

PREDICTION OF BUS ARRIVAL TIMES AT BUS STOP

Ruslawati Abdul Wahab^{1, 2, 3*}, Muhamad Nazri Borhan^{1, 2}, Riza Atiq Abdullah O.K. Rahmat^{1, 2}

¹*Department of Civil and Structural Engineering, Faculty of Engineering & Built Environment,
Universiti Kebangsaan Malaysia (UKM), 43600 Bangi, Selangor, Malaysia*

²*Sustainable Urban Transport Research Centre (SUTRA), Universiti Kebangsaan Malaysia (UKM),
43600 Bangi, Selangor, Malaysia*

³*Department of Polytechnic Education, Ministry of Higher Education, Persiaran Perdana, Presint 4,
62100 Putrajaya, Malaysia*

(Received: May 2016 / Revised: January 2017 / Accepted: January 2017)

ABSTRACT

Bus arrival time is very important to passengers, not only at the origin terminal but also at every stop. If there is no predicted arrival time at a stop, the headway designed should match the bus frequency at the stop. The uncertainty in bus arrival time can hinder the headway from matching the bus frequency at the stops. Moreover, lack of information on the service route and actual arrival times at stops leads to difficulty for passengers in planning their trips. Observation surveys were conducted to collect data on the problems of bus arrival frequency and uncertain arrival times at a selected stop with multiple routes during off-peak hours in Putrajaya Malaysia. This paper proposes a method to estimate arrival times at bus stops using the adaptive neuro fuzzy inference system (ANFIS) and several models are proposed to predict arrival times using MATLAB Curve Fitting Tool. All the proposed models exhibited RMSE close to 0 and R^2 close to 1.

Keywords: Arrival prediction; Arrival time; Delay; Headway; Waiting time

1. INTRODUCTION

Headway refers to the time interval in bus operations. Bus frequency is the number of buses travelling through a service route. Bus stops are the medium or structures that are very familiar and close to the bus passengers. The relationships among these parameters lead to the bus arrival times at the different stops on a route. Problems arise if the arrival times of the buses at stops cannot be predicted by passengers owing to lack of provided information. Furthermore, there are also issues resulting from inconsistent bus-arrival times due to several factors, such as delayed bus departures, delayed pick-ups and drop-offs of passengers at bus stops, and delays in traffic and traffic lights. These problems have been observed at bus stops in the city of Jaipur, India, (Wagale et al., 2013) and at system stops in Hong Kong (Yu et al., 2011).

Inaccurate bus arrival times at stops significantly affect the headway designed by the bus management. Previous studies (Berrebi et al., 2015; Wagale et al., 2013; Yu et al., 2011) showed that the bus companies are usually more focused on minimizing bus operations while the users are more interested in shorter wait times, accuracy of arrival times, and frequency of bus arrivals. Bus managements design bus operations with convenient headway and release the bus schedule based on bus operation manual standards. However, the implementation is

*Corresponding author's email: ruslawati@siswa.ukm.my, Tel. +603-8921-5555, Fax. +603-8921-4097
Permalink/DOI: <https://doi.org/10.14716/ijtech.v8i1.4026>

dependent on the bus speed, bus operator, and the state of traffic on the highway (Yu et al., 2011). Factors such as punctuality of bus departures, delays on the road, and delays at stops affect the accuracy of arrival times at stops. Previous research on the prediction accuracy on arrival time used a model of support vector machine (SVM), artificial neural network (ANN), k nearest neighbors' algorithm (k-NN), and linear regression (LR) (Yu et al., 2011). The bus arrival time is directly related to the passenger waiting time at the bus stop. To minimize this waiting time, Berrebi et al. (2015) applied stochastic formulation to optimize the headway with high frequency. Wagale et al. (2013) applied the demand and travel time responsive (DTR) model to optimize social bus boarding using arrival times and departure times at stops. They found that the model was unsuccessful in improving the bus schedule system (headway) because the arrival times at stops are strongly influenced by traffic jams.

To ensure the comfortability of passenger, an information system of bus routes must be provided from various sources; printed, telephone and website (Rahmat, 2015). The availability of real-time bus information at stops will help passengers plan their bus trips efficiently and reduce long queues and waiting times at stops (Yu et al., 2011). The absence of information on arrival times at stops or inaccurate information may result in uncertainty regarding the frequency of buses at stops. For example, in the case of 30-min bus frequency, if the first bus arrives too early and the next bus arrives quite late, the service frequency will not comply with the information of one bus every 30 minutes. The passengers will have a negative impression regarding the service because they had assumed that the bus would arrive within 30 min. The reliability of the bus schedule is a very important for bus passengers (Soh et al., 2014).

This paper focuses on the basic development of optimum time arrival prediction models for multiple routes at a selected stop using the adaptive neuro fuzzy inference system (ANFIS). The variables involved are departure time of the bus from the origin terminal, headway schedule, travel times, and actual arrival times at the stop. Then, the variables are analyzed of their validity and the model's polynomial, exponential, and Gauss functions are created.

2. METHODOLOGY

There are three steps in the study: data collection on the study site, model development, and evaluation of the model functions. Data collection is carried out using the time-series method to collect empirical data. The time series can provide an understanding of the underlying forces and structure that produced the observed data (Bin et al., 2006). The parameters involved are the perceived off-peak time bus routes with best design planning such as headway, frequency, and bus arrival times at the stop, route distance, time services, speed, stops and comfortability of passengers (Rahmat, 2015; Kikuchi, 1985).



Figure 1 Map of the bus route, and location of the ALMD stop in Putrajaya

The study was conducted in Putrajaya, Malaysia, at the ALMD stop, as shown in Figure 1. The multiple routes involved in this study are L01, L02, L05, L07, and L10. The bus headway started at 6:30 am until 6:30 pm, except for the L05 that began at 6:45 am until 6:45 pm. This study was conducted over four days from 28 to 31 July 2015. The bus arrival times and number of passengers were observed at 30-min intervals.

2.1. Development of Adaptive Neuro Fuzzy Inference System, ANFIS

ANFIS was used to develop the forecast arrival time model at the stop. The grid partition structure model was developed using Sugeno, which produced only one output. The least-squares method was applied in the modelling process of the optimized hybrid composite. The epoch's value was used to achieve the desired criteria of the model. The framework of the forecast model was developed in this study by referring to the model proposed by Yu et al. (2011), but this study slightly differed in the data collection method of bus departure as the headway was considered according to Equation 1 given below:

$$\hat{T}_{l,n}^s = T_{l,n}^a + \hat{t}_{l,n}^{\text{running}} \quad (1)$$

where $\hat{T}_{l,n}^s$ is the predict time arrival, $T_{l,n}^a$ is the time departure following headway from origin terminal, and $\hat{t}_{l,n}^{\text{running}}$ is the predict time travel from origin terminal to stop

In the calculation of travel time to the ALMD stop, traffic congestion is assumed as per the past studies (Yu et al., 2010), and the data collection of the arrival is classified as routine data, which are valid and significant to the bus arrival times at the stops.

3. CASE STUDY

The proposed models for predicting the bus arrival time at a bus stop with multiple routes was evaluated using real-world data from Putrajaya, Malaysia. There are more than 100 bus stops in Putrajaya with variable designs, as shown in Figure 2a. This study was carried out using data for the ALMD stop. This is a popular transit stop for passengers because of its location in the main commercial center, and it has medium to high bus frequency, as shown in Figure 2b. During the study, the ALMD stop did not have information about bus arrival times compared to the other 17 stops, which have real-time electronic displays.

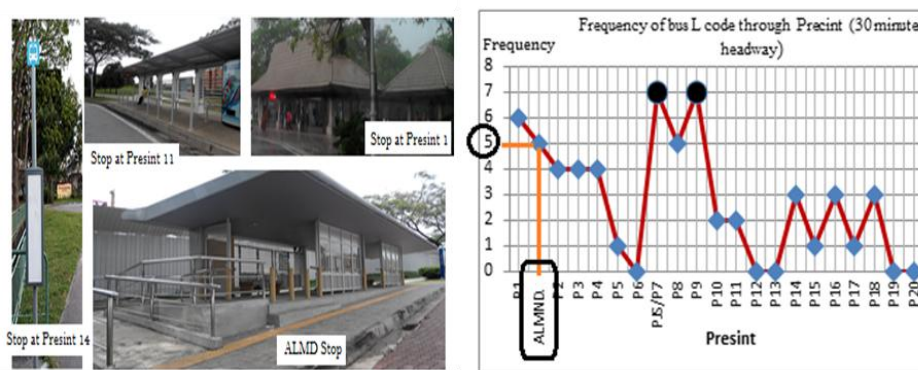


Figure 2 (a) Various design of stops; (b) frequency of service at different stops in Putrajaya

There are multiple bus routes passing through this stop. Passengers with no prior experience of travelling in such a bus network will be confused with the routes and bus arrival times at the

stop. They may require querying the bus operator repeatedly to get information regarding the targeted bus destinations. The study results are discussed in terms of the relationship of the headway parameter with the bus arrival times, passenger waiting time, and bus frequency.

4. RESULTS AND DISCUSSION

4.1. Headway vs. Arrival Times at Stops

The interval time or headway of 30 min was designed for the L route during off-peak hours. The headway of the route started at 6.30 am according to the allocation of routes and schedule of the bus operator. All the codes are expressed by the stated period, except for the L05, for which the headway started at 6.45 am.

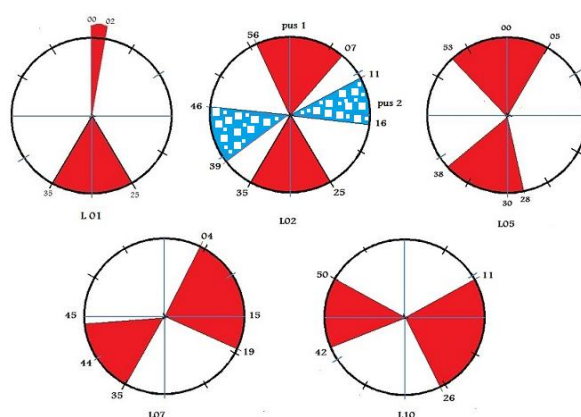


Figure 3 Average minute arrival time at the ALMD stop

There are four routes of the L code that pass by the ALMD stop at various times. Figure 3 shows the result of actual average arrival times in minutes at the ALMD stop from the observations. The minute clock is drawn to simplify visualization of the real arrival time data collection of the bus routes. The minutes in the clock highlighted in red and dotted blue are the ranges of arrival times for multiple bus routes, namely, L01, L02, L05, L07 and L10. In this minute clock, the first 30 minutes is divided from 1–6 on the clock as the first headway and the next round as the second headway from 6–12 on the clock. The study found that the bus arrival times for L01 for the first headway (first 30 min) is between minutes 00:00 and 00:02 and the second bus (next 30 min) arrived between minutes 00:25 and 00:35. The first bus for L01 was consistent because the range of the bus only spanned two minutes but for the next bus the range was larger, averaging at 10 min. Passengers without this information cannot predict the actual arrival time of the buses, unless they are regular travelers by the same bus or collect information regarding the bus routine time beforehand. The routes L05 and L02 have an average arrival between minutes 00:55 and 00:05, and the next bus between minutes 00:25 to 00:35. The average arrival of buses on other routes L02, L07, and L10 are between minutes 00:04 and 00:26, with the next bus arriving between minutes 00:35 and 00:50.

PjS and PR14 are two terminals that provide park and ride and operate for bus service. For route L02, the highlighted dotted blue color is shown on the clock to depict the overlapping buses from both origins that depart simultaneously and pass through the same stop, as shown in Figure 4. The studied ALMD stop is located between these two origin terminals. Apart from the L02 route, the L05 route also has both origins, from where the number of buses was distributed equally for each terminal. The buses thus intersect on the road with different destinations. The headway and boarding times are the same for the buses from different origins. To avoid problems in counting data, the researchers marked the same L02 code from different

destinations as L021 from the origin terminal P&R14 to PjS and L022 in the opposite direction. In this case, passengers without such information may choose their destination based on the same code, but end up travelling in the wrong direction.

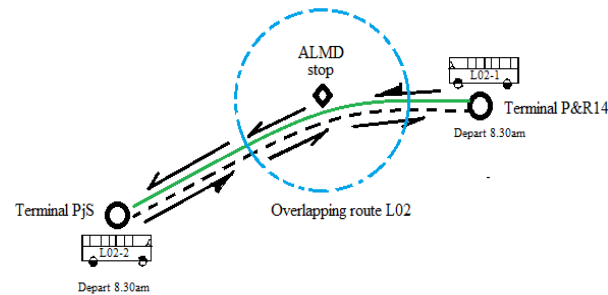


Figure 4 Average arrival time at the ALMD

4.2. Waiting Time and Frequency at ALMD Stop

From the observation data, the researchers also found that the uncertain arrival times of each route do not meet the headway or the frequency of one bus every 30 min of period service. This is because the arrival times of the buses depend on serious traffic and delays during passengers alighting and boarding the buses. The headway of 30 min will thus result in a long waiting time if the passengers miss the bus schedule.

The waiting time is related to the frequency per round of an exhaustive 60 min. The bus headway of 30 min means there are two frequencies for each line in one trip per hour. The bus trips have the first frequency bus every 30 min, and another bus, which is referred to as the second frequency, arrives 30 min later.

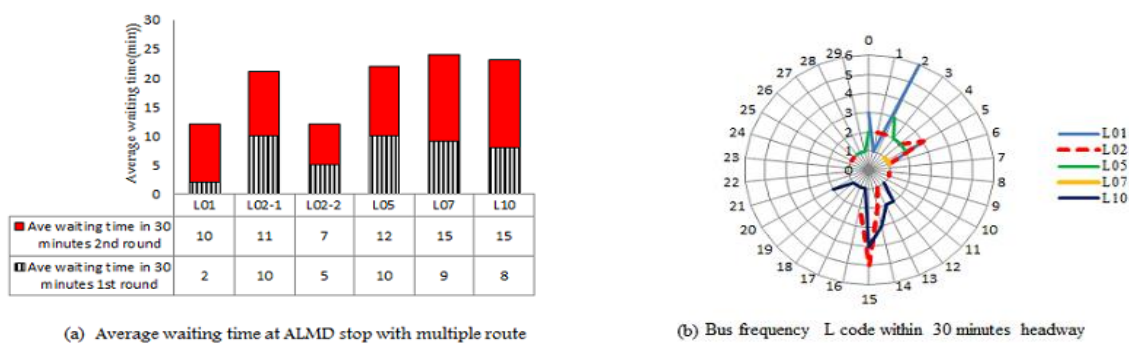


Figure 5 (a) Average waiting time; (b) Frequency at the ALMD stop

Figure 5a shows the average waiting time at the ALMD stop for multiple routes, and Figure 5b shows the frequency of the bus. The minimum waiting time dedicated to the L01 bus in the first round is an average of 2 min and the second bus round is an average of 10 min. The longest waiting time is an average of 15 min for the second rounds of L07 and L10.

Study has showed that inconsistent bus arrival times can result in anxiety for passengers waiting at bus stops (Yu et al., 2011). This is because a bus either does not fulfil the punctuality or delay the boarding and lead to the uncertainty arrival. It affects the bus frequencies, such that there will be two buses in the first round of 30 min and no buses in the following 30-min interval. An example of this situation can be seen for the L01 route, where the first bus reaches the stop at 8:02 am and the next bus arrives at 8:25 am. Information on arrival times should be provided by the bus operator to minimize the passenger waiting time at the bus stop and to avoid the queuing due to simultaneous arrivals. Improvement of the arrival schedule or the

employment of a holding strategy is recommend to avoid the buses from arriving too early or too late in the 30-min interval.

4.3. Arrival Time Prediction Model using ANFIS

Previous research (Yu et al., 2011) suggested that radial basis function (RBF) kernel be used in the SVM model. ANFIS was therefore used as a model in this study to evaluate the performance. A predictive model was developed with actual arrival times from the survey data at the ALMD stop. This model had two input parameters and one output parameter.

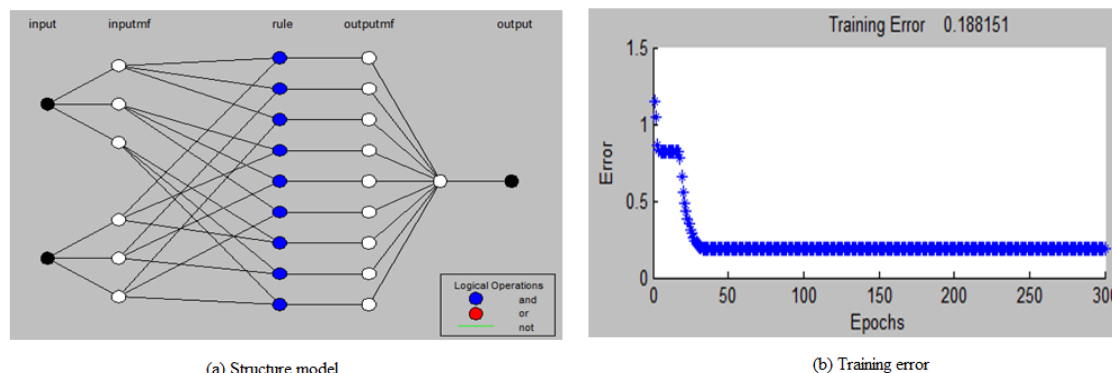


Figure 6(a) Structure model and (b) Training error of predicted arrival times using ANFIS

To determine the inputs of the ANFIS model, sensitivity tests were conducted, and the results are shown in Figure 6. Figure 6a shows the division of the structure and the predictive input model using the basic structure of the Fuzzy–Mandani system to map the characteristics of the model input to the membership function input, the next input membership functions to the rules, rules to the characteristics, and the output set to the value of the output or results. This method used a fuzzy model to understand the data set, and its function was calculated using modified membership function parameters. The ANFIS model has separate test data and training data. Analysis of the data input and output must meet the requirements of the model used to insert the test data in advance so that researchers get the desired output. The training data can then be carried out, and the resulting decisions are displayed in Figure 6b. The display shows that the validity of the predictive models can be accepted because the error was 0.1881. Therefore, ANFIS could be very useful in the predictions when it may be difficult or even impossible to mathematically formulate the relationship between the input and output.

4.4. Comparison of Function Model between Forecast and Actual Arrival Time

Several functions can be generated from the actual and forecast arrival times. All three models were generated using the non-linear least square method in MATLAB Curve Fitting Tool with 95% constant border. The trust region algorithm was selected to build the optimum function corresponding to the validation data and display goodness of fit statistics. The result of algorithm models will evaluate by the validity of R^2 and RMSE respectively. These variables function are fitting in three different models; Gaussian 1, Polynomial, and Exponential forms.

MATLAB Curve Fitting Tool does not require a specific form of function. This tool can create, plot, and compare multiple fits. In this research, we loaded the output value from ANFIS, and then selected the fitting curve, which automatically generated the option of the different model and changed the degree to fit the model.

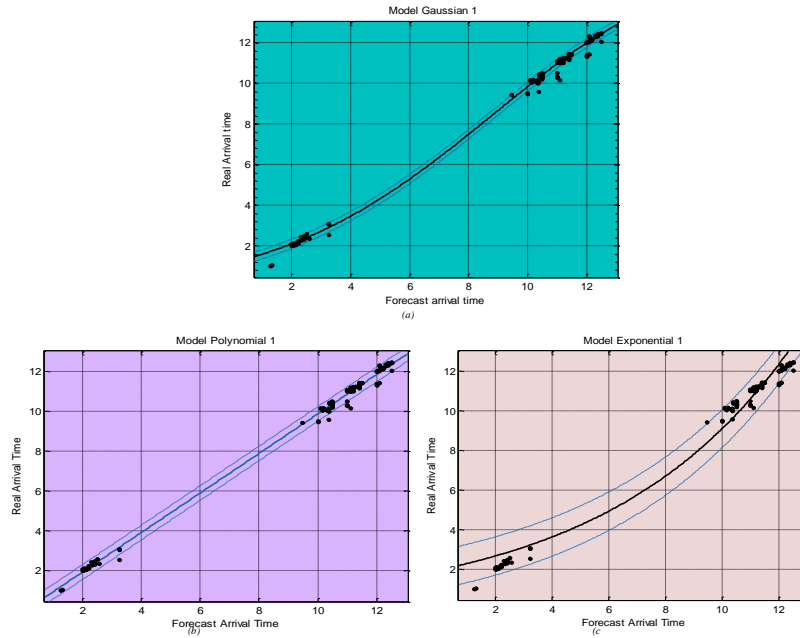


Figure 7 Performance of the three models on real-time and predicted arrival times

Table 1 Model function and correlation coefficients and R^2 and RMSE values

Model (degree 1)	Function	R-square	RMSE
Gauss	$f(x) = a_1 \times \exp(-((x-b_1)/c_1)^2)$ $a_1 = 14.2 \text{ (13.82, 14.57)}$ $b_1 = 1.884 \text{ (1.776, 1.992)}$ $c_1 = 2.75 \text{ (2.681, 2.819)}$	0.9987	0.1359
Polynomial	$f(x) = p_1 \times x + p_2$ $p_1 = 0.9989 \text{ (0.997, 1.001)}$ $p_2 = -0.03743 \text{ (-0.05613, -0.01873)}$	0.9999	0.1158
Exponential	$f(x) = a \times \exp(b \times x)$ $a = 1.973 \text{ (1.807, 2.139)}$ $b = 0.1529 \text{ (0.1455, 0.1604)}$	0.9763	0.5766

Figure 7 and Table 1 show the results of the bus running times predicted by the three models against the observed running times. It can be seen from the figure that the results of the ANFIS forecast model are much closer to the observed data. The correlation coefficients and the R^2 and RMSE values, which reflect the accuracy of the bus running time prediction of the three models, are 0.9987, 0.9999, and 0.9763, respectively. This result implies that the predicted bus running times from each model are well fitted with the observed bus running times. Further, the RMSE values are close to zero. In summary, based on the validation results, the performances of all the ANFIS models for bus arrival time prediction at the stop with multiple routes are found to be satisfactory.

5. CONCLUSION

This study investigated the problems associated with delayed bus arrival times and proposed a model to forecast arrival times at a stop with multiple routes. The aim of the study was to improve bus operation and maintain good services to the passengers by decreasing their waiting

time. Travel time predictions should consider poor schedule or headway adherence. Considering the inconsistent timings of bus departure, this research used the arrival time of the current bus to estimate the travel time and developed prediction models that used the ANFIS output. The MATLAB Curve Fitting Tool was then used to create the model because it can generate input/output relationships without requiring an explicit function form and it exhibits good general performance. The model was constructed using data of real-time bus arrival, departure time from the terminal, headway, and travelling time to the stop. Several factors such as the traffic on the road, the time delay at every stop, and traffic lights were included in the travel time predictions. Time arrival data at the stop was based on the daily routine of the bus routes. In the future, this model can be improved by comparison with other prediction models. Further, input data can be improved to obtain more appropriate and varied forecast arrival times.

6. ACKNOWLEDGEMENT

The authors wish to thank the Department of Civil and Structural Engineering, Faculty of Engineering and Built Environment, Universiti Kebangsaan Malaysia (UKM) and Pengangkutan Awam Putrajaya Sdn. Bhd. for their support in this research project. The project was funded by UKM under the Project GUP-019-2016.

7. REFERENCES

- Berrebi, S.J., Watkins, K.E., Laval, J.A., 2015. A Real-time Bus Dispatching Policy to Minimize Passenger Wait on a High Frequency Route. *Transportation Research Part B: Methodological*, Volume 81, pp. 377–389
- Bin, Y., Zhongzhen, Y., Baozhen, Y., 2006. Bus Arrival Time Prediction using Support Vector Machines. *Journal of Intelligent Transportation Systems*, Volume 10(4), pp. 151–158
- Kikuchi, S., 1985. Relationship between the Number of Stops and Headway for a Fixed-route Transit System. *Transportation Research Part A: General*, Volume 19(1), pp. 65–71
- Rahmat, O.K., R.A.A., 2015. *Urban Transport Planning*. Universiti Kebangsaan Malaysia, UKM Press, Bangi, Malaysia, pp. 81–96
- Soh, K.L., Chong, C.L., Wong, W.P., Hiew, Y.H., 2014. Proclivity of University Students to Use Public Bus Transport Service. *Comprehensive Research Journal of Education and General Studies (CRJEGS)*, Volume 2(2), pp. 24–34
- Wagale, M., Singh, A.P., Sarkar, A.K., Arkatkar, S., 2013. Real-time Optimal Bus Scheduling for a City using a DTR Model. In: Proceedings of the 2nd Conference of the Transportation Research Group of India (2nd CTRG), Uttar Pradesh, 12th to 15th December 2013, India
- Yu, B., Lam, W.H.K., Tam, M.L., 2011. Bus Arrival Time Prediction at Bus Stop with Multiple Routes. *Transportation Research Part C: Emerging Technologies*, Volume 19(6), pp. 1157–1170
- Yu, B., Yu, B., Lu, J., Yang, Z., 2010. An Adaptive Bus Arrival Time Prediction Model. *Journal of the Eastern Asia Society for Transportation Studies*, Volume 8, pp. 1126–1136