Proceedings of the 2023 International Conference on Machine Learning and Automation
DOI: 10.54254/2755-2721/30/20230095

Prediction-based early warning system for overflow of people in metro stations

Di Lu

Southeast University, 2 Southeast University Road, Jiangning District, Nanjing, Jiangsu, 211189,China

204758365@qq.com

Abstract. As cities rapidly urbanize, handling the management of crowded transportation hubs, notably metro stations, has become an immediate concern. High passenger traffic can lead to severe risks, including stampedes. While many past passenger flow forecasting systems aim to enhance prediction accuracy, the inherently noisy nature of passenger flow data makes it challenging for existing technologies to provide stable and precise predictions, making passenger flow management based solely on these models risky. This paper introduces a novel system that mitigates this risk by integrating a predictive model with managerial methods. The proposed management framework formulates a unique model for each station, determining a risk deviation coefficient grounded in the station's historical prediction accuracy. Station management is then holistically executed based on this coefficient and the predictive model. This paper employs the LSTM model for station-specific passenger flow prediction and defines the risk-related parameter α, taking prediction accuracy into account. This adjusted LSTM prediction is then utilized to proactively streamline resource allocation, targeting improved passenger safety and overall station experience.

Keywords: LSTM, station flow prediction, deep learning.

1. Introduction

As urbanization intensifies, city transit issues have increasingly come into spotlight. Especially for urban transport hubs like stations, which accommodate a large flow of individuals, it becomes critical to precisely anticipate the crowd concentration to enhance station administration and improve commuting experience. Hence, this study aims to construct an intelligent station traffic prediction system grounded on long and short-term memory (LSTM) networks, a robust aid for station management.

LSTM, a specialized structure of recurrent neural networks (RNN), is devised to counter the vanishing gradient issue encountered when conventional RNNs process lengthy sequences. The proposed intelligent station crowd estimation system harnesses the long short-term memory proficiency of the LSTM model, amalgamated with real-time surveillance data and historical records within the station, to project the future population within the station in real time. This provides the station master with a visual representation of the projected numbers.

Nevertheless, the forecasted results frequently deviate from the reality, especially during bouts of high footfall in the station, leading to an alarmingly low predicted value by the model. This can potentially mislead the station manager into assuming a safe crowd number in the forthcoming period and might preclude essential measures. This study introduces an innovative tactic to bolster the reliability of passenger flow prediction in railway stations by integrating a coefficient-based amendment for LSTM models. Unlike traditional techniques that depend solely on LSTM predictions, the proposed model utilizes historical passenger flow data to compute station-specific coefficients. These coefficients, as modification factors, refine LSTM predictions allowing better management corresponding to each station's unique features. The paper delineates the process of deriving and applying these coefficients, emphasizing their importance in supporting station management personnel with proactive decision making.

In this article:

1) The prediction model is constructed, which can be used to predict the passenger flow of the station by using the recent historical data of the station.

2) Parameter α is introduced to modify the predicted value of the model and improve the accuracy of the model.

3) The management framework is meticulously designed to utilize the adjusted LSTM forecasts for proactive optimization of resource allocation. This includes strategically deploying personnel and effectively utilizing infrastructure in advance, specifically before peak hours, to proactively mitigate the risks associated with high passenger flow density.

2. Literature review

Short-term passenger flow prediction plays a key role in a transportation system. The predictions are beneficial for several areas of system management, like operation strategizing and the planning of passenger flow control at stations. Recently, deploying deep learning for passenger flow forecasts is emerging as a focal point in research. For instance, Wei amalgamated empirical mode decomposition (EMD) with backpropagation neural network (BPNN) and implemented this EMD-BPNN hybrid forecast approach for short-term passenger flow in subway systems [1]. Polson et al. proposes an architecture that uses regularization to fit a linear model and a series of tanh layers [2]. In transportation flow, the prediction of sudden non-linear shifts due to transitions between free movement, breakdown, recovery, and overcrowding is crucial. Furthermore, incorporating object detection technology with neural networks can help estimate the count of remaining passengers on the subway platform. Sipetas et al. deployed image processing and object detection software to derive the number of passengers lingering on the platform from surveillance footage. [3].

Several other researches have employed encoder models. Ke et al. pointed out in his study that the autoencoder, a unique structure of deep neural network (DNN), holds the capability to derive non-linear traits embedded within the input through profound abstraction. [4]. In this study, various features such as time, scene, and passenger flow are synergistically incorporated and trained in distinct stacked autoencoders (SAEs). The pre-trained SAEs are subsequently utilized to initialize a supervised deep neural network (DNN) with real-time passenger flow data as the label. Nevertheless, Hao et al. suggests an innovative end-to-end deep learning framework that leverages attention mechanisms embedded in sequence-to-sequence models as the fundamental structure. [5]. The model is used to predict the number of people getting off at each station in the near future, and a large number of real data are used to verify the model, which also has good scalability and robustness.

Some studies show that the short-term prediction of large passenger flow needs to consider the spatial dependence, time dependence and exogenous dependence. Zhang et al. introduces a hierarchical modeling framework for passenger flow prediction, comprising a two-layer fuzzy model. The global model predicts the overall situation, while the local model predicts changes in passenger flow due to specific factors, such as events and weather. [6]. Some studies have focused on pedestrian behavior. Zhou et al. discussed the behavior of waiting passengers through on-site observation and video recording of two subway stations in Beijing under normal conditions, as well as the intervention of tour guides. Under normal and intervention conditions, traffic behavior parameters were collected and analyzed to evaluate the influence of manual intervention on waiting passenger behavior [7].

To cater to the decreased willingness of passengers to board crowded trains during the COVID-19 pandemic, there is an increased need for real-time congestion information. Hence, the development of a real-time traffic flow prediction model becomes crucial. Noursalehi et al. suggested the implementation of a decision support platform that offers real-time congestion predictions for trains and platforms, effectively communicating this information to passengers and considering their reactions to the given information [8]. In another study, Wang et al. used large-scale smart card data to simulate changes in passenger flow in extended subway sections [9]. A semi-supervised joint training model (S-MLR-XGBoost) is proposed to address the issue of limited training samples. Wen et al. utilizes time series decomposition for transfer learning, aiming to enhance the accuracy of short-term passenger flow prediction in high-speed railways. Initially, the time series is decomposed into linear and nonlinear components. The linear component is predicted using the SARIMA model, while the nonlinear component is transformed into feature label samples. Feature selection is then applied to facilitate transfer learning. [10].

3. Methodology

3.1. Proposed framework

In the training and prediction stage of the model, the sensor collects both historical and real-time data on the number of people entering the station, as depicted in Figure 1. The historical data is stored for training the time series model, while the real-time data is inputted into the trained model to predict future passenger flow. The model's predictions on the test set are uniformly stored for evaluation purposes.

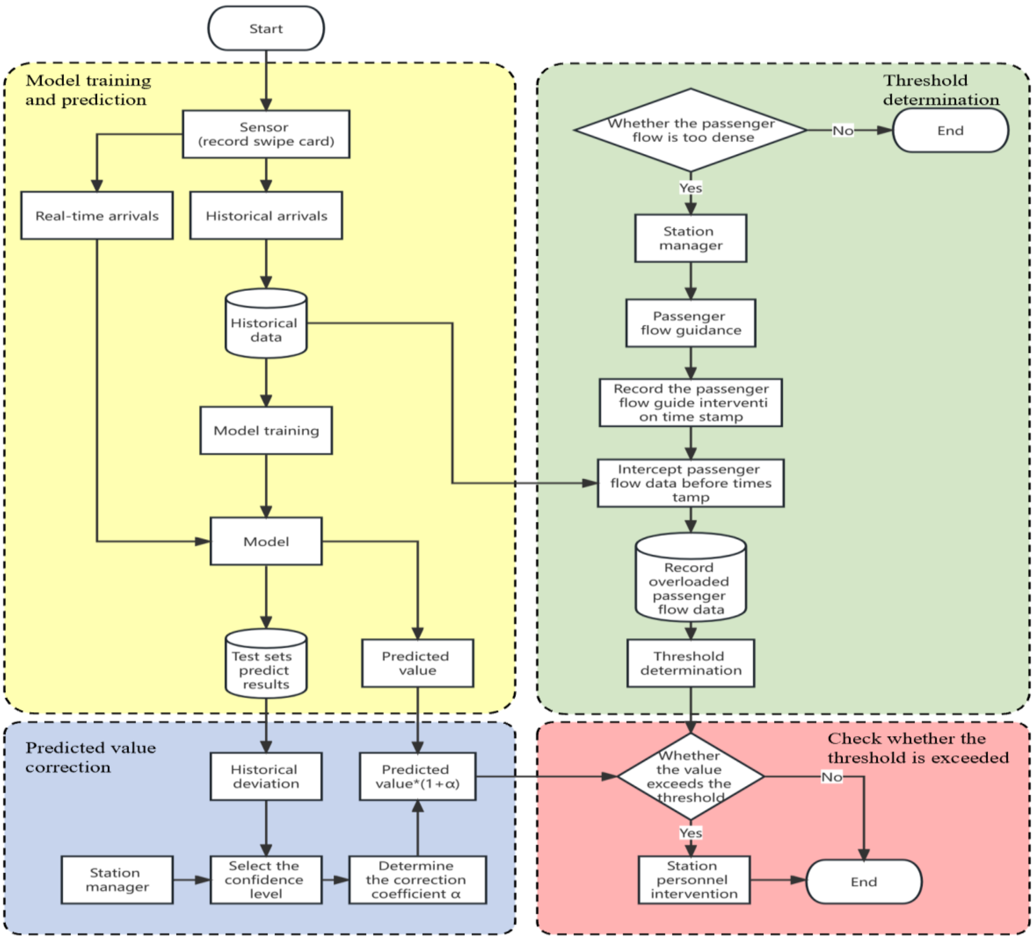

Figure 1. Structure of station management system.

In the threshold determination stage, the station management personnel first judge whether the passenger flow of the station is congested according to the experience, and then judge the passenger flow of the station is too dense, and then the station management personnel intervene and guide the passenger flow. After the intervention of the manager, the station manager records the time stamp of the intervention, obtains the passenger flow data before the time stamp from the historical data, and determines the threshold of passenger flow overload.

In the predicted value correction stage, due to the deviation of the model prediction, it is necessary to correct the predicted value. The historical deviation is obtained by comparing the predicted results of the test set with the true values. At the same time, the station manager selects the confidence degree of the model, and the system selects different correction coefficients α according to different confidence degrees. Finally, the prediction results are corrected.

In the stage of judging whether the value exceeds the threshold, the system determines whether the revised predicted value exceeds the threshold. If the value exceeds the threshold, the station management personnel will intervene to manage it.

3.2. Model selection

3.2.1. LSTM model. An adaptive gating mechanism is used in LSTM to selectively forget and update information using a sigmoid neural network layer. Feature extraction of the input vector is performed using LSTM to extract the confirmed cases count. The LSTM model includes three gated units: the forget gate, input gate, and output gate. (Figure 2).

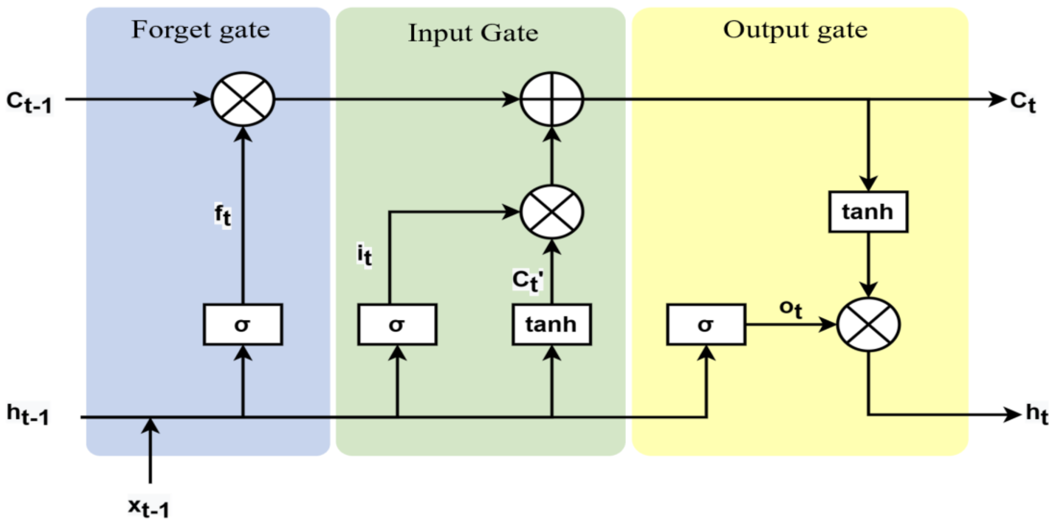

Figure 2. The structure of LSTM.

The forget gate selects the information to be forgotten from the cell state based on the current input x_t and the previous output h_{t-1}.

$$f_t = \sigma\left(W_f[h_{t-1}, x_t] + b_f\right) \tag{1}$$

The calculation process involves selecting forgotten information from the cell state based on x_t and h_{t-1} with the help of the forget gate. The input gate determines the new information stored in the cell state, which is then controlled by the tanh function based on the update degree.

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i) \tag{2}$$

$$c_t' = \varphi(W_c[h_{t-1}, x_t] + b_c) \tag{3}$$

$$c_t = f_t \cdot c_{t-1} + i_t \cdot c_t' \tag{4}$$

The current output h_t is determined by x_t and h_{t-1}, which are used to filter and calculate the new unit state through the output gate. The calculation formula is:

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o) \tag{5}$$

$$h_t = o_t \cdot \varphi(c_t) \tag{6}$$

In the formula, f_t, i_t, and o_t represent the forget gate, input gate, and output gate respectively. c_t represents the memory unit. x and h refer to the input layer and hidden layer respectively. W and b represent the weight matrix and bias vector respectively. σ and φ represent the sigmoid and tanh activation functions respectively.

3.2.2. Evaluation index. Root Mean Square Error (RMSE)

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2} \tag{7}$$

Mean Absolute Error (MAE)

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|\hat{y}_i - y_i| \tag{8}$$

Mean Absolute Percentage Error (MAPE)

$$MAPE = \frac{100\%}{n}\sum_{i=1}^{n}\left|\frac{\hat{y}_i - y_i}{y_i}\right| \tag{9}$$

R-Square (R2)

$$R2 = 1 - \frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{\sum_{i=1}^{n}(\bar{y}_i - y_i)^2} \tag{10}$$

Adjusted R-Square (Adjusted R2)

$$Adjusted\ R2 = 1 - \frac{(1 - R^2)(n-1)}{n-p-1} \tag{11}$$

$\hat{y}$ is predicted value, y is observed value, n is the sample quantity, p is the characteristic number.

4. Result analysis

4.1. Experimental result

In this study, the data of passengers entering and leaving 64 stations in Chongqing rail transit from June 1, 2018 to December 31, 2018 (214 days in total) were selected, with a statistical interval of 1 minute. The original data were combined at a time interval of 5 minutes, and finally the data of each station was sorted into a data with 61,632 time points and the corresponding number of people entering and leaving the station.

The number of LSTM units is set as 50. The characteristic values of the past 5 time steps are taken as the input sequence (n_{in}=5), and the model is made to predict the characteristic value of the next time step ($n_{out} = 1$), where the characteristic value is the sum of the number of people entering and leaving the station, and the characteristic number is 1(M=1). The number of original data rows is 61632(N=61632), and the original data is converted into data for supervised learning. After conversion, the data shape used to input LSTM model is $(n - n_{in} - n_{out} + 1, M * 5)$, namely (61627,5). The loss function was defined as root mean square error (RMSE) and mean absolute error (MAE), the optimizer was set as *adam*, batch_size as 20, and the number of training rounds (epochs) as 50.

The first 85% of the data is the training set (52,388 data pieces), and the last 15% is the verification set (9244 data pieces).

Using LSTM model training, get results.

4.2. Experimental design

Take the data of Station 102 on the 183rd day (which is Friday, November 30, 2018) and analyze the generated data.

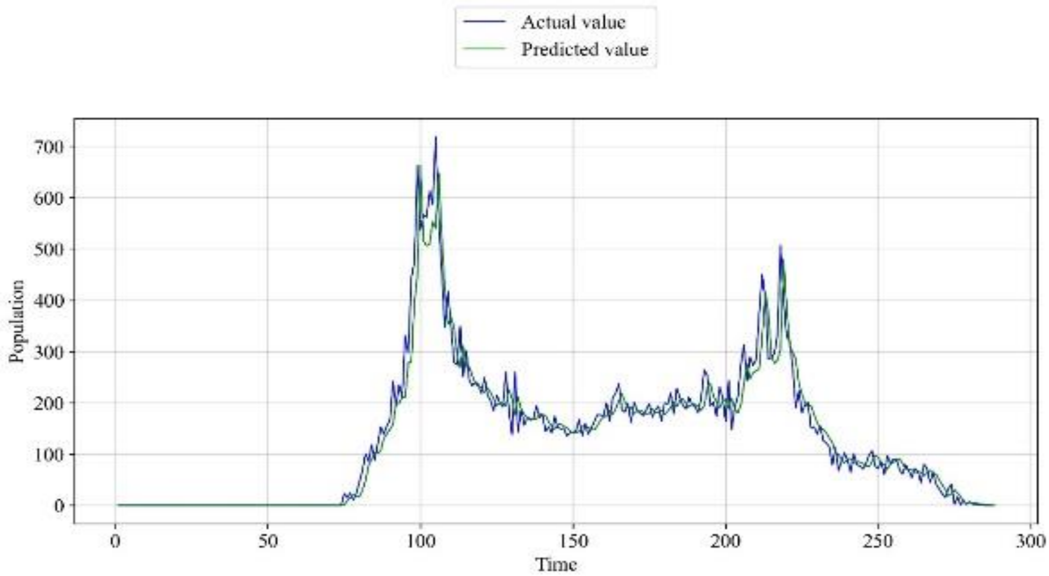

Figure 3. Traffic at Station 102 on weekdays.

Around Time 75 (6:15 a.m.) in Figure 3, passengers started entering the station. The blue line represents the actual count of people entering the station, while the green line represents the predicted count. As you can see, the green line shows the same trend as the blue line. At the 100th Time (corresponding to 8:20 a.m.), the morning rush peaks with about 600 people. However, at about 120-digit Time (corresponding to 10:00 AM), the number of people entering and leaving the station recovered to a flat peak of about 200 people. At about 220 Time (corresponding to 18:20 PM), the evening peak peak was about 400 people.

The data of the inbound and outbound passenger flow of this station on Saturday (the 184th day, i.e. December 1, 2018) is obtained as Figure 4:

DOI: 10.54254/2755-2721/30/20230095

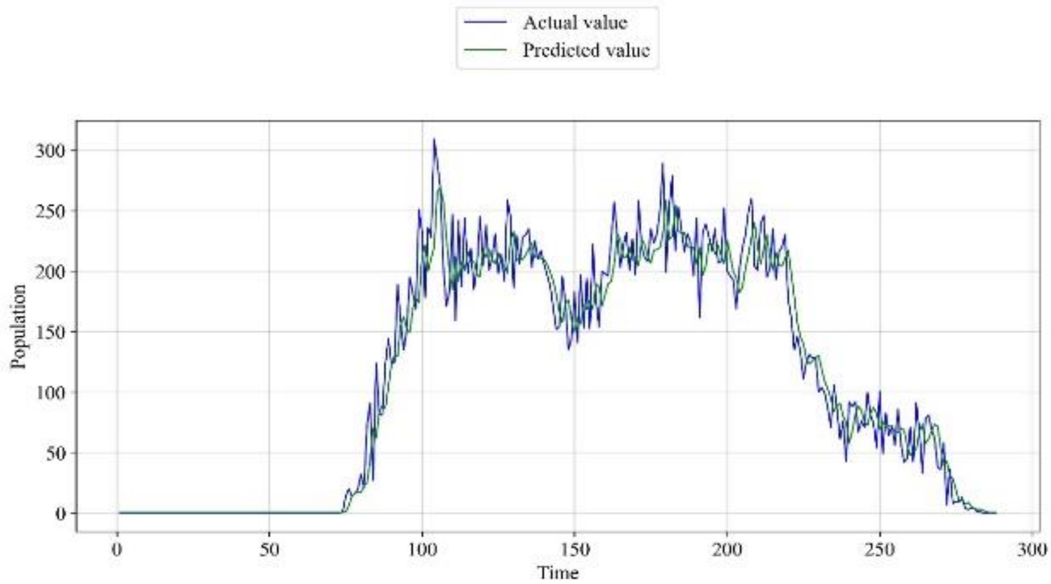

Figure 4. Traffic at Station 102 on weekends.

It can be found that the peak and peak time of human flow on weekends is similar to that on weekdays. However, the peak crowds are much smaller than on weekdays, for example, the morning peak crowds (about 300 people) are only half as large (about 600 people) as on weekdays. By contrast, traffic on weekdays is roughly the same -- about 200 people.

In this model, R^2=0.922, Adjusted R^2=RMSE=36.396, MAE=20.780. When deleting the actual value of 0, MAPE= 26.497%. It can be found that the value of R^2 has exceeded 0.8, so the prediction effect is good.

It is worth noting that the curve representing the predicted value in the figure can well reflect the changing trend of the human flow in the next time period. The LSTM model appears to struggle in accurately predicting large fluctuations in passenger flow, as evidenced by the predicted values being lower than the actual values during sudden increases in human flow and higher than the actual values during sudden decreases in human flow.Similarly, take the data of the 183rd day (Friday) of Station 108 and analyze the generated data.

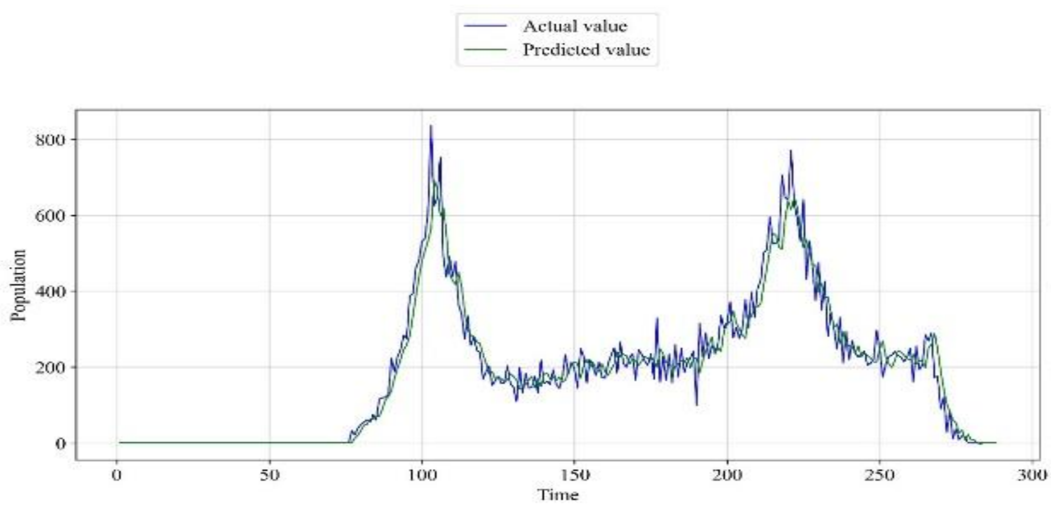

Figure 5. Traffic at Station 108 on weekdays.

DOI: 10.54254/2755-2721/30/20230095

It can be seen that during the morning peak, the time series model's prediction of the peak is always smaller than the actual value (it can be seen that the green line is always below the blue line during the morning peak at about the 100th time and the evening peak at about the 220th time).

The data of the inbound and outbound passenger flow of this station on Saturday (the 184th day, i.e. December 1, 2018) is obtained as Figure 6:

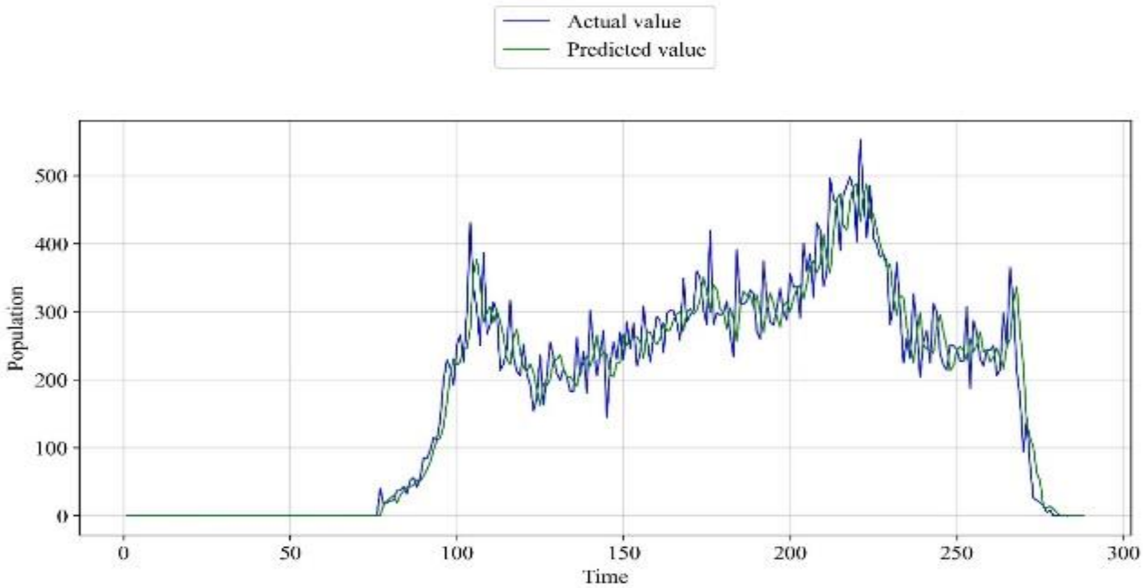

Figure 6. Traffic at station 108 on weekends.

It can be seen that the passenger flow in peak hours is significantly reduced, and the passenger flow tends to be evenly distributed throughout the day.

5. Correction factor

From the above results, we can find that when the number of passengers in and out of the station reaches its peak in a day, the predicted value of the model is always smaller than the actual value, which will increase the potential security risks in the period of a day with large passenger flow. To this end, we define a correction factor for the management system to correct the predicted value to be closer to the actual value. Since the correction coefficient is to reduce the forecast deviation of the model when the passenger flow is high, we only consider the case where the predicted value is greater than 10 when defining the correction coefficient.

The correction factor is defined as α.

$$\alpha_i = \frac{|y_i - \hat{y}_i|}{\hat{y}_i}$$

Where, $\hat{y}$ is the predicted value and y is the observed value. The values used to calculate the correction coefficient α are the predicted values and observed values corresponding to the non-zero time of the predicted passenger flow of the day.

From the definition of the correction coefficient, we can find that when $y > \hat{y}$, $y = (1+\alpha)\hat{y}$. The difference between the actual value and the predicted value can still be measured by taking the absolute value of $\hat{y} - y$ if $y < \hat{y}$. We can assume that there is a possible actual value y. We calculate a series of A according to all the periods when the predicted value is greater than 10 in the verification set (9,244 data in total) of the training data of the model. And sort it out.

Taking No. 102 station as an example, after removing the samples whose predicted value is no more than 10, the value of 6,125 a copies is obtained, as shown in Table 1 below.

DOI: 10.54254/2755-2721/30/20230095

Table 1. The values of α.

	α
Sample number	6125
Mean value	0.1729397
variance	0.1622483
Minimum value	3.77e-07
1% subsite	0.0023235
5% subsite	0.0122375
10% subsite	0.0245471
25% subsite	0.0632314
50% subsite	0.1298933
75% subsite	0.2324098
90% subsite	0.362056
95% subsite	0.4684686
99% subsite	0.8196731
Maximum value	1.597572

Station managers can determine different points based on different levels of confidence. For example, the number of people predicted at a certain time in station 102 is 100. If the station management hopes that the probability that the actual value will not exceed the predicted value is 75%, then the corresponding α is 0.23, and the predicted number should be revised to 123. If this probability is expected to be 90%, then the corresponding α is 0.36, then the predicted number should be revised to 136. The higher the probability that the actual value will not exceed the predicted value, the higher the corresponding α value. Under this condition, the security of the model will be higher, but the corresponding management cost will also increase.

6. Conclusion

In this paper, an intelligent station flow prediction system with long and short term memory is established by using LSTM. The system combines real-time monitoring data with historical data to provide accurate and reliable crowd prediction information for station managers. At the same time, in order to avoid the situation of low predicted value, the author introduces the introduction of correction coefficient α to correct the predicted value of the system. This revision uses historical passenger flow data to generate parameter α, which helps increase the fault tolerance and security of the system by reducing the deviation between predicted and actual values. Because the station manager can flexibly choose the value of α by choosing the safety degree of the model. The parameter α introduced in this study not only makes the prediction model highly intelligent, real-time and accurate, but also has strong scalability and universality, which can be widely used in various station and yard scenarios.

By accurately predicting future passenger flow, station managers can dynamically allocate resources, enhance operational efficiency, improve service quality and passenger experience, thereby alleviating urban traffic congestion and promoting sustainable traffic development.

References

[1] Y. Wei and M.-C. Chen, ‘Forecasting the short-term metro passenger flow with empirical mode decomposition and neural networks’, Transportation Research Part C: Emerging Technologies, vol. 21, no. 1, pp. 148–162, Apr. 2012, doi: 10.1016/j.trc.2011.06.009.

[2] N. G. Polson and V. O. Sokolov, ‘Deep learning for short-term traffic flow prediction’, Transportation Research Part C: Emerging Technologies, vol. 79, pp. 1–17, Jun. 2017, doi: 10.1016/j.trc.2017.02.024.
[3] C. Sipetas, A. Keklikoglou, and E. J. Gonzales, ‘Estimation of left behind subway passengers through archived data and video image processing’, Transportation Research Part C: Emerging Technologies, vol. 118, p. 102727, Sep. 2020, doi: 10.1016/j.trc.2020.102727.
[4] J. Ke, H. Zheng, H. Yang, and X. (Michael) Chen, ‘Short-term forecasting of passenger demand under on-demand ride services: A spatio-temporal deep learning approach’, Transportation Research Part C: Emerging Technologies, vol. 85, pp. 591–608, Dec. 2017, doi: 10.1016/j.trc.2017.10.016.
[5] S. Hao, D.-H. Lee, and D. Zhao, ‘Sequence to sequence learning with attention mechanism for short-term passenger flow prediction in large-scale metro system’, Transportation Research Part C: Emerging Technologies, vol. 107, pp. 287–300, Oct. 2019, doi: 10.1016/j.trc.2019.08.005.
[6] Q. Zhang, X. Liu, S. Spurgeon, and D. Yu, ‘A two-layer modelling framework for predicting passenger flow on trains: A case study of London underground trains’, Transportation Research Part A: Policy and Practice, vol. 151, pp. 119–139, Sep. 2021, doi: 10.1016/j.tra.2021.07.001.
[7] M. Zhou, S. Ge, J. Liu, H. Dong, and F.-Y. Wang, ‘Field observation and analysis of waiting passengers at subway platform — A case study of Beijing subway stations’, Physica A: Statistical Mechanics and its Applications, vol. 556, p. 124779, Oct. 2020, doi: 10.1016/j.physa.2020.124779.
[8] P. Noursalehi, H. N. Koutsopoulos, and J. Zhao, ‘Predictive decision support platform and its application in crowding prediction and passenger information generation’, Transportation Research Part C: Emerging Technologies, vol. 129, p. 103139, Aug. 2021, doi: 10.1016/j.trc.2021.103139.
[9] K. Wang, B. Guo, H. Yang, M. Li, F. Zhang, and P. Wang, ‘A semi-supervised co-training model for predicting passenger flow change in expanding subways’, Expert Systems with Applications, vol. 209, p. 118310, Dec. 2022, doi: 10.1016/j.eswa.2022.118310.
[10] K. Wen, G. Zhao, B. He, J. Ma, and H. Zhang, ‘A decomposition-based forecasting method with transfer learning for railway short-term passenger flow in holidays’, Expert Systems with Applications, vol. 189, p. 116102, Mar. 2022, doi: 10.1016/j.eswa.2021.116102.