Research Article

An ELM-Based Approach for Estimating Train Dwell Time in Urban Rail Traffic

Wen-jun Chu, Xing-chen Zhang, Jun-hua Chen, and Bin Xu

State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing 100044, China

Correspondence should be addressed to Wen-jun Chu; 12114228@bjtu.edu.cn

Received 22 July 2014; Accepted 9 October 2014

Academic Editor: Tao Chen

Copyright © 2015 Wen-jun Chu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Dwell time estimation plays an important role in the operation of urban rail system. On this specific problem, a range of models based on either polynomial regression or microsimulation have been proposed. However, the generalization performance of polynomial regression models is limited and the accuracy of existing microsimulation models is unstable. In this paper, a new dwell time estimation model based on extreme learning machine (ELM) is proposed. The underlying factors that may affect urban rail dwell time are analyzed first. Then, the relationships among different factors are extracted and modeled by ELM neural networks, on basis of which an overall estimation model is proposed. At last, a set of observed data from Beijing subway is used to illustrate the proposed method and verify its overall performance.

1. Introduction

Dwell time is the time that a public transport vehicle spends at a station or a stop for passenger alighting and boarding [1]. In any mode of public transportation, it is an important parameter, which determines the system performance and service quality to a large extent. On one hand, dwell time constitutes a significant part of the total trip time, which is the key criterion for service quality of public transit. On the other hand, dwell time determines the capacity utilization of infrastructure, thus affecting the efficiency of the whole transit system. Therefore, reasonable estimation of dwell time plays an important role in operation of various public transit systems.

A number of studies have been conducted on dwell time estimation in various types of public transportation and corresponding research approaches can be roughly classified into two categories: regression approach and microsimulation approach.

Regression approach is to establish regression model with observed data to describe the relationship between dwell time and corresponding factors. This approach is first used in the estimation of bus dwell time. Levinson [2] proposed a linear regression model to estimate bus dwell time, in which the bus dwell time is formulated as a linear function of two primary contribution factors—number of alighting and boarding passengers and the amount of time required for bus doors opening and closing. Since then, a number of studies were carried out to take into account some other contributing factors for the bus dwell time estimation. For example, Guenthner and Hamat [3] investigated the relationship between the bus dwell time and bus fare collection system. Levine and Torig [4] analyzed impact of bus floor types on the bus dwell time. Jaiswal et al. [5] examined influence of platform walking on bus rapid transit stations on bus dwell time. Tirachini [6] studied impact of fare payment technology in urban bus services. Most previous studies on urban rail dwell time estimation also applied the regression approach. Weston [7] proposed a polynomial regression model using the survey data of London Metro, in which various contributing factors, including the number of alighting and boarding passengers, passenger distribution, and on-board crowdedness, are considered. Lam et al. [8] proposed a linear regression model on basis of observed data from two LRT stations. Lin and Wilson [9] compared linear and nonlinear regression models with observed data of MBTA Green Line and proved that crowdedness has a nonlinear effect on urban rail dwell time. On this basis, Puong [10] proposed a nonlinear dwell time model that can fit 90% of observed data from MBTA Red Line.
As can be seen, almost all proposed regression models on dwell time estimation are polynomial. In these studies, the model structure is first determined through certain hypothesis and then corresponding parameters are calibrated. Under this condition, though these models fit respective field data well, the generalization performance of them cannot be ensured.

Microsimulation approach is to calculate the required dwell time on basis of single passenger behavior description under computer environment. In recent years, computer-based pedestrian simulation technology rapidly develops and is gradually introduced into dwell time estimation. Li et al. [11] applied Monte Carlo simulation to simulate the bus dwell process, in which a binary door choice model predicting the proportion of alighting passengers through front or rear door is integrated. Zhang et al. [12] proposed a cellular automaton based alighting and boarding microsimulation model for passengers in Beijing subway stations, which is proven effective in estimating urban rail dwell time. Bae et al. [13] investigated the influence of different boarding/alighting strategies on urban rail dwell time on basis of a microsimulation model, in which an inclination function governing passengers’ movement in a two-dimensional queue is introduced. In addition, some commercial pedestrian simulation software programs, such as VISSIM and Legion, are applied to calculate dwell time in many related studies.

Theoretically speaking, microsimulation models have better generalization performance than regression model. If the behavior of passengers is described properly, the model can be used in any scenario. However, existing microscopic simulation theory is still insufficient in describing pedestrian behavior under crowded condition. As a result, the accuracy of microsimulation dwell time estimation models cannot be ensured at present.

In urban rail transit system, train operation is typically based on timetables which are made in advance and the dwell time at each station is assigned beforehand. Under this condition, the reasonability of preassigned dwell time may have a significant influence on the performance of the whole system. If the assigned dwell time is insufficient for passenger alighting and boarding, delay will happen and complicated adjustments need to be made in the predesigned timetable so as to ensure the following train operation. On the other side, if the assigned dwell time is too long, the headway between two consecutive trains will also be overlong, consequently limiting the capacity of the whole transit line. Therefore, in all urban rail transit systems, especially in those with heavy traffic such as Beijing subway, reasonable estimation of dwell time is essential to create effective timetables and make a compromise between service quality and transportation capacity.

Artificial neural network is a widely used method of data fitting. It can approximate complex nonlinear mappings directly from the input sample without making much hypothesis beforehand. In this paper, a new proposed artificial neural network method ELM is used in urban rail dwell time estimation. The outline of the paper is as follows. In Section 1, previous research regarding dwell time estimation of public transportation is reviewed. Section 2 elaborates the principles and steps of ELM. Section 3 makes a detailed analysis on the factors of train dwell time at urban rail stations and Section 4 presents the structure of the proposed model. In Section 5, several data sets on Beijing subway are used to evaluate the proposed model. Conclusions and discussions are given in Section 6.

2. Extreme Learning Machine

Single-hidden layer feedforward network (SLFN) is a widely used type of artificial neural network, which has been proven effective in complex nonlinear approximation [14–16]. Figure 1 illustrates the structure of a standard SLFN. In this network, \( n \) input nodes and \( m \) output nodes are included, corresponding to \( n \)-dimensional input vector and \( m \)-dimensional output vector. \( L \) nodes are contained in the hidden layer and \( \xi_i \) is the threshold of the \( i \)th hidden node.
As named in Huang and Babri [17], $H$ is called the hidden layer output matrix of the SLFN and the $i$th column of it corresponds to the output of $i$th hidden node and the output nodes.

Given $N$ arbitrary training samples $(x_j, e_j)$, where $x_j = [x_{j1}, x_{j2}, \ldots, x_{jm}]^T \in \mathbb{R}^n$ and $e_j = [e_{j1}, e_{j2}, \ldots, e_{jm}]^T \in \mathbb{R}^m$, the output of the above SLFN is

$$o_j = \sum_{i=1}^L k_i g(w_i \cdot x_j + \xi_i).$$

If this SLFN can approximate these $N$ samples with zero error, that is, $\sum_j^N \| o_j - e_j \| = 0$, then there exist $k_i$, $w_i$, and $\xi_i$ such that

$$\sum_{i=1}^L k_i g(w_i \cdot x_j + \xi_i) = e_j,$$  \hspace{1cm} (2)

These $N$ equations can be written compactly as

$$HK = E,$$  \hspace{1cm} (3)

where

$$H = \begin{bmatrix}
g(w_1 x_1 + \xi_1) & \ldots & g(w_L x_1 + \xi_L) \\
\vdots & \ddots & \vdots \\
g(w_1 x_N + \xi_1) & \ldots & g(w_L x_N + \xi_L)
\end{bmatrix}_{N \times L},$$

$$K = \begin{bmatrix}
k_1^T \\
\vdots \\
k_L^T
\end{bmatrix}_{L \times m},$$

$$E = \begin{bmatrix}
e_1^T \\
\vdots \\
e_N^T
\end{bmatrix}_{N \times m}.$$  \hspace{1cm} (6)

As named in Huang and Babri [17], $H$ is called the hidden layer output matrix of the SLFN and the $i$th column of it corresponds to the output of $i$th hidden node with respect to $N$ inputs. As proven by Huang et al. [18], given arbitrary $w_i$ and $\xi_i$, the least square solution of $K$ in formula (3) can be obtained by formula (7):

$$\hat{K} = H^T E,$$  \hspace{1cm} (7)

where $H^T$ is the Moore-Penrose generalized inverse of matrix $H$. On this basis, a simple and efficient training algorithm for SLFN called ELM is proposed [18], whose procedure can be summarized as follows.

**Step 1.** Randomly assign input weight $w_i$ and bias $\xi_i$, $i = 1, 2, \ldots, L$.

**Step 2.** Calculate the hidden layer output matrix $H$ according to formula (4).

**Step 3.** Calculate the output weight $K$ according to formula (7).

Due to the fast training speed, ELM has been widely used for many applications [19]. In this paper, ELM is applied to approximate the complex relationship between the factors of urban rail dwell time.

### 3. Factors of Urban Rail Dwell Time

Urban rail dwell time is typically defined as the time elapsed between the door opening and closing of a train sitting at a station [10]. In this period, several tasks need to be accomplished, as shown in Figure 2.

In Figure 2, the horizontal axis represents time and $t_0$ represents the time when the train stops and doors begin to open. On the vertical axis, four types of task are listed. The duration
of door opening and closing process is mainly determined by the mechanism of the vehicles. The confirmation process represents the interval between the end of passenger alighting at all doors and the beginning of door closing process, which is used for operators confirming the completion of passenger alighting. The start time of this process depends on the door at which passenger boarding completes last, that is, the door $i^\ast$. The times of alighting and boarding tasks vary across doors. According to previous research, this is mainly because the numbers of alighting, through, and boarding passengers differ from door to door. In other words, the duration of alighting and boarding process at a door is mainly decided by the number of passengers alighting and boarding from this door and the crowdedness of corresponding vehicle. And these parameters will be affected by the passenger flow and platform pattern of this station and previous stations.

Nevertheless, in practical terms, there exist overlaps between some consecutive tasks. As shown in Figure 2, the overlap between door opening and passenger alighting represents that some passengers begin to alight before the door is fully open and the overlap between passenger alighting and boarding represents that some passengers do not obey the “get off and then on” rule. Under this condition, times of these processes cannot be separately considered, no matter from the perspective of survey or estimation. Therefore, an overall concept, passenger service time, is proposed here, which represents the period from the beginning of door opening to the end of passenger boarding at single or all doors.

On basis of the above analysis, the factors of urban rail dwell time and their interaction can be concluded, which is shown in Figure 3.

### 4. Urban Rail Dwell Time Estimation

#### 4.1. Notations

The key notations used in the dwell time estimation are shown in Notation Definitions section.

#### 4.2. Problem Statement

Generally speaking, in practical operation of urban rail system, the operation-related parameters, that is, platform pattern, vehicle performance, and operation efficiency, are relatively stable. Therefore, only the influence of the traffic-related parameters which is the concern of most previous research is taken into account here. On this basis, the urban rail dwell time estimation problem can be described as follows.

Consider a $n$-door urban rail train that will make a stop on a station. On the train, $A$ passengers will alight at the station and $C$ passengers will not. On the platform of the station, $B_j$ passengers who enter the platform through entrance $j$ are waiting to get on this train. In addition, the train needs $\tau_1$ to close all its doors and operators need to spend $\tau_2$ to confirm the full close of all doors. Thus, assign a minimum dwell time $D$ for the train, which is sufficient for passengers alighting and boarding at the station.

According to the analysis in Section 3, the required dwell time $D$ can be seen as the accumulation of three parts: the maximum single-door passenger service time, duration of door closing process, and confirmation time; that is,

$$D = \max_i t_i + \tau_1 + \tau_2,$$

where the passenger service time at $i$th door $t_i$ is determined by the number of boarding, alighting, and through passengers at this door; that is,

$$t_i = F(a_i, b_i, c_i).$$

Furthermore, for a specific station, the distribution of boarding passengers on the platform is always accorded with certain rules [18], which means certain mapping exists between the vector $\beta = [b_1, b_2, \ldots, b_n]^T$ and the boarding passenger vector $B = [B_1, B_2, \ldots, B_m]^T$; that is,

$$\beta = g_1(B) = g_1(B_1, B_2, \ldots, B_m).$$
By contrast, the distribution of alighting and through passengers on board, which is determined by platform pattern of previous stations, is more complicated. In previous research, the alighting and through passengers on board are usually assumed to be uniformly distributed [10] or distributed with constant proportion [7]. In this paper, the uniform distribution is adopted for $\alpha$ and $\gamma$; that is, $\alpha = [a_1, a_2, ..., a_n]^T = g_2(A) = \left[\frac{A}{n}, \frac{A}{n}, ..., \frac{A}{n}\right]^T$, $\gamma = [c_1, c_2, ..., c_n]^T = g_2(C) = \left[\frac{C}{n}, \frac{C}{n}, ..., \frac{C}{n}\right]^T$.

To summarize, the required dwell time $D$ can be described as follows:

$$D = \max_i \left\{ F \left( \frac{A}{n} [g_1(B_1, B_2, ..., B_m)]_{i,1}, \frac{C}{n} \right) \right\} + \tau_1 + \tau_2. \quad (12)$$

As can be seen, the key to dwell time estimation is to approximate the mappings $F$ and $g_1$.

4.3. ELM-Based Estimation Model. In this section, two ELM neural networks are designed to approximate the mappings shown in formula (12). On this basis, an overall estimation model is proposed.

4.3.1. Single-Door Passenger Service Time (SDPST) Model. In order to approximate the relationship between $t_i$ and $(a_i, b_i, c_i)$, that is, $(a_i, b_i, c_i)$, an ELM neural network is designed, whose structure is shown in Figure 4(a). As illustrated in this figure, the model has an input vector of three dimensions which represent $a_i$, $b_i$, and $c_i$, respectively, and a single-dimensional output vector $t_i$. Sigmoid function is chosen as the activation function of the hidden nodes and the number of hidden nodes $p$ needs to be determined through $k$-fold cross-validation with training data set.

4.3.2. Platform Passenger Distribution (PPD) Model. Another ELM neural network is designed to describe the distribution rule of passengers on platform, as shown in Figure 4(b).

This model has an input vector of $m$ dimensions which represent the numbers of boarding passengers from each entrance and an output vector of $n$ dimensions which represent the number of boarding passengers at each door. Besides, the activation function of this model is also sigmoid function and the number of hidden nodes is $q$, which also needs to be determined through cross-validation.

4.3.3. Overall Dwell Time Estimation Model. On basis of the previous two models, an overall model for urban rail dwell time estimation is proposed, which is shown in Figure 5. In this model, the mappings $F$ and $g_1$ in formula (12) are replaced by SDPST model and PPD model, respectively, and this two ELM neural networks need to be trained separately with corresponding data sets.

5. Model Evaluation

5.1. Data Collection and Processing. A survey is conducted on the outbound platform of Zhichunlu station of Line 13, Beijing subway. This platform is a typical side platform with three stairways and one escalator acting as entrances and exits, as shown in Figure 6. In the survey, 24 recorders are assigned to observe the 24 doors of trains, respectively, and another two are assigned to record the number of boarding passengers entering from the two entrances. After 10 days’ survey, a raw data set containing 8304 instances from 346 trains is obtained, whose structure is illustrated in Table 1. It should be noted that the actual number of through passenger cannot be observed precisely from platform. Therefore, the attribute $c$, which is used to describe the crowdedness on the vehicle, is replaced by the number of through passengers that stand on board near the door.
5.1.1. Operation-Related Parameters. Firstly, the confirmation and door closing times are derived. Considering the effect of scheduled dwell time, only the records in which actual dwell time exceeds scheduled dwell time are used and the sum of constant parameters $\tau_1$ and $\tau_2$ is assigned with the average of differences between $T$ and PST; that is,

$$\tau_1 + \tau_2 = \sum_{r} \frac{(T - PST)}{\{r \mid T - T_0 > 60\}}.$$  \hfill (13)

5.1.2. SDPST Data Set. This data set has 8304 instances, each of which represents a passenger service process at a single door. Four attributes, $a$, $b$, $c$, and PST, are contained and corresponding data can be extracted directly from the raw data set. This data set can be used to train the SDPST model.

5.1.3. PPD Data Set. 346 instances are contained in this data set, each of which corresponds to an observed train. There are 26 attributes per instance. Two of them are the numbers of boarding passengers entering from the two entrances (named as $B_1$ and $B_2$) and the rest represent the number of boarding passengers at each door (named as $b_i$, $i = 1, 2, \ldots, 22$). In this way, the distribution of boarding passengers for each observed train can be described by the instances of this data set. Therefore, this data set can be used to train the PPD model.

5.1.4. Dwell Time Data Set. This data set concerns the relationship between dwell time of trains and corresponding passenger flow. Therefore, 346 records corresponding to 346 observed trains are included and each of them has five attributes: the total number of alighting passengers $A$, the numbers of boarding passengers entering from Entrance 1 and Entrance 2, that is, $B_1$ and $B_2$, the total number of through standees $C$, and required dwell time $D$. The former three attributes can all be obtained through accumulating the corresponding single-door data of the raw data set, while $D$ is obtained according to

$$D = \begin{cases} T & T - T_0 > 60 \\ PST + \tau_1 + \tau_2 & \text{otherwise} \end{cases}.$$ \hfill (14)

5.2. Training of SDPST Model. With the SDPST data set, ELM is used to train the SDPST model. Meanwhile, for comparison, another two popular algorithms, LMBP and SVM, are also applied on this specific regression problem. All the attributes in this data set are normalized into range $[0, 1]$ and the data set is divided into two parts: 4000 observations are used for training and the rest are used for testing. For ELM, the number of hidden nodes $p$ is gradually increased by an interval of 5 and the optimal number 65 is obtained using 3-fold cross-validation method, which is illustrated in Figure 7. Similarly, the number of hidden nodes in the BP network is also determined through repeated cross-validations. For SVM, RBF is used as kernel function and the cost parameter and kernel parameter are both chosen from set $\{2^{-10}, 2^{-9}, 2^{-8}, \ldots, 2^9, 2^{10}\}$ through repeated tests.

From this raw data set, the operation-related parameters and three useful data sets are derived.
Figure 7: Tuning the number of hidden nodes in the ELM-based SDPST model.

Table 2: Comparison of performance of ELM, LMBP, and SVM on SDPST data set.

<table>
<thead>
<tr>
<th>Algorithms</th>
<th>Number of nodes/SVs</th>
<th>Time (s) Training</th>
<th>RMSE Training</th>
<th>Time (s) Testing</th>
<th>RMSE Testing</th>
</tr>
</thead>
<tbody>
<tr>
<td>ELM</td>
<td>65</td>
<td>0.1358</td>
<td>0.0865</td>
<td>0.0945</td>
<td></td>
</tr>
<tr>
<td>LMBP</td>
<td>20</td>
<td>2.1722</td>
<td>0.0904</td>
<td>0.1218</td>
<td></td>
</tr>
<tr>
<td>SVM</td>
<td>310.24</td>
<td>4.6375</td>
<td>0.0853</td>
<td>0.1046</td>
<td></td>
</tr>
</tbody>
</table>

All the simulations are carried out in MATLAB 8.2 environment running in a Core2 Quad, 2.67 GHz CPU, and corresponding results are shown in Table 2. As shown in this table, no matter in training speed or generalization performance, ELM is remarkably better than the other two algorithms. In other words, the ELM-based SDPST performs better in estimating the single-door passenger service time.

For further comparison, a basic social force model [20] is established to simulate passengers alighting and boarding at single door of urban train. The parameters of this model are calibrated according to the observed data of a basic case, in which the numbers of alighting, boarding, and through passengers are all 5; that is, \( a = b = c = 5 \). On this basis, different cases are tested on this microsimulation model and the results are compared with the proposed model. In the test, the numbers of alighting and through passengers are all set to be 5; that is, \( a = c = 5 \). The number of boarding passengers is gradually increased and corresponding PST outputted by the microsimulation model is compared with the result estimated by the ELM-based SDPST model, which is shown in Figure 8. As can be seen, the results of the proposed model are in good accordance with the observed data. The microsimulation model fits the observed data well when \( b \leq 16 \), but it does not perform well when \( b > 16 \).

Furthermore, using the SDPST model trained by ELM, the relationship between passenger service time (PST) and corresponding factors \( (a, b, c) \) at single door is also investigated. With the other two factors fixed at 5, the variation of PST with each factor is tested. As shown in Figure 9, PST is in nonlinear relationship with each of the three factors, which is much different with previous studies [8–10].

5.3. Training of PPD Model. With the PPD data set, the PPD model is trained to describe the boarding passenger distribution on the outbound platform of Zhichulu station (Line 13). The data set is also normalized into \([0,1]\) and divided into two parts: 200 observations are used for training and the rest are used for testing. The other two algorithms, LMBP and SVM, are also applied on this data set and their performances are compared with ELM in Table 3. As can be seen, the training speed of ELM is still remarkably faster than that of the other two algorithms. As for generalization performance, ELM is similar to the SVM and slightly better than LMBP. In conclusion, the ELM-based model obtains best performance on the PPD data set.

5.4. Evaluation of Overall Estimation Model. With the above two models trained by ELM, the overall model can be used to estimate the train dwell time of Line 13 at Zhichunlu station. The proposed overall model is compared with two polynomial models. One is proposed by Lam et al. [8] and
shown as formula (15). The other is proposed by Puong [10] and shown as formula (16):

\[ D = \mu_1 + \mu_2 A + \mu_3 B, \quad (15) \]
\[ D = v_1 + v_2 \frac{A}{n} + v_3 \frac{B}{n} + v_4 \left( \frac{C}{n} \right)^3 \frac{B}{n}, \quad (16) \]

Using the dwell time data set, least squares method is used to calibrate the parameters of the above two models. Considering the outputs of these three models are all single-dimensional, the coefficient of determination which is usually denoted as \( R^2 \) is adopted to evaluate their regression performance. The model whose \( R^2 \) is closer to 1 is considered better. The results are listed in Table 4. As can be seen, the ELM-based model proposed in this paper performs much better than the other two polynomial models.

### 6. Conclusions

This paper proposed a new model to estimate urban rail dwell time. In this model, two crucial relationships among the factors of urban rail dwell time are modeled by two SLFNs, which are trained with ELM. Using a set of observed data from Beijing subway, the training of these two networks is illustrated, during which ELM is proven more effective than other two algorithms, and advantage of the proposed approach is also verified by comparing with an existing estimation model.

### Notation Definitions

- \( i \): Index of doors
- \( n \): Number of doors
- \( j \): Index of platform entrances
- \( m \): Number of platform entrances
- \( A \): Total number of alighting passengers
- \( B_j \): Number of boarding passengers entering the platform through \( j \)th entrance
- \( B \): \( m \)-dimensional column vector, whose \( j \)th component is \( B_j \); that is, \( B = [B_1, B_2, \ldots, B_m]^T \)
- \( C \): Total number of through passengers
- \( a_i \): Number of alighting passengers at \( i \)th door
- \( b_i \): Number of boarding passengers at \( i \)th door
- \( c_i \): Number of through passengers at \( i \)th door
- \( \alpha \): \( n \)-dimensional column vector, whose \( i \)th component is \( a_i \); that is, \( \alpha = [a_1, a_2, \ldots, a_n]^T \)
- \( \beta \): \( n \)-dimensional column vector, whose \( i \)th component is \( b_i \); that is, \( \beta = [b_1, b_2, \ldots, b_n]^T \)
- \( \gamma \): \( n \)-dimensional column vector, whose \( i \)th component is \( c_i \); that is, \( \gamma = [c_1, c_2, \ldots, c_n]^T \)
- \( t_f \): Passenger service time at \( i \)th door
- \( \tau_1 \): Duration of door closing process
- \( \tau_2 \): Confirmation time
- \( D \): Required dwell time.

### Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

### Acknowledgments

The authors are grateful to the editor and reviewers for their valuable suggestions which improved the paper. This work is supported by National Natural Science Foundation of China (U1361114).

### References


Submit your manuscripts at
http://www.hindawi.com