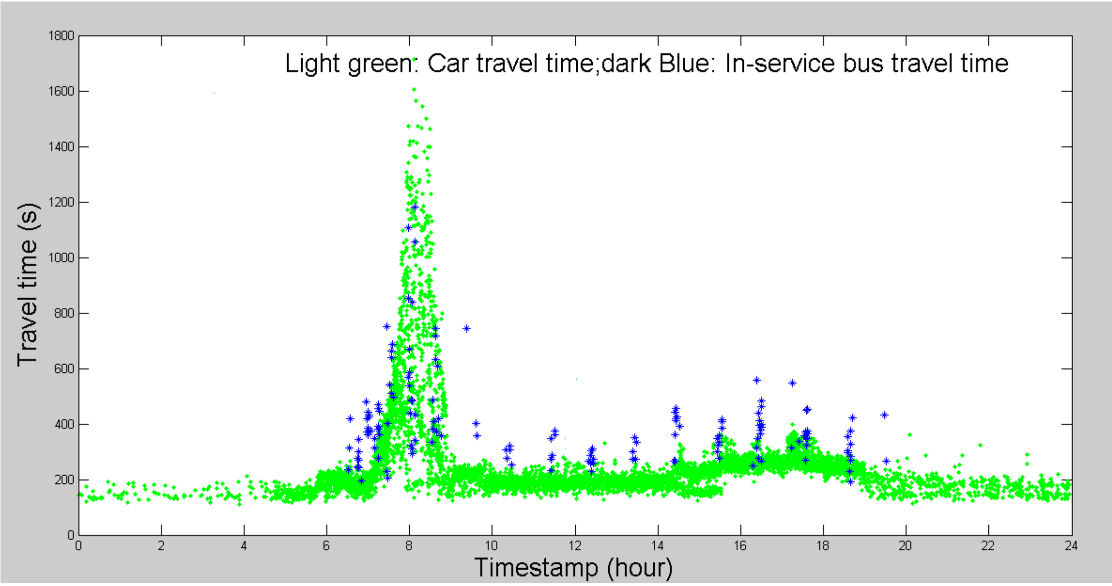
Figure 4: Cars travel time from Bluetooth data and Not-in-service Bus travel time from VID data on weekdays of July 2011



The not-in-service buses are the buses which do not stop at bus stops within the study area. They are basically operating in a similar way to a car, except in the differences in their operating capabilities as heavy vehicles. In the Figure 4, it could be noted that the not-in-service bus travel times are within the range of the cars travel times. Hence, data of not-in-service bus travel time could be used to estimate car travel time. Further analysis is needed for validating this hypothesis and a formulation of the relationship between these two should be considered.

Between the intersections A and B there is only one bus servicing. The Figure 5 shows the comparison between car travel time (light green dots) and in-service bus travel time (dark blue dots).

Figure 5: Cars travel time from Bluetooth data and In-service Bus travel time from VID data on weekdays of July 2011



The following two patterns could be identified from the Figure 5:

- The in-service bus travel times are usually higher than the car travel times, but getting closer to the car travel times in peak hours. In morning peak hours, the in-service buses seem to spend less time for travelling between our two studied intersections than the cars, even though they have to stop at bus stops. The reason for that is a temporal bus lane on a short section between the two intersections, where only buses could use it between 7-9 AM
- During the off-peak period, look at the timestamp axis we could find some groups of in-service bus in which each group forms a horizontal line of travel time. The time difference between each group is around 1 hour and they appear from 9 AM to 5 PM. The reason for that could be the fact that the bus could stay with the schedule during off peak hours. As there is not much congestion at upstream intersection during off peak hours, every day a bus shows up at the same time at upstream intersection. As the bus travel from upstream intersection to downstream intersection the travel time could be difference, but the arrival time to the upstream intersection is mostly the same every day, and then these groups are formed. The sample with smallest travel time in each group always stays within the largest cars travel time samples. Hence, these in-service bus samples could also be used to estimate cars travel time, and vice versa.

7. Conclusion

This paper provides the fundamental understanding of the benefits and issue of bus travel time estimation and prediction. The existing approaches in bus travel time estimation and prediction are reviewed in this paper. The literature review revealed that there is a need for a comprehensive study on the relationship between bus and car travels. Most of the studies explore the utility of the GPS-based data source such as AVL or APC data. Other data sources such as VID and the data of not-in-service buses should be considered to obtain better understanding of the bus operation and its relationship with other transport modes. A short analysis of an urban corridor in Brisbane showed that analysing VID and Bluetooth data could be useful for the estimation of the car travel time from the bus travel time and vice versa. From these understanding, an integrated method which could reflect the differences in types of bus links and bus stops need to be proposed. Regarding the factors that affect bus travel time, the authors have identified that the traffic signal delay, bus queuing time and bus bunching phenomenon need more consideration in bus travel time prediction and estimation models.

References

Abdelfattah, A. M., & Khan, A. M. (1998). Models for predicting bus delays *Transit: Bus, Paratransit, Rural, Intermodal, Rail, Commuter and Intercity Rail, Light Rail* (pp. 8-15).

Abkowitz, M. D., & Engelstein, I. (1983). Factors affecting running time on transit routes. *Transportation Research Part A: General, 17*(2), 107-113. doi: 10.1016/0191-2607(83)90064-x

Bąk, R. (2010). Simulation Model of the Bus Stop. [10.2478/v10174-010-0001-6]. *Archives of Transport, 22*(1), 5-25.

Baker, R. J., Collura, J., Dale, J., Head, L., Hemily, B., Ivanovic, M., . . . Smith, L. (2002). An overview of transit signal priority. In J. Chang (Ed.), (Prepared by the ATMS, APTS of ITSA ed.).

Bates, J., Polak, J., Jones, P., & Cook, A. (2001). The valuation of reliability for personal travel. *Transportation Research Part E: Logistics and Transportation Review, 37*(2–3), 191-229. doi: 10.1016/s1366-5545(00)00011-9

Bertini, R. L., & El-Geneidy, A. M. (2004). Modeling Transit Trip Time Using Archived Bus Dispatch System Data. *Journal of Transportation Engineering, 130*(1), 56-67.

Bertini, R. L., & Tantiyanugulchai, S. (2004). Transit buses as traffic probes: Use of geolocation data for empirical evaluation. *Transportation Research Record: Journal of the Transportation Research Board, 1870*(-1), 35-45.

Bhaskar, A. (2009). *A methodology (CUPRITE) for urban network travel time estimation by integrating multisource data.* PhD, Ecole polytechnique fédérale de Lausanne EPFL. Retrieved from http://books.google.com.au/books?id=9FRHcgAACAAJ

Billings, D., & Jiann-Shiou, Y. (2006, 8-11 Oct. 2006). *Application of the ARIMA Models to Urban Roadway Travel Time Prediction - A Case Study.* Paper presented at the Systems, Man and Cybernetics, 2006. SMC '06. IEEE International Conference on.

Burges, C. J. C. (1998). A Tutorial on Support Vector Machines for Pattern Recognition. *Data Min. Knowl. Discov., 2*(2), 121-167. doi: 10.1023/a:1009715923555

Chakroborty, P., & Kikuchi, S. (2004). Using bus travel time data to estimate travel times on urban corridors. *Transportation Research Record: Journal of the Transportation Research Board, 1870*(-1), 18-25.

Chang, H., Park, D., Lee, S., Lee, H., & Baek, S. (2010). Dynamic multi-interval bus travel time prediction using bus transit data. *Transportmetrica, 6*(1), 19-38. doi: 10.1080/18128600902929591

Chen, G., Teng, J., Zhang, D., Yang, X., & Yang, X. (2012). *Dwell Time Estimation with Consideration of Bus Bunching.* Paper presented at the Transportation Research Board 91st Annual Meeting, U.S.

Chen, G., Yang, X., Zhang, D., & Teng, J. (2011). *Historical Travel Time Based Bus-Arrival-Time Prediction Model.* Paper presented at the ASCE Conf. Proc. ICCTP 2011: Towards Sustainable Transportation Systems. Proceedings of the 11th International Conference of Chinese Transportation Professionals.

Chen, M., & Chien, S. (2001). Dynamic Freeway Travel-Time Prediction with Probe Vehicle Data: Link Based Versus Path Based. [10.3141/1768-19]. *Transportation Research Record: Journal of the Transportation Research Board, 1768*(-1), 157-161.

Chen, M., Liu, X., Xia, J., & Chien, S. (2004). A Dynamic Bus-Arrival Time Prediction Model Based on APC Data. *Computer-Aided Civil and Infrastructure Engineering, 19*(5), 364-376. doi: 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. [Article]. *Journal of Transportation Engineering-Asce, 128*(5), 429-438. doi: 10.1061/(asce)0733-947x(2002)128:5(429)

Cleveland, W. S. (1979). Robust Locally Weighted Regression and Smoothing Scatterplots. *Journal of the American Statistical Association, 74*(368), 829-836.

Collura, J., Vlachou, K., & Mermelstein, A. (2008). Improving Transportation Mobility, Safety, and Efficiency: Guidelines for Planning and Deploying Traffic Signal Priority Strategies. Published by UMass-Amherst (Final Report).

Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine learning, 20*(3), 273-297.

Currie, G., & Shalaby, A. (2008). Active Transit Signal Priority for Streetcars: Experience in Melbourne, Australia, and Toronto, Canada. *Transportation Research Record: Journal of the Transportation Research Board, 2042*(-1), 41-49.

Daganzo , C. F. (2009). A Dynamic Approach to Eliminate Bus Bunching.

Danaher, A. R. (2010). Bus and rail transit preferential treatments in mixed traffic (Vol. Volume 83 of TCRP synthesis): National Research Council . Transportation Research Board. Transit Cooperative Research Program. United States. Federal Transit Administration. Transit Development Corporation.

Dueker, K. J., Kimpel, T. J., Strathman, J. G., & Callas, S. (2004). Determinants of bus dwell time. *Journal of Public Transportation, 7*(1), 21-40.

Fernández, R., De Los Angeles Del Campo, M., & Swett, C. (2008). *Data collection and calibration of passenger service time models for the Transantiago system.* Paper presented at the European Transport Conference.

Fernandez, R., & Tyler, N. (2005). Effect of Passenger–Bus–Traffic Interactions on Bus Stop Operations. *Transportation Planning and Technology, 28*(4), 273-292. doi: 10.1080/03081060500247747

Fitzpatrick, K., Hall, K., Perkinson, D., & Nowlin, L. (1996). TCRP Report 19: Guidelines for the Location and Design of Bus Stops: TRB, National Research Council, Washington, DC.

Fitzpatrick, K., & Nowlin, R. L. (1997). Effects of Bus Stop Design on Suburban Arterial Operations. *Transportation Research Record: Journal of the Transportation Research Board, 1517*.

Furth, P. G., & SanClemente, J. L. (2006). Near side, far side, uphill, downhill : impact of bus stop location on bus delay. *Transportation research record.*(1971), 66-73.

Gao, L., & Liang, H. (2011, 22-24 July 2011). *Mining travel time from smart card fare data.* Paper presented at the Control Conference (CCC), 2011 30th Chinese.

Higatani, A., Kitazawa, T., Tanabe, J., Suga, Y., Sekhar, R., & Asakura, Y. (2009). Empirical Analysis of Travel Time Reliability Measures in Hanshin Expressway Network. *Journal of Intelligent Transportation Systems: Technology, Planning, and Operations, 13*(1), 28-38. doi: citeulike-article-id:6098112

Horváth, G. (2003). *CMAC: Reconsidering an old neural network.* Paper presented at the Proc. Intell. Control Syst. Signal Process.

Jaiswal, S., Bunker, J. M., & Ferreira, L. (2009). *Modelling the relationships between passenger demand and bus delays at busway stations*. Paper presented at the TRB 88th Annual Meeting.

Jeong, R. (2004). *The prediction of bus arrival time using Automatic Vehicle Location Systems data.* Texas A&M University. Retrieved from http://hdl.handle.net/1969.1/1458

Jeong, R., & Rilett, L. (2005). Prediction model of bus arrival time for real-time applications *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.

Kalaputapu, R., & Demetsky, M. J. (1995). *Modeling schedule deviations of buses using automatic vehicle location data and artificial neural networks*: Transportation Research Board.

Levinson, H. (1983). Analyzing transit travel time performance. *Transportation Research Record*(915).

Levinson, H., Zimmerman, S., Clinger, J., Rutherford, S., Smith, R. L., Cracknell, J., & Soberman, R. (2003). Tcrp report 90: Bus rapid transit, volume 1: Case studies in bus rapid transit. *Transit Cooperative Research Program. Washington DC: Transportation Research Board*.

Lin, W. H., & Zeng, J. (1999). Experimental Study of Real-Time Bus Arrival Time Prediction with GPS Data. [10.3141/1666-12]. *Transportation Research Record: Journal of the Transportation Research Board, 1666*(-1), 101-109.

Liu, H. (2008). *Travel time prediction for urban networks.* PhD Thesis, Netherlands Research School for Transport, Infrastructure and Logistics, TU Delft, Netherlands.

Mazloumi, E., Rose, G., Currie, G., & Sarvi, M. (2011). An Integrated Framework to Predict Bus Travel Time and Its Variability Using Traffic Flow Data. *Journal of Intelligent Transportation Systems, 15*(2), 75-90. doi: 10.1080/15472450.2011.570109

McKnight, C., Paaswell, R., Ali, M., Kamga, C., & Cruz, W. (1997). Impact of Congestion on New York Bus Service: Prepaired for MTA New York City Transit. University Transportation Research Center, New York.

Milkovits, M. (2008). Modeling the Factors Affecting Bus Stop Dwell Time: Use of Automatic Passenger Counting, Automatic Fare Counting, and Automatic Vehicle Location Data. [10.3141/2072-13]. *Transportation Research Record: Journal of the Transportation Research Board, 2072*(-1), 125-130.

Park, D., & Rilett, L. R. (1999). Forecasting freeway link travel times with a multilayer feedforward neural network. *Computer‐Aided Civil and Infrastructure Engineering, 14*(5), 357-367.

Patnaik, J., Chien, S., & Bladikas, A. (2004). Estimation of Bus Arrival Times Using APC Data. *Journal of Public Transportation, Vol. 7, No. 1, 2004*.

Pilachowski, J. M. (2009). *An Approach to Reducing Bus Bunching.* PhD Thesis, University of California at Berkeley

Rajbhandari, R. (2005). *Bus arrival time prediction using stochastic time series and Markov chains.* PhD, New Jersey Institute of Technology.

Reinhoudt, E. M., & Velastin, S. (1997). *A dynamic predicting algorithm for estimating bus arrival time.* Paper presented at the 8th IFAC/IFIP/IFORS symposium on transportation systems, 16-18 June, Chania, Crete, Greece.

Seneviratne, P. N. (1988). Simulation of fixed route bus travel time. *Journal of Advanced Transportation, 22*(1), 39-53. doi: 10.1002/atr.5670220104

Shalaby, A., & Farhan, A. (2004). *Bus Travel Time Prediction Model for Dynamic Operations Control and Passenger Information Systems*. Paper presented at the TRB 2003 Annual Meeting CD-ROM, Washington D.C., 2003.

Shalaby, A., Lee, J., Greenough, J., Hung, S., & Bowie, M. D. (2006). Development, evaluation, and selection of advanced transit signal priority concept directions. *Journal of Public Transportation, 9*(5), 97.

Smith, B. L., & Demetsky, M. J. (1994, 2-5 Oct 1994). *Short-term traffic flow prediction models-a comparison of neural network and nonparametric regression approaches.* Paper presented at the Systems, Man, and Cybernetics, 1994. 'Humans, Information and Technology'., 1994 IEEE International Conference on.

Sunkari, S. R., Beasley, P. S., Urbanik, T. I., & Fambro, D. B. (1995). *Model to evaluate the impacts of bus priority on signalized intersections*. Washington, DC, ETATS-UNIS: National Research Council.

Suwardo, Madzlan, N., & Ibrahim, K. (2010). ARIMA models for bus travel time prediction. *The Journal of the Institution of Engineers, Malaysia, vol. 71(2), 2010, pages 49-58, vol. 71(2), 2010*, 9.

Tétreault, P. R., & El-Geneidy, A. M. (2010). Estimating bus run times for new limited-stop service using archived AVL and APC data. [Journal]. *Transportation Research Part A: Policy and Practice, 44*(6).

TRB. (2003). *Transit Capacity and Quality of Service Manual*.

Van Lint, H. (2004). *Reliable travel time prediction for freeways.* PhD Thesis, Netherlands TRAIL Research School, T U Delft, Netherlands.

Vapnik, V. N. (1999). An overview of statistical learning theory. *Neural Networks, IEEE Transactions on, 10*(5), 988-999.

Vu, N. H., & Khan, A. M. (2010). Bus running time prediction using a statistical pattern recognition technique. *Transportation Planning and Technology, 33*(7), 625-642. doi: 10.1080/03081060.2010.512225

Wall, Z. R., & Dailey, D. J. (1999). *An algorithm for predicting the arrival time of mass transit vehicles using automatic vehicle location data*. Paper presented at the 78th Annual Meeting of the Transportation Research Board CD-ROM, National Research Council, Washington D.C.

Wilkie, C., & Van Ewijk, F. (2009). Public Transport & Climate Change *TTF Transport Position Paper. Tourism & Transport Forum, Australia.*

Yu, B., Yang, Z. Z., & Yao, B. Z. (2006). Bus arrival time prediction using support vector machines. *Journal of Intelligent Transportation Systems, 10*(4), 151-158. doi: 10.1080/15472450600981009

Zhou, G., Gan, A., Lue, D., & Shen, L. D. (2008). *Design of Transit Signal Priority at Intersections with Queue Jumper Lanes.* Paper presented at the Transportation Research Board 87th Annual Meeting, Washinton D.C, USA.