A SURVEY ON APPLICATION OF ARTIFICIAL INTELLIGENCE FOR BUS ARRIVAL TIME PREDICTION

SEYED MOJTABA TAFAGHOD SADAT ZADEH, TONI ANWAR, MINA BASIRAT

1 Faculty of Computer Science and Information Systems, Universiti Teknologi Malaysia (UTM), Johor Bahru, Johor, Malaysia
2 Faculty of Electrical, IT & Computer Science, Qazvin Islamic Azad University (QIAU), Qazvin, Iran
E-mail: mojtaba.sadat@hotmail.com, tonianwar@utm.my, minabasirat@gmail.com

ABSTRACT

With the intention of satisfying mobility requirements for trustworthy, healthy and secure transport, there are more considerations on the establishment of intelligent transport systems (ITS) currently. Advanced traveller information systems (ATIS), as a part of ITS, is to provide travel time information as precisely as possible. Basically, there are reasons leading to delay in bus arrival time, e.g. traffic jam, ridership distribution, and climate situation. Consequently, these issues impress on growing travellers waiting time, postponement in timetable, rise in transit’s expense and private vehicles' uses, dissatisfaction of passengers and reduction of passengers, providing of precise transit travel time information are significant since it will result in further transit passages and upsurge the acquiescence of passengers. In this paper, we first explore the importance of arrival time for passengers and present a new taxonomy of bus arrival prediction models, and then review some recent works. Finally, summary of the main technologies illustrate big picture of the studies.

Keywords: Bus Travel Time Prediction, Intelligent Transportation Systems (ITS), Advanced Traveller Transportation Systems (AITS), Kalman filtering, machine learning, statistical methods, ANN

1. INTRODUCTION

In the 21st century, as we are faced with the extreme expansion of modern society and urbanization, transport has been one of the pillars of civilization [1,2]. Transport or transportation is the movement of people, animals and goods from one location to another. Transport is important since it enables trade between peoples, which in turn establishes civilizations. IBM research [3] in over 50 developed and developing world cities reveals that most people strive for cleaner, less congested cities and improved traffic flow, primarily through increased use of enhanced public mass transit systems.

Generally, passenger transport may be public or private. In fact, in developed countries still most of the people use private vehicles. Also, in developing countries the level of vehicle ownership is rising at a faster rate [3,4]. However, to come up such problems is that we need to make people interest to use public transport by offering enhanced services.

So far, different solutions for encouragement of people to use public transportation systems (i.e., buses) have been forwarded. One of them is providing travelers with reliable travel information by means of Advanced Public Transport System (APTS) and Advanced Traveler Information System (ATIS), which are the primary key components in Intelligent Transportation Systems (ITS) [5,6]. APTS/ATIS applications has the ability of making people more interested to change from private vehicles to public transport by providing convenient service [7]. One such ATIS application includes pre-trip and real-time passenger information systems, arrival time notification systems. While bus services play an essential role in the provision of public transport, a diminution in travel-time erraticism consequences to cuts passengers’ anxiety and stress caused by uncertainty in decision making about departure time and route choice [8].

As has been explained on [9], a precise travel time prediction is valuable for both travelers and logistic operators, while its aid outcomes in evading congested route to lessen transport outlays and upsurge facility excellence. For traffic managers, travel time information is a significant index of traffic system operation. Especially, travel-time data is critical for pre-trip and en route information which is highly informative to drivers and travelers.
to have smart choice or design better schedules [10]. According to Mazloumi [11,8], buses are contingent on environments’ situation; to provide precise forecasts of bus travel times, there is a wide range of determinants that must be considered, such as passenger demand, weather conditions, road condition, and most significantly, traffic flow at signalized intersections. These issues influence on public transportation system efficiency consequential in, growing passengers waiting time, worsening of schedule adherence, irregular intermodal transfers, raising operation’s cost, traffic delays, etc. [12]. Figure 1 demonstrates the major demand and capacity which impact directly on public transport efficiency.

![Figure 1. Demand And Capacity Related Determinants Of Bus Travel-Time Variability][11]

After a brief introduction on the importance of the bus arrival time for advance public transportation systems, this paper intend to represent the diversity of the approaches taken to resolve the difficulties related to bus schedule. The rest of the paper is structured as follows: Section 2, an overview of bus arrival time prediction models is addressed. Section 3, a table consist of overview showed as conclusion.

2. TAXONOMY OF BUS TRAVEL TIME PREDICTION MODELS

In previous researches, a variety of studies have been focused to address the bus arrival time prediction drawback. A number of approaches had been proposed for travel time prediction over the years; these existing methodologies have been categorized into five mostly wide types used of prediction models: (a) historical data based models, (b) statistical models, (c) Kalman filtering model, (d) machine learning models, and (e) hybrid models. Obviously, there are a variety of classifications for bus arrival time prediction models which were introduced in other studies. For instance, Lee [9] grouped them into four categories: regression method, time series estimation method, hybrid of data fusion or combinative model and artificial intelligence method like neural network; Sun [13] stated them into three types of prediction models: models based on historical data, multi-linear regression models, and artificial neural network models; Jin in [14] refers these methods mainly included nonparametric statistical methods, time series model, neural-network models and others; the most similar to our aforementioned categories can be found in [7] that Zhu explained the methods such as historic and real-time approaches, machine learning techniques (Artificial Neural Network (ANN), support vector machines), model based approaches (Kalman filtering) and statistical methods (regression analysis, time-series) for the Prediction of bus arrival time. However, we will pursue our five categories in the discussion presented on the next subsections.

2.1. Historical Data Based Models

Commonly, a bus arrival time prediction algorithm is based on static and dynamic information. Static information, donates to Historical Data here, provides from bus schedule information, recurrent traffic circumstances and average dwell time. While, dynamic information, donates to real-time models here, includes: real-time bus location data, delay at bus stops, and current traffic circumstances [15]. In this type of prediction models, it gives the current and future bus travel time from the historical travel time of the same bus of the previous journeys on the same time period. Historical approach predicts the travel time at a particular time as the average travel time for the same period over different days. The results of these models are satisfiable under expected circumstances, but the precision of these models in unexpected congestion and delay situations, is seriously decreased [16]. The real-time approach predicts the next time interval travel time to be the same as the present travel time [7].

In these models, we are faced with a cyclical traffic patterns and the ratio of the historical travel time on a specific route to the current travel time stated in real-time will repeat again. The procedure needs an extensive set of historical information and it is not easy to establish the system in a new setting [17]. In real-time models, the almost recently observed journey travel times will remain the same in the future. There is a requirement to provide a precise auto bus arrival time estimation system by real-time data based on diverse traffic circumstances. Padmanaban et al [16] proposed a method that uses real-time data from diverse traffic circumstances and plus consideration of delays for estimation. It fundamentally was found the future

---

www.jatit.org
travel time to be the equal as the current one. This proposed method is not identical when data is unavailable of (i.e. loss in reception, equipment failure, etc.) [16]. Chen et al developed a prediction algorithm that combined these two approaches. Primarily, from a historical database, an estimated travel time was obtained. It was adjusted as real-time location data are obtained [18].

Lin and Zeng in [19], used GPS-based bus location data, and other information as input data, including bus schedule information, bus delay patterns, and bus stop type information (a time-check stop vs. a regular stop) However, they did not consider the effect of traffic congestion and dwell time at bus stations. Also they found that treating time-check stops and regular stops differently would make a big difference in the performance of the algorithm. Kidwell [20] presented an algorithm for predicting bus arrival times based on real-time vehicle location. The algorithm worked by dividing each route into zones and recording the time that each bus passed through each zone. Predictions were based on the most recent observation of a bus passing through each zone. However, this algorithm was not suitable for large cities where both travel time and dwell time could be subjects to large variations.

On the whole, these models are reliable only when the traffic pattern in the area of interest is relatively stable. One of their main limitations is that it requires an extensive set of historical data, which may not be available in practice, especially when the traffic pattern varies significantly over time.

2.2. Statistical Models

According to Karlaftis [21], Glymour claims that "statistics is the mathematics of collecting, organizing and interpreting numerical data, particularly when these data concern the analysis of population characteristics by inference from sampling" [22]. Statistics have solid and widely accepted mathematical foundations and can provide insights on the mechanisms creating the data. However, they frequently fail when dealing with complex and highly nonlinear data (curse of dimensionality) [21]. Bus arrival time is also impressed by several factors including driver behavior, carriage way width, intersections, signals and etc. and those factors are used as independent variables in many studies. The precision in these methods depends to all the dependent variables that they are recognized and incorporated in the model, which is a tough procedure [16]. The most literatures about time series model and regression model have been published before 1990s however, recently, new studies worked on combination of these models with others that we discuss them in hybrid models.

2.2.1. Time series model

Time series models depend on the data which is from historical time periods and forecast the future time periods. In this model, it is assumed that a pattern or mixture of patterns happens occasionally over time and these patterns can be provided by mathematical functions and for this purpose historical data can be used. Time series models assume that the historical traffic patterns will remain the same in the future. In the time series models, its precision highly depends on a function of the correspondence between the real-time and historical traffic patterns [23]. Variation in historical data or changes in the relationship between historical data and real-time data could significantly cause inaccuracy in the prediction results [24], and the problem in these methods usually back to its short time delay if the prediction model is in the real time [25], [26]. D’Angelo used a non-linear time series model to predict a corridor travel time on a highway [27]. He compared two cases: the first model used only speed data as a variable, while the second model used speed, occupancy, and volume data to predict travel time. It was found out that the single variable model using speed was better than the multivariable prediction model.

2.2.2. Regression models

This approach is applying mathematical models to predict the expected travel times between stops and then the expected bus arrival times at individual stops. Regression models use multiple independent variables (regressors) to clarify a single dependent variable using a linear or non-linear relationship. This model is usually provided by regressing travel times against a set of independent variables, such as traffic circumstances, passenger, number of bus stops, and climate situation. Regression models are conventional approaches for predicting travel time [28]. Regression models predict a dependent variable with a mathematical function formed by a set of independent variables. To make a regression model, the dependent variables require being independent. This obligation confines the applicability of the regression model to the transportation systems because variables in transportation systems are highly inter-correlated [23]. Patnaik in [29] developed a set of regression models to estimate bus arrival times with collected...
data by automatic passenger counters (APC) embedded on buses. The attained results demonstrated that the proposed models could be employed to predict bus arrival times based on different situations. However, this method is dependable only in case such equations could be provided, that it might not be probable for any application environments where in the most system variables, they are characteristically correlated. Kwon et al. [30] proposed the linear regression method and tree-based methods to estimate travel time on freeways based on the flow and occupancy data (measured by loop detectors).

2.3. Kalman Filtering Model

Kalman filtering models [31,6,32] have been used extensively in travel time estimation research. As Kalman [33] argues the basic function of Kalman filtering model is to provide prediction of the present status in the system, nonetheless it correspondingly works as the basis for estimation future values or for mending prediction of variables in former times. [17] used Kalman filtering techniques to forecast auto travel time. The Kalman filtering model has the potential to adapt to traffic fluctuation with time-dependent parameters [34]. These models are effective in predicting travel time one or two time periods ahead, but they deteriorate with multiple time steps [32]. Park and Rilett compared artificial neural network (ANN) models with other prediction models including Kalman filtering techniques to predict freeway link travel time. In contrast with, the average mean absolute percentage error (MAPE) of ANN altered from 8.7 to 16.1 for 1 and 5 time periods respectively , while in Kalman filtering, it altered from 5.7 to 20.1 [32]. Bae and Kachroo [35] prescribed Kalman filter model to estimate arterial travel time for buses with equipped buses to AVL as probe vehicles. Cathey and Dailey proposed a method which had three components namely tracker, filter and predictor [36].The Kalman filter applied in the filter section. In [6], they predict bus arrival time on Indian traffic circumstances by using GPS data installed on buses. Their method was using a model-based algorithm and the estimation was relying on Kalman filtering method. Shalaby and Farhan collected data by AVL and APC for the prediction of bus arrival time [37]. They employed Kalman filter for estimation of bus running time and passenger dwell time in an integrated framework and compared the results with a historical model, a regression model. Nanthawichit et al. [38] applied a macroscopic traffic flow model along with Kalman filtering algorithm to forecast travel time with combination of detector data and probe vehicle data. There are further studies that can be found in combination Kalman filtering with other models and we discuss later on hybrid models.

2.4. Machine Learning Models

Machine learning (ML) is studying how the computer to simulate or to realize the study behavior of human being. Figure 2 illustrates the basic structure of ML. The environment provides certain information to the learning section in system, and the learning section reviews knowledge library with consideration such this information. ML methods contain of two stages, i.e., choosing a candidate model, and next, prediction the parameters of the model through learning process on existing data [39,40]. ML methods have certain benefits with respect to statistical methods: dealing with complex relationships between predictors that can come up within a huge volume of information; processing non-linear relationships between predictors; processing complicated and noisy data [41]. These models can be used for prediction of travel time, without implicitly addressing the traffic processes. Results obtained for one location are normally not transferable to the next, because of location specific circumstances, e.g., geometry or traffic control. Artificial Neural Network (ANN) and Support Vector Machine methods are presented under these categories.

![Figure 2. Learning System Basic Structure](image)

2.4.1. Artificial neural network model

ANNs emulate the learning process of the human brain [32]. They are good at pattern recognition, prediction, classification, etc. ANNs have two stages, training and testing. During the training stage, inductive learning principles are used to learn patterns from a training set data. There are two types of learning processes used: unsupervised and supervised learning. In unsupervised learning, the network attempts to classify the training set data into different groups based on input patterns. In supervised learning, the desired output from output layer neurons is known, and the network adjusts the weight of connections between neurons to produce the desired output. During this process, the error in the output is propagated back from one layer to the previous layer by adjusting the weights of the
connections. This is called the back-propagation method, which is the most frequently used technique in transportation applications. The learning process of ANNs can be continuous so that the models can adapt to changes in environmental characteristics. In other words, ANN models can be considered dynamic prediction models because they can be updated and modified using new online data.

Due to their ability to solve complex non-linear relationships, artificial neural network models (ANNs) have been developed for transportation since the early 1990s [42]. ANN models had better results than those of existing link travel time techniques, including a Kalman filtering model, an exponential smoothing model, a historical profile, and a real-time profile [32]. In addition, ANN model showed better performance than historical average and autoregressive integrated moving average (ARIMA) models to predict short-term traffic flow. While other models are dependent on cyclical traffic data patterns or need independence between dependent and independent variables, ANNs do not require that variables are uncorrelated and/or that they have a cyclic pattern.

2.4.2. Support vector machines

Support vector machines (SVMs) are a set of related supervised learning methods used for classification and regression. While other machine learning techniques, such as ANN, have been extensively studied, the reported applications of SVM in the field of transportation engineering are very few. Since support vector machines have greater generalization ability and guarantee global minima for given training data, it is believed that support vector regression will perform well for time series analysis. In [12], the authors developed Support Vector Machine (SVM) as a new learning machine algorithm to predict the bus arrival time. They pointed out that unlike the traditional ANN, SVM is not amenable to the over-fitting problem, and it could be trained through a linear optimization process. However, they [12] indicated that when SVM is applied for solving large-size problems, a large amount of computation time will be involved. In addition, the methods for selecting input variables and identifying the parameters should be further researched.

2.5. Hybrid Models

Several researchers suggested hybrid frameworks that integrated two or more models for travel time prediction. Liu et al. proposed a hybrid model based on State Space Neural Networks (SSNN) and the Extended Kalman filter (EKF) as trainer [43]. The issue in SSNN is that the model requires large data set for offline training. Van Lint [44] mixed linear regression model and locally weighted linear regression model in a Bayesian framework to enhance forecast precision and reliability. Although their method may produce larger prediction errors weather each sub-model in the model layer is biased. Jeong and R.Rilett [45] proposed a travel time prediction model with consideration schedule adherence and dwell times. Also they compared a historical data based model, Regression Models, and (ANN) Models. As result, they found that ANN Models outperformed the historical data based model and the regression models in the case of estimation precision. Ramakrishna et al [46] developed a multiple linear regression model and an ANN model on heterogeneous Indian traffic circumstances with limited dataset for bus travel time prediction. Park and Lee [47] claims Bayesian model and neural network model can be good combination to estimate for urban arterial link travel times. Chen and Chien [48] compared link-based and path-based travel time prediction models using Kalman filtering algorithm with simulated data. Chu et al. [49] considered the model system noises and developed an adaptive Kalman filtering-based travel time prediction method that fuses both point detector data and probe vehicle data. Kuchipudi and Chien [50] proposed a hybrid model with combination of path-based and link-based models on real data and under different traffic conditions. In [51], [52], reported a short term transit vehicle arrival times prediction algorithm by combination of real-time AVL with historical data source in Seattle, Washington. They used a Kalman filter model to track a vehicle location and statistical estimation for prediction of bus arrival time purpose.
3. CONCLUSIONS

In the previous section, an overall review highlighted some efforts which have been made up to now for prediction bus arrival travel. Besides, there were other may many more studies that we could not explain all them here. However, it is worth mentioning them in a glance that we prepared in the following. Predicting transit arrival/travel times has been the focus for many existing studies. Table 1 summarizes existing studies with respect to author, year of published,

<table>
<thead>
<tr>
<th>Author</th>
<th>Year</th>
<th>Model / Architecture</th>
<th>Data Source Considered</th>
<th>Feature(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zhu et al. [7]</td>
<td>2011</td>
<td>Historical data</td>
<td>GPS and FCD</td>
<td>Traffic flow, signal delay and dwell time</td>
</tr>
<tr>
<td>Chien et al. [17]</td>
<td>2003</td>
<td>Hybrid</td>
<td>RST</td>
<td>Real time data, and previous time interval data, Kalman filter and Historic data used</td>
</tr>
<tr>
<td>Lin [19]</td>
<td>1999</td>
<td>Historical data</td>
<td>GPS</td>
<td>bus schedule, bus delay patterns</td>
</tr>
<tr>
<td>Kidwell [20]</td>
<td>2001</td>
<td>Historical data</td>
<td>GPS</td>
<td>dividing each route to zones</td>
</tr>
<tr>
<td>Chien [23]</td>
<td>2002</td>
<td>Machine learning</td>
<td>CORSIM</td>
<td>ANN, link-based and stop-based data</td>
</tr>
<tr>
<td>Sun [28]</td>
<td>2003</td>
<td>Statistical</td>
<td>AVI</td>
<td>Local linear regression, Result compared with K-NN</td>
</tr>
<tr>
<td>Patnaik [29]</td>
<td>2004</td>
<td>Statistical</td>
<td>APC</td>
<td>Regression, number of stops, dwell times, number of passengers, &amp; weather</td>
</tr>
<tr>
<td>Park et al. [32]</td>
<td>1999</td>
<td>Machine learning</td>
<td>AVI</td>
<td>Correlation analysis between neighboring links, feed forward ANN, one hidden layer</td>
</tr>
<tr>
<td>Vanajakshi [53]</td>
<td>2007</td>
<td>Hybrid</td>
<td>AVI</td>
<td>Comparison with ANN, SVR1 used with radial basis kernel function (RBF)</td>
</tr>
<tr>
<td>Lee et al. [9]</td>
<td>2009</td>
<td>Hybrid</td>
<td>LBS</td>
<td>real-time and historical data, data mining technique used, knowledge base</td>
</tr>
<tr>
<td>Shalaby et al [37]</td>
<td>2003</td>
<td>Kalman filter</td>
<td>AVL and APC</td>
<td>dwell times, route divides into multi-link</td>
</tr>
<tr>
<td>Sun et al. [13]</td>
<td>2007</td>
<td>Hybrid</td>
<td>GPS</td>
<td>Route Linearization, finite state machine</td>
</tr>
<tr>
<td>Wu et al. [10]</td>
<td>2004</td>
<td>Hybrid</td>
<td>Loop detector</td>
<td>Using regression combination (SVR), RBF as kernel function</td>
</tr>
<tr>
<td>Bin et al. [12]</td>
<td>2007</td>
<td>Machine learning</td>
<td>GPS (not clear)</td>
<td>SVM, Estimation of traffic congestion, RBF</td>
</tr>
<tr>
<td>Yu et al. [54]</td>
<td>2011</td>
<td>Machine learning</td>
<td>AVI</td>
<td>SVM, multiple routes, comparison with ANN, k-NN and linear regression (LR)</td>
</tr>
<tr>
<td>Yu et al. [5]</td>
<td>2009</td>
<td>Hybrid</td>
<td>GPS</td>
<td>Linear regression, Historical data</td>
</tr>
<tr>
<td>Fei et al. [55]</td>
<td>2011</td>
<td>Machine learning</td>
<td>Loop detectors</td>
<td>Bayesian, Historical data for traffic in conjunction</td>
</tr>
<tr>
<td>Padmanaban et al. [16]</td>
<td>2009</td>
<td>Hybrid</td>
<td>GPS</td>
<td>dwell time, heterogeneous traffic conditions, Kalman filtering, Historic data</td>
</tr>
<tr>
<td>Zhang et al. [56]</td>
<td>2009</td>
<td>Hybrid</td>
<td>GPS (not clear)</td>
<td>Real time data, Historic data</td>
</tr>
</tbody>
</table>
prediction model, data source, and considered features (consist of important details in article).

REFERENCES:


[37] Amer Shalaby and Ali Farhan. "Bus Travel


[58] Peng Chen, Zhao Lu, and Junzhong Gu, "Vehicle Travel Time Prediction Algorithm Based on Historical Data and Shared Location," in *Fifth International Joint Conference on INC, IMS and IDC*, Shanghai, China, 2009, pp. 1632-1637.
