Research on Lane-level Travel Time Prediction Method Based on Gated Recurrent Neural Network

Shupeng Zhang, Deyong Guan*, Haiqing Liu

Transportation Academy, Shandong University of Science and Technology, Qingdao 266000, China
651392268@qq.com
*guandeyong@sdust.edu.cn

Abstract. In order to meet the needs of refined traffic guidance in the environment of the Internet of Vehicles, a lane-level travel time prediction method based on Gated Recurrent Neural Network is proposed. The research is based on the vehicle information collected by electronic police equipment, and analyzes and integrates the trip time of a single vehicle as model input; builds a gated recurrent unit (GRU) model deep learning framework, and performs multiple iterations through deviation correction and parameter update; The vehicle travel time data is verified and compared with the ARIMA model, BPNN model and LSTM model. The results show that the MAE and RMSE values of the GRU model are lower than other models, so it has good prediction accuracy.

1. Introduction
With the growth of private vehicles and the increase in people's travel frequency, road traffic congestion has intensified and the traffic environment has become complicated. In the current development of the Internet of Vehicles technology, in order to solve this traffic problem, related research uses the intelligent transportation system's vehicle travel time prediction method for traffic guidance. Traditional travel time prediction methods mainly include ARIMA model, Markov model, Kalman filter model and so on. Cui Qinghua et al [1] used the ARIMA model to predict the time-varying confidence interval of road travel time. Tran Duy et al [2] used Hidden Markov Model (HMM) to predict the driver's intention. The traditional travel time prediction method does not have real-time performance, cannot handle large-scale data well, and the model itself lacks the ability to update.

In recent years, with the development of artificial intelligence and deep learning, Support Vector Regression (SVR) models, Back propagation Neural Network (BPNN), wavelet neural networks and other models have been increasingly applied to travel time prediction. Miao Xu et al [3] proposed a comprehensive model of pre-order section state and two-layer BPNN to predict the bus arrival time. Yu Quan et al [4] used particle swarm optimization wavelet neural network (PSO-WNN) in order to improve the accuracy of expressway travel time prediction. Xie Jiaming et al [5] input historical data and real-time data into the neural network for traffic in Hong Kong Make predictions. Wang Xueyuan et al [6] used the cuckoo algorithm to optimize the Elman neural network, and used the collected on-board OBD data to perform map matching to complete the travel time prediction of urban roads. Huang Wenhao et al [7] proposed a deep structure composed of deep belief network (DBN) and multi-task regression to complete the traffic flow prediction. Deep learning methods can improve the predictive ability to a certain extent, but they have certain limitations in dealing with time series.
Lane-level travel time data is a non-linear time series, while cyclic neural networks (RNN) can handle variable-length time series problems, but RNNs have gradient disappearance and explosion phenomena. Current researches mostly use long and short-term memory units (LSTM) solves this problem. Chen Yuan-yuan et al [8] used the stacked long-term memory network model to learn and predict traffic conditions based on the traffic flow conditions obtained from the Web map. Petersen et al [9] combined LSTM neural network and convolutional neural network to predict the bus travel time. However, the LSTM network also has too many parameters and a complex structure. Based on the above research, this paper takes into account the complexity of lane-level travel time data, in order to improve the prediction accuracy and speed up the prediction, the gated recurrent unit neural network is applied to the lane-level travel time prediction research.

2. Travel time data collection and processing

Use the electronic police equipment installed on the road to collect data on the travel time of the vehicle. First, export the vehicle information collected by the electronic police on the road section and perform data preprocessing; then calculate the vehicle travel time between the two electronic police and record the vehicles, drive lane location; finally integrate the travel time of each vehicle section and related information about the subordinate lane. Process data according to standardized technology:

\[
\bar{x}_i = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}}, \quad i \in [1, n]
\]

Where: \( \bar{x}_i \) is the standard value of the sample data, \( x_i \) is the original sample data, \( x_{\min} \) and \( x_{\max} \), are the maximum and minimum values of the original sample data, respectively. The time series data set obtained after processing is divided into historical data training set, model test set, and method verification set.

3. Lane-level travel time prediction model

3.1. The principle of gated recurrent unit neural network

In the current deep learning related research, RNN recurrent neural network is mostly used to process time series data. This network allows the internal connection of hidden units and can handle the relationship between discontinuous data. Because RNN has the problem of gradient disappearance, and LSTM neural network can solve this problem, it is widely used in the field of time series prediction. In recent years, GRU neural network has been used more and more. Compared with LSTM, this network has fewer training parameters and better prediction effect. GRU uses a gated recurrent neural network structure, and combines the input gate and the forget gate in LSTM into one update gate, and the other is a reset gate. The update gate controls the state information retention degree at the previous moment, and the reset gate determines the degree of combination of previous information and current information.

![Figure 1 GRU internal structure diagram](image-url)
In figure 1, the input of the GRU unit at the current moment in the recurrent neural network is $x_t$. Last moment state activation value is $h_{t-1}$, The current state activation value is $h_t$; $U$ is the recursive connect weight matrix, $r_t$ and $z_t$ are the reset gate and update gate of GRU at time $t$:

$$r_t = \text{sigmoid}(U_r h_{t-1} + W_r x_t + b_r)$$

$$z_t = \text{sigmoid}(U_z h_{t-1} + W_z x_t + b_z)$$

Where: $U_r$ is the output weight matrix of the reset gate at $t-1$, $W_r$ is the input weight matrix of the reset gate, $b_r$ represents the bias vector of the reset gate; $U_z$ is the output weight matrix of the update gate at $t-1$, $W_z$ is the update gate's weight matrix, which $b_z$ represents the bias vector of the update gate.

At this time, the candidate hidden state and the current state activation value in the GRU unit are:

$$\tilde{h}_t = \tanh(U_r (r_t h_{t-1}) + W_r x_t + b_r)$$

$$h_t = z_t h_{t-1} + (1 - z_t) \tilde{h}_t$$

Where: $U_r$ is the output weight matrix of the candidate hidden state at $t-1$, $W_r$ is the input weight matrix of the candidate hidden state, and $b_r$ represents the bias vector of the candidate hidden state.

The value range of the reset gate and update gate in the GRU unit is controlled by the sigmoid function. Through the reset gate, the input of the GRU unit $x_t$ is combined with the state activation value $h_{t-1}$ at the previous moment, and the value of $r_t$ is between 0 and 1, which tends to 1, The larger the proportion of $h_{t-1}$; on the other hand, the information transfer from $h_{t-1}$ to $h_t$ in the update gate control, which is similar to the memory unit in LSTM, which memorizes the hidden state information in the RNN.

3.2. Model settings

This paper uses the travel time data of the preamble period to realize the prediction of the vehicle lane-level travel time by the GRU neural network. In order to predict the travel time of the vehicle during $t+1$ period, it is necessary to predict the travel time during $t$ period and $t-n$ period. The number of neurons in the input layer and output layer in the GRU unit is $n+1$ and 1, respectively, and the model input is expressed as $x = (x_1, x_2, x_3, \cdots, x_n)$.

Initialize the connection weight value between each layer and the bias vector of each neuron, select the appropriate activation function, and determine the mean square error as the loss function:

$$H_{MSE} = \frac{1}{N} \sum_{i=1}^{N} (o_i - p_i)^2$$

Where: $N$ is the number of training samples of the model, $o_i$ is the observed value, and $p_i$ is the predicted (output) value.

Input the training set and the validation set into the adaptive gradient optimization model for algorithm training. Set the global learning rate as $\eta$, the model initialization parameter as $w_0$, and the time step count $t$; the gradient calculation performed in the loop is:

$$g_t = \nabla J(w_{t-1})$$

Where: $g_t$ is the gradient value at the $t$th time step, $w_{t-1}$ is the weight matrix at the $t-1$ time step, and $J$ is the cost function of the gradient model.

After the time step is updated, the first-order and second-order biased estimation matrices are:

$$m_t = \varphi m_{t-1} + (1 - \varphi) g_t$$

$$v_t = \varphi v_{t-1} + (1 - \varphi) g_t^2$$

Where: $\varphi$ is the momentum parameter.
Where: $m_1$, $v_1$ are the first-order and second-order momentum terms, respectively, and $\varphi_1$, $\varphi_2$ are the exponential decay rates of the first- and second-order moment estimates, respectively.

The deviation correction and parameter update of the first-order and second-order matrices are:

$$m'_1 = \frac{m_1}{1 - \varphi_1}$$  \hspace{1cm} (10)

$$v'_1 = \frac{v_1}{1 - \varphi_2}$$  \hspace{1cm} (11)

$$\Delta w = \frac{\eta}{\sqrt{v'_1 + \theta}} m'_1$$  \hspace{1cm} (12)

Where: The default value is $10^{-8}$.

To update the parameters:

$$w_{i+1} = w_i + \Delta w$$  \hspace{1cm} (13)

The trained model is tested and de-standardized, and finally the predicted value of travel time is obtained. In order to better evaluate the prediction accuracy of the model, this paper adopts two indicators, mean absolute error (MAE) and root mean square error (RMSE), to evaluate the prediction results of the model.

$$MAE = \frac{1}{N} \sum_{i=1}^{N} |f_i - y_i|$$  \hspace{1cm} (14)

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (f_i - y_i)^2}$$  \hspace{1cm} (15)

Where: $y_i$ is the actual observation value, $f_i$ is the predicted value, and $N$ is the sample size.

4. Case analysis

According to the electronic bayonet data of Hong Kong Middle Road, Shinan District, Qingdao, the road section is two-way four-lane, the length is 825m, the distance between electronic bayonet is 200m, and the data upload time is 15s/time. This article selects the vehicle travel time data during the peak period from 10:00 to 11:00 in the morning for a continuous month as the research object, and divides the data into historical data training set, model test set, and historical data training set at a ratio of 5:2:3. The historical data training set is shown in the figure 2:

![Figure 2 Lane-level travel time data map](image)

In order to verify the prediction effect of the GRU model, the study uses the traditional differential autoregressive moving average ARIMA model, shallow neural network BPNN model, and LSTM model to predict the same test sequence. According to the AIC criterion, the parameters of the ARIMA model are selected as follows $p = 1 \; d = 3 \; q = 2$; the number of neurons in the BPNN model is 240, and the hidden layer is 2 layers; the number of hidden neurons in the LSTM is set to 136, and the number
of neurons in the GRU model is 105, using the activation function sigmoid and tanh. Use Adam model for optimization. This article uses PYTHON to program the operating platform for Inter(R)Core(TM) i5-8250 CPU@2.4GHz, 8GB memory.

![GRU model prediction results](image1.png)

**Figure 3 GRU model prediction results**

The solid line and the dashed line in Figure 3 represent the original data and model prediction results, respectively. According to the curve in the figure, it can be known that the GRU model has a better prediction effect.

![Training and verification data loss curve](image2.png)

**Figure 4 Training and verification data loss curve**

It can be seen from Figure 4 that as the time step increases, the loss values of the historical data training set and the historical data training set become smaller and smaller and then stabilize.

| Model  | 3     | 6     | 9     | 12    |
|-------|-------|-------|-------|-------|
| ARIMA | 83.25 | 75.36 | 79.67 | 82.48 |
| BPNN  | 71.52 | 65.15 | 67.43 | 68.45 |
| LSTM  | 53.67 | 46.48 | 49.12 | 51.55 |
| GRU   | 44.28 | 33.62 | 35.20 | 38.65 |

**Table 1 Comparison of MAE indicators of different models**

| Model  | 3     | 6     | 9     | 12    |
|-------|-------|-------|-------|-------|
| ARIMA | 98.46 | 91.06 | 93.97 | 96.39 |
| BPNN  | 82.54 | 76.35 | 78.83 | 80.29 |
| LSTM  | 69.31 | 63.46 | 65.43 | 67.54 |
| GRU   | 56.57 | 50.91 | 52.56 | 54.92 |

**Table 2 Comparison of RMSE indicators of different models**
The MAE and RMSE indicators of ARIMA model, BPNN model, LSTM model and GRU model at different time intervals are shown in Tables 1 and 2. Among them, the LSTM model has the same structure as the GRU model in this paper. It can be seen from the table that compared with other models, the model in this paper obtains the highest prediction accuracy, and the two error indicators are significantly lower than other models. As the time step increases, the prediction accuracy of the model first increases and then decreases. When the model input length is 6, MAE and RMSE indicators are the lowest. Therefore, the use of the GRU model can better predict the lane-level travel time.

5. Conclusion
This paper proposes a lane-level travel time prediction method based on a gated recurrent neural network. A time series set is established based on the nonlinear characteristics of the lane-level travel time data, and the GRU neural network is applied to the travel time prediction. It provides a new way of thinking in terms of class travel time prediction. Therefore, the collected data is used to predict the travel time of Hong Kong Middle Road, Shinan District, Qingdao. The experimental results show that the MAE and RMSE indicators of the GRU model are significantly better than other models, and the effect of predicting travel time at the lane level is better.

Acknowledgments
This work was financially supported by the Project of Shandong Provincial Natural Science Foundation (ZR2019QF017).

References
[1] Cui Qinghua, Xia Jingxin. Time-varying confidence interval forecasting of travel time for urban arterials using ARIMA-GARCH model[J]. Journal of Southeast University,2014,30(3):358-362(in Chinese).
[2] Tran Duy, Sheng Weihua, et al. A Hidden Markov Model based driver intention prediction system[C]. 2015 IEEE International Conference on Cyber Technology in Automation, Control, and Intelligent Systems (CYBER). IEEE, Shenyang China, 2015:115-120(in Chinese).
[3] Miao Xu, Wang Zhongyu, et al. Double-layer BPNN prediction model of bus arrival time considering the state of preceding sections[J]. Journal of Transportation Systems Engineering and Information Technology, 2020,20(02):127-133(in Chinese).
[4] Yu Quan, Sun Yao. Travel time prediction based on particle swarm optimization wavelet neural network[J]. Transportation research, 2020,6(02):40-47+59(in Chinese).
[5] Xie Jiaming,Choi Yi-King, et al. Hybrid traffic prediction scheme for intelligent transportation systems based on historical and real-time data[J]. International Journal of Distributed Sensor Networks,2017,13(11) :1-11(in Chinese).
[6] Wang Xueyuan. Study on Travel Time Prediction Model of Urban Road Based on OBD Data Map Matching[D]. Bei Jing: Beijing jiaotong university,2019(in Chinese).
[7] Huang Wenhao, Song Guojie, et al. Deep Architecture for Traffic Flow Prediction: Deep Belief Networks With Multitask Learning[J]. IEEE Transactions on Intelligent Transportation Systems,2014,15(5):2191-2201(in English).
[8] Chen Yuan-yuan, Lv Yisheng, et al. Long short-term memory model for traffic congestion prediction with online open data[C]. 2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC), 2016:132-137(in English).
[9] Niklas Christoffer Petersen, Filipe Rodrigues, et al. Multi-output bus travel time prediction with convolutional LSTM neural network[J]. Expert Systems With Applications,2019,120:426-435(in English).