J. COMBIN. MATH. COMBIN. COMPUT. 127a (2025) 6215--6228

Journal of Combinatorial Mathematics and Combinatorial Computing
www.combinatorialpress.com/jcmcc

Optimal electric bus frequency setting combining LSTM prediction and two-layer planning

Yixing Bao[1, ✉]

[1] *Department of Data and Systems Engineering, The University of Hong Kong, Hong Kong, 999077, China*

ABSTRACT
As an environmentally friendly and efficient public transport, the optimization of the operating frequency of electric buses is of great significance for improving passenger satisfaction and reducing operating costs. This paper proposes an optimal electric bus frequency setting method that combines LSTM prediction and two-layer planning. First, LSTM neural network is utilized to predict the passenger flow of electric buses. Second, a two-layer planning model is constructed, with the upper model aiming at frequency optimization and the lower model aiming at electric bus frequency setting. Finally, this two-layer planning model is solved by genetic algorithm to obtain the optimal electric bus frequency setting. The inbound and outbound passenger flow data of the 5th station of 363 electric bus in Q city are used for practical verification. The prediction results of the LSTM model on inbound and outbound passenger flow on weekdays and natural days are basically consistent with the actual values. The optimal frequency of 62 trips was solved using genetic algorithm. The maximum deviation of the actual capacity supply from the actual capacity demand curve is only 0.09% when the frequency setting is verified under the scenario of thousands of passenger flows. From the above analysis, it is shown that it is practical to design the optimal electric bus frequency using LSTM prediction and two-layer planning model.

Keywords: LSTM neural network, two-layer planning model, genetic algorithm, electric bus frequency

1. Introduction

LSTM is a variant of recurrent neural network (RNN) that can efficiently deal with long-term dependencies in time-series data. In LSTM, the states of neurons are designed to memorize information from previous time steps in order to correctly predict the output at future time steps [1-4]. Bi-level programming is a system optimization problem with a two-level recursive structure, where the upper level problem and the lower level problem have their own decision variables, constraints and objective functions. Bi-level system optimization studies the planning and management problems

✉ Corresponding author.
E-mail address: sjmy4046@connect.hku.hk (Y. Bao).
Received 25 March 2024; Revised 05 June 2024; Accepted 15 December 2024; Published Online 15 April 2025.
DOI: 10.61091/jcmcc127a-342

of systems with two levels [5-8]. The upper-level decision maker only guides the lower-level decision maker through his/her own decision, and does not directly interfere with the lower-level decision making. The lower level decision maker only needs to take the upper level decision as a parameter, and he is free to make decisions within his own possibilities [9]. The combination of LSTM and two-layer planning is of great significance in the field of electric buses.

Traditional fuel vehicles cause serious pollution to the environment due to the consumption of large amounts of oil. In today's increasingly environmentally friendly world, electric buses have received extensive attention from all over the world due to their advantages in energy security, environmental protection and sustainable development [10-13]. With the growing global concern about air quality and carbon emissions, more cities are focusing on electric buses as a solution to replace traditional fuel buses. In the rapid development of electric buses, the design of electric motor, control technology and the comprehensive performance of the vehicle are important factors constraining the power, economy and range of electric buses [14-17]. Therefore, the optimal electric bus frequency setting is of great significance, and the combination of LSTM prediction and two-layer planning can play an important role in this process.

The study designed the optimal electric bus frequency setting method by combining LSTM prediction and two-layer planning. The passenger flow of electric buses is accurately predicted by LSTM neural network, which provides data support for the frequency setting of electric buses. Further, a two-layer planning model is constructed in this paper, with the upper model and lower model aiming at frequency optimization and electric bus frequency setting, respectively, and the model takes into account the passenger cost and enterprise operation cost. Finally, the genetic algorithm is used to solve the two-layer planning model, and the feasibility of the optimal electric bus frequency is analyzed through practical verification.

2. LSTM-based passenger flow prediction

LSTM model is a time series model, electric bus passenger flow itself has time series characteristics, the current passenger flow is affected by the historical passenger flow sequence, so LSTM can effectively use the historical passenger flow sequence to predict the future passenger flow.

2.1. Data set construction

For the time series can be mined from three time dimensions (hourly, daily, weekly) of historical passenger flow information, the specific sequence composition is as follows:

1) Preceding time flow sequence. Preceding time series can better reflect the changes of passenger flow in a short period of time, but it is most likely to be affected by the error data, and at the same time, it will ignore the time distribution characteristics of passenger flow data itself, such as the difference between morning peak, evening peak and flat peak, as well as the difference between weekday and non-weekday passenger flow. From the historical passenger flow correlation analysis of the first [1, 3, 4, 12, 13, 15] hours of historical passenger flow and the current passenger flow correlation is strong, so the use of the above historical moments of the passenger flow data composed of the sequence as an input, then the data of a sample input vector $HX(d,h)$ is:

$$HX(d,h) = [S(d,h-15), S(d,h-12), \cdots, S(d,h-1)] \tag{1}$$

Sample Label $hy(d,h)$ is:

$$hy(d,h) = S(d,h) \tag{2}$$

2) Preceding day passenger flow sequence. Considering that it is easy to be interfered by errors and there is obvious regularity in the distribution of passenger flow over time in a day, the sequence of passenger flow on the preceding days extracts the passenger flow data at the same moment of the neighboring days, which can better reflect the distribution of passenger flow in time, but ignores the

differences between the passenger flow on weekdays and holidays. From the historical passenger flow correlation analysis of the preceeding days of passenger flow sequence and the current moment of passenger flow are strongly correlated, easy to deal with the use of the first four days of the same moment of passenger flow data composed of the sequence as an input, then the data of a sample of the input vector DX_d^h is:

$$DX(d,h) = [S(d-4,h),\cdots,S(d-1,h)] \tag{3}$$

Sample Label $dy(d,h)$ is:

$$dy(d,h) = S(d,h) \tag{4}$$

3) Preceding Weekly Passenger Flow Sequence. Considering the difference of passenger flow between week and weekend, the preorder weekly passenger flow sequence extracts the data of the same moment of the same week in the neighboring weeks, which can better reflect the difference of passenger flow between week and weekend, but the statistical time span is large, and there is no way to reflect the change of passenger flow in a short period of time. By the historical passenger flow correlation analysis of the previous day passenger flow sequence and the current moment of passenger flow are very strong correlation, easy to deal with the use of the previous 4 weeks of the same week the same moment of passenger flow data composed of the sequence as input, then the data in a sample input vector WX_d^h is:

$$WX(d,h) = [S(d-28,h),\cdots,S(d-7,h)] \tag{5}$$

Sample Label $wy(d,h)$ is:

$$wy(d,h) = S(d,h) \tag{6}$$

Combining the preordered hourly flow series, preordered day flow series and preordered weekly flow series to simultaneously satisfy the temporal distribution characteristics of passenger flow and the change of passenger flow within a short period of time can effectively make up for the shortcomings of a single passenger flow series. Therefore, the input vector $UX(d,h)$ of a sample in the data set is:

$$UX(d,h) = [HX(d,h), DX(d,h), WX(d,h)] \tag{7}$$

Sample Label $uy(d,h)$ is:

$$uy(d,h) = S(d,h) \tag{8}$$

For the construction of the LSTM dataset, not only the influence of the historical passenger flow sequence is taken into account, but also the extrinsic influences (weather data and holiday data). At this time, the constructed dataset adds more dimensions than the above dataset, and each passenger flow data in the 14×1 dimensional input vector composed of passenger flow sequences is added to the 56 dimensional extrinsic factors of the decision tree input vector X to compose a 14×57 dimensional input vector, and the input vector $X(d,h)$ of a sample in the dataset is:

$$X(d,h) = \left[[S(d,h-15),X],\cdots,[S(d-4,h),X],\cdots,[S(d-28,h),X]\right] \tag{9}$$

Sample Label $y(d,h)$ is:

$$y(d,h) = S(d,h) \tag{10}$$

2.2. LSTM model construction

The LSTM model is constructed based on the TensorFlow framework, and the network structure of the LSTM model is shown in Fig. 1. The basic structure of the LSTM model contains three layers of LSTM, three layers of Dropout and one layer of Dense [18]. The first LSTM layer also serves as the input layer, and the input layer is a three-dimensional variable: the first dimension is the sample size taking the value of None, the second dimension is the dimension of each sample taking the value of 14, and the third dimension is the dimension of each variable taking the value of 1. A sample in the input layer is a vector of 14 numbers, and the dimension of the input sequences is 14×1. The return sequences of the first two LSTM layers are set to True, and the Dense attribute is set to True, Dropout layer is set to 0.2, the number of discarded connections is 20% of the full connections, the activation function is set to ReLU, the optimizer is set to Adam, and the passenger flow prediction results are output through the Dense layer.

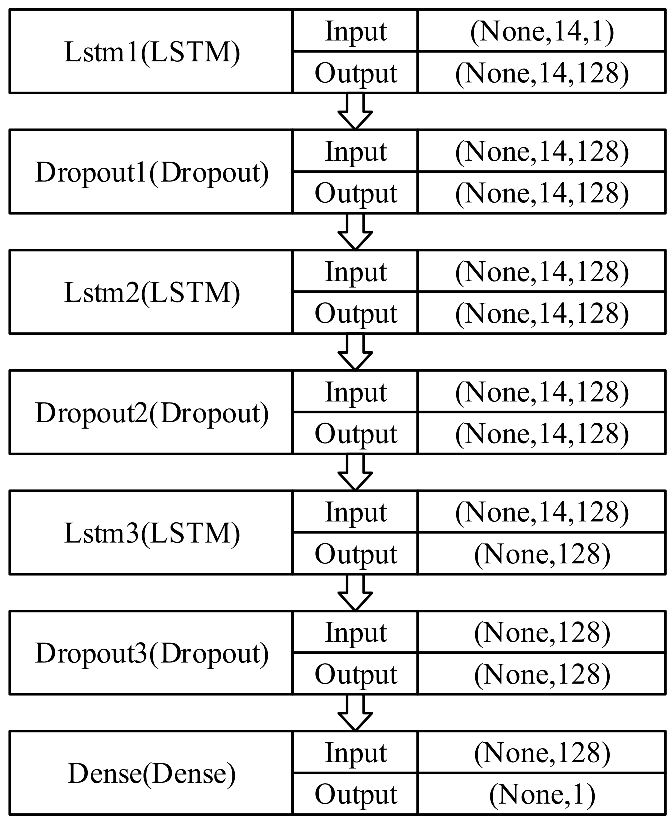


Fig. 1. Basic structure of LSTM model

3. Optimal electric bus frequency setting based on two-layer planning

Bi-level programming is a system optimization problem with a two-level recursive order structure, where the upper and lower models have their own objective functions and constraints [19]. The decisions of the upper model influence the decisions of the lower model, and the optimal solutions of the lower model are fed back to the upper model.

3.1. Lower level model

The electric bus frequency setting is performed on the given electric bus passenger flow data, and the electric bus passenger flow data predicted by the LSTM neural network above is applied to the lower model. The classical electric bus frequency setting model is shown below:

$$\begin{cases} min\sum_{a\in A} c_a V_a + \sum_{k\in I} \omega_i \\ \sum_{a\in\Lambda_i^+} v_a - \sum_{a\in\Lambda_i^-} v_a = g_i \\ \omega_i = \dfrac{V_i}{\sum_{a\in\Lambda_i^+} f_i x_a} \\ v_\alpha \le f_\alpha \omega_i, i\in I, \alpha\in A_i^+ \\ v_\alpha \ge 0, \alpha\in A \\ X_\alpha = 0 \text{ or } 1, \alpha\in A \end{cases} \tag{11}$$

Where: α, A is the arc segment and the set of arc segments, respectively. i,I are the nodes and the set of nodes, respectively. c_a,v_a are the travel time and passenger flow on arc segment α, respectively. f_a is the joint departure frequency on arc segment α. g_i is the passenger flow generated by node i. V_i is the passenger flow at node i. A_i^+, A_i^- are the set of routes leaving node i (attraction route set) and the set of routes entering node i, respectively.

3.2. Upper layer model

Before modeling, the following assumptions are first made:
1) Electric bus vehicles travel along a specified route during a specific time period.
2) Crossing stops and overtaking are not allowed on all electric bus routes.
3) Passengers queue up consciously and get on the buses first, without queue jumping.
4) There is only one way for passengers to travel, and passengers do not leave the electric bus routes due to long waiting times.

3.2.1. Parameter definitions. Departure frequency optimization is a complex process, where the supplier (electric bus company) wants to use the largest possible departure interval to reduce the variable cost, while the passengers require a higher departure frequency to reduce the waiting cost. Based on this idea, this paper establishes an electric bus departure frequency optimization model with the objective of minimizing the total cost of passengers and electric bus companies.

The variables used in the model are as follows: T is the time interval during the optimization period. f is the departure interval. π is the arrival frequency of passengers at station k, where passengers are considered to arrive uniformly during the optimization period. And let the arrival frequency be equal to the ratio of the number of boardings at each station during the optimization period (obtained from the lower model) to the optimization period (T). q_k is the passenger alighting ratio at station k, which is a percentage based on the number of passengers alighting at that station during the optimization period (obtained from the lower model) as a proportion of the number of passengers alighting at all stations between that station and the end point, $D_k t_k$ is the distance from station k to station $k+1$ and the average running time, respectively. $\overline{u},\overline{d}$ is the average passenger boarding and alighting time, respectively. V_m is the capacity of the mth vehicle. I_m^k, U_m^k, R_m^k are the number of occupants, boarding and remaining passengers of the m th vehicle at the k th station, respectively, where:

$$U_m^k = \begin{cases} r_k f + \sum_{n=1}^{\infty} r_k f (r_k u)^n + R_{m\text{-}1}^k V > r_k f + \sum_{n=1}^{\infty} r_k f (I_k u)^n + R_{m-1}^k + I_m^k (1-q_k) \\ V_m - I_m^k \ (1-q_k) \ \ other \end{cases} \tag{12}$$

$$R_m^k = \begin{cases} r_k f - (1-r_k u)U_m^k + R_{m\text{-}1}^k U_m^k < r_k f + \sum_{n=1}^{\infty} r_k f (r_k u)^n \\ 0 \ other \end{cases} \tag{13}$$

the stopping time of a vehicle at stop k, $S_m^k = \max(\bar{u}U_m^k, \bar{d}I_m^k q_k, S_{\min})$, ω is the passenger comfort coefficient, $1.5 > \omega > 1$ and here the indicator is expressed in terms of crowding: φ is the penalty coefficient, which increases with the number of vehicles missed, reflecting the additional cost to passengers who are unable to board the bus due to insufficient vehicle capacity, and can be expressed either as an integer greater than 1 or by multiplying the frequency of departures by the number of vehicles waiting. C_p^w, C_p^t, C_p^r are the passenger's wait time cost factor, ride time cost factor, and ride cost factor (fare), respectively. C_c^v, C_c^o are the electric bus vehicle cost factor and the electric bus company operating cost factor, respectively.

3.2.2. Total Passenger Costs. In addition to fares, the cost to passengers is primarily the cost of time spent waiting for and traveling on the bus; the cost of time spent transferring or walking is not considered here because these are not related to the frequency of departures.

1) Passenger's cost of waiting time (off-board costs):

$$M_o = \tau_w + \tau_\mu + \tau_\varphi \tag{14}$$

In the formula:

$$\tau_w(f, U_m^k) = C_p^w \sum_m \sum_k \sum_{n=1}^{U_m^k} (f + t + S_m^{k-1} - n / r_k) \tag{15}$$

$$\tau_\mu(U_m^k) = C_p^w \sum_m \sum_k \sum_{n=1}^{U_m^k} (n-1)\bar{u} \tag{16}$$

$$\tau_\varphi(f\, R_m^k) = C_p^w \varphi \sum_m \sum_k \left(\sum_{n=U_m^k+1}^{U_m^k+R_m^k} (f + S_m^{k\text{-}1} - (n-1)/r_k) + R_m^k f \right) \tag{17}$$

M_o is divided into three parts:

(1) Waiting for Vehicle Costs τ_w is the cost of time spent by passengers (excluding passengers left over from the previous trip) waiting on the platform before the vehicle arrives at the station.

(2) Waiting for Boarding Cost τ_μ is the cost of time spent by passengers waiting outside the vehicle to board the vehicle during the time the vehicle is stopped on the platform after the vehicle arrives.

(3) Extra cost τ_φ is the cost of time spent by left-behind passengers (passengers who are left on the platform due to maximum capacity constraints) waiting for the current vehicle and the previous vehicle and for the current vehicle to stop at that station. Here, a penalty factor φ is used to indicate the weight of adding the extra cost to the total cost.

2) Passenger's time cost of traveling (in-vehicle cost):

$$M_i = \tau_i + \tau_s \tag{18}$$

In the formula:

$$\tau_i = \omega C_p^t \sum_m \sum_k (I_m^k t + S_m^k I_m^k (1 - q_k)) \tag{19}$$

$$\tau_s = C_p' \sum_m \sum_k \left(\sum_{n=1}^{U_m^k} (S_m^k - (n-1)u) + \sum_{n=1}^{I_m^k \cdot q_k} dn \right) \tag{20}$$

M_i includes two components:

(1) Ride time cost τ_i, which includes the cost of the time that passengers in the car spend running

from the previous stop to that stop and the cost of the time spent stopping at that stop.

(2) Boarding and alighting time cost τ_s is the cost of the time passengers boarding (or alighting) at that station spend waiting on the bus.

3) Total Passenger Time Cost. The total time cost of a passenger is the sum of his off-vehicle time cost and on-vehicle time cost, viz:

$$Z_p = M_o + M_i \tag{21}$$

3.2.3. Costs of electric bus companies. When setting the frequency of departure on a route, an electric bus enterprise must reduce its own costs in addition to reducing the travel costs of passengers. This paper proposes a model of minimum enterprise cost, which mainly considers the revenue (passenger fare income) and operating costs (including the enterprise's resource allocation, personnel wages and fuel consumption and other comprehensive factors) of the electric bus enterprise, and this objective function is closer to the actual operating conditions.

1) Revenue of the enterprise M_{p}. Depends on the passenger flow it carries, which is equal to the total passenger flow multiplied by the fare:

$$M_p = C_p^r \sum_m \sum_k U_m^k \tag{22}$$

2) Business Operating Costs M_c. Includes costs related to depreciation and maintenance of vehicles, salaries of drivers and conductors, and fuel consumption:

$$M_c = C_c^v T / f + C_c^o \sum_k D_k T / f \tag{23}$$

3) The goal of a business Z_c. Is to minimize the difference between expenses and revenues:

$$Z_e = M_e - M_p \tag{24}$$

3.3. Integration model

Minimizing the total time cost of passengers and minimizing the cost of electric bus companies are somehow contradictory to each other, to minimize the cost of the whole system, it is necessary to find out the equilibrium of the two optimization problems, and here a kind of linear weighting method is used to consider the costs of these two aspects together:

$$\min: \ Z = \sum_1 (\lambda Z_p^1 + (1-\lambda) Z_e^1) \tag{25}$$

Where: 1 is the number of the electric bus line. λ is the weighting factor used to control the two objectives, $0 \le \lambda \le 1$.

3.4. Model solving

The flow of solving a two-layer planning model using a genetic algorithm is shown in Fig. 2, where the number of genetic generations is used as an algorithmic stopping criterion, and the other key modules are described as follows.

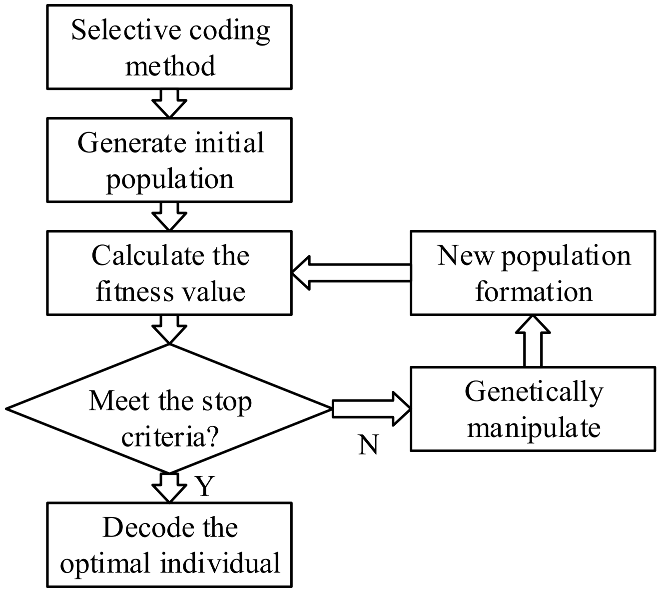


Fig. 2. Genetic algorithm solution process

Selection of coding method. Due to the large number of lines in the actual electric buses, the decision variable of the model, i.e., the frequency of electric bus line departure, is huge, so the real number coding is used to improve the search efficiency of the genetic algorithm, and the initial population is generated by using the $\{f_1, f_2, \cdots, f_n\}$-individual form, where n is the total number of lines. In order to make the randomly generated initial population are satisfied with the constraints, the first randomly generated line operating vehicles and then converted into the line frequency of the idea. The number of vehicles operating on the $i(i \in l)$ rd line is T_i. The specific operation is as follows: firstly, 1 integer is randomly selected from $[1, T-n+1]$, which is labeled as T_1, and is used as the number of vehicles on the 1st line. Then in $[1, T-T_1-n+2]$ randomly selected 1 integer, recorded as T_2, as the number of vehicles in the second line. In the same way, 1 integer is randomly selected from $[1, T-T_1-T_2-\cdots+T_2]-T_i-n+i](2<i\le n)$, which is recorded as T_i, as the number of vehicles of the i th line, until the number of vehicles of the last line is generated. After generating the number of vehicles for a line, the number of vehicles for each line is then converted into the frequency of departure. The initial value obtained from the above process is one individual, and the above steps are repeated to produce an initial feasible population that meets the required number of individuals.

Calculate the fitness value. The genetic algorithm in this paper uses the smallest fitness value as the search objective, so the layer objective value is used instead. It can be seen that before determining the fitness value it is necessary to match the flow of electric buses for each individual of departure frequency. If the generated individual does not satisfy the constraints then a penalty value needs to be added to the fitness value with a view to its elimination in the genetic operation.

Genetic manipulation. Based on the fitness value, genetic operation is performed to generate a population of good offspring. In this paper, the following operations are adopted sequentially: selection, replication, crossover, and mutation. Specifically, the individuals with the best fitness values in the current generation are replicated as superior offspring based on the elite principle. The individuals with better fitness values in the current generation are selected as the parents by roulette method, and then two-point crossover and Gaussian mutation operations are performed on the parents to generate crossover and mutation offspring. The above three sub-generations together form the next generation population.

4. Case studies

4.1. *Experimental data*

Passenger flow data are collected for inbound and outbound passengers at station 5 of electric bus route 363 in city Q. The daily inbound and outbound passenger flow data from 07:00 to 22:30 from August 5, 2023 to September 29, 2023 are shown in Fig. 3. The trend of inbound and outbound passenger flow at station 5 of electric bus route 363 is basically the same, with obvious volatility. Affected by the opening of the new subway line nearby, the inbound passenger flow is smaller than the outbound passenger flow. The minimum passenger flow occurs in early September and the maximum passenger flow occurs in early August. The passenger flow is divided into several intervals, and the LSTM model is used to predict the passenger flow in an interval with strong regularity.

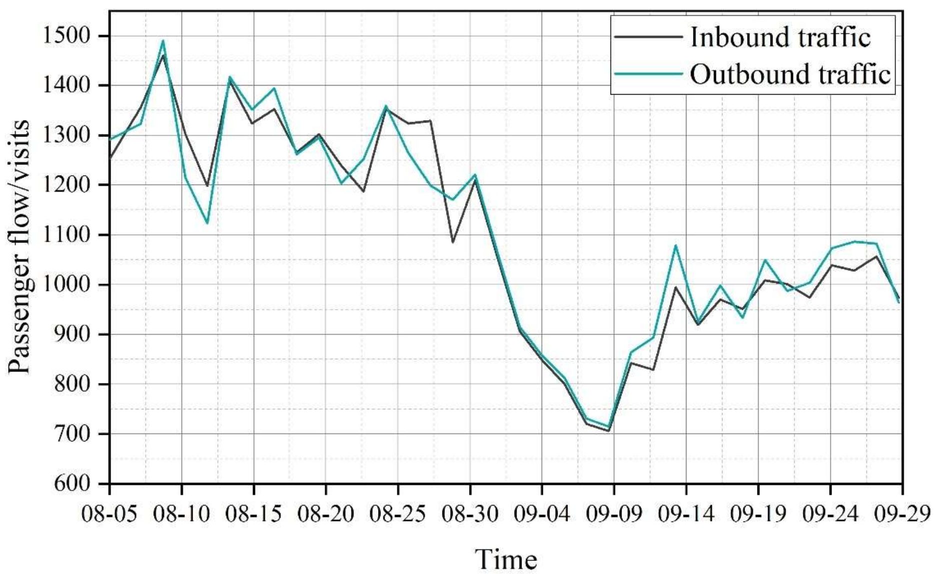
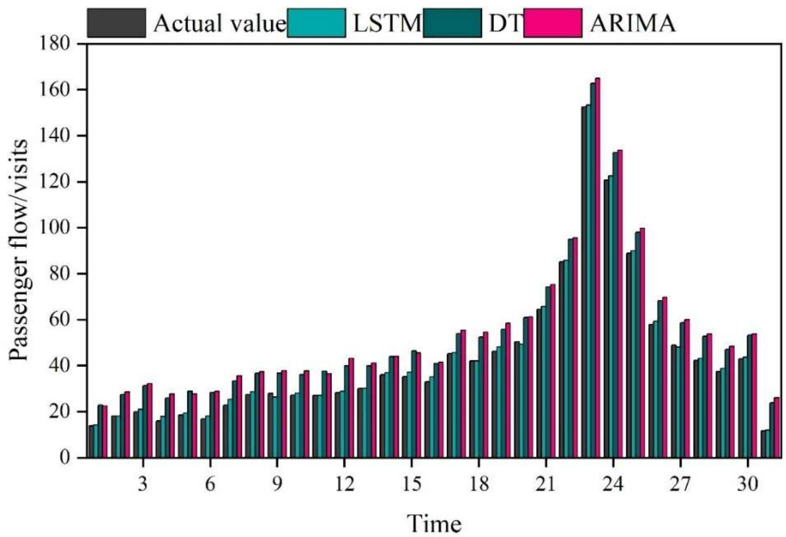


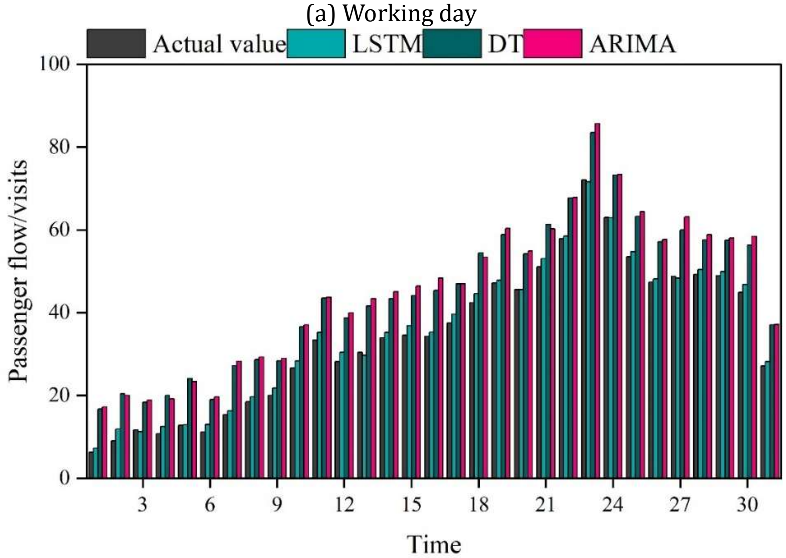
Fig. 3. Passenger flow data in and out

4.2. *Analysis of forecast effects*

4.2.1. Feasibility verification. In order to verify the feasibility of the prediction models, the autoregressive moving average (ARIMA) model, the decision tree (DT) model and the LSTM passenger flow prediction model were used to predict the passenger flow of the metro stations on weekdays and natural days, respectively. The inbound passenger flow of 2023-08-12-2023-08-13 is the training set for weekdays, the inbound passenger flow of 2023-08-27 is the test set for natural days, and the inbound passenger flow of 2023-08-27 is the test set for natural days.The prediction results of the inbound passenger flow of the ARIMA model, the DT model and the LSTM model are shown in Fig. 4. (a) and (b) represent the prediction results of inbound passenger flow on weekdays and natural days, respectively.The average errors of the LSTM model in the prediction of inbound passenger flow on weekdays and natural days are 0.8092 and 1.108, respectively.The deviation of the prediction results of the LSTM model from the actual inbound passenger flow is smaller compared with those of the ARIMA model and the DT model.



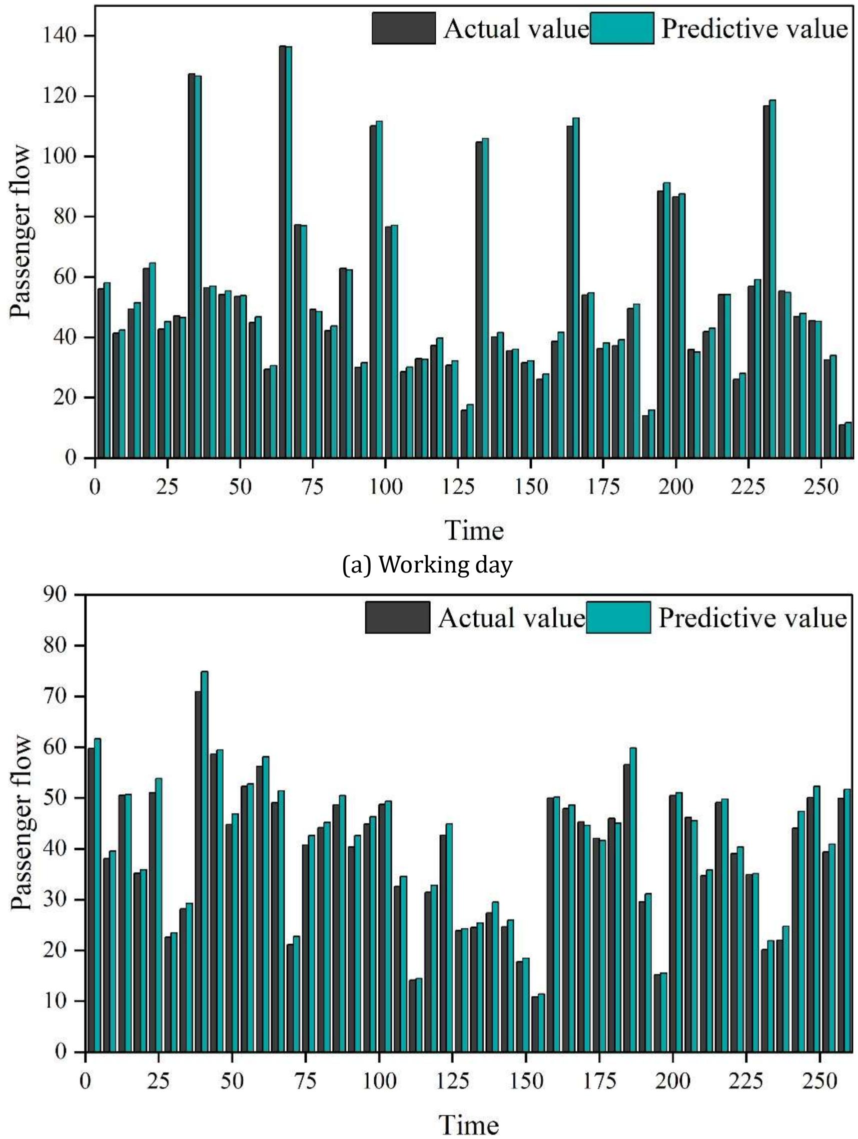
(a) Working day



(b) Natural day

Fig. 4. Inbound passenger traffic forecast results

4.2.2. Validation of applicability. Validate the applicability of LSTM passenger flow prediction model. The outbound passenger flows from 2023-08-07-2023-08-11 were used as the training set for weekdays, and the outbound passenger flows from 2023-08-14, 2023-08-21, 2023-08-28, 2023-09-04, 2023-09-11, 2023-09-18, 2023-09-25, 2023-09-28 outbound passenger volumes for the weekday test set. The outbound passenger flows of 2023-08-12 and 2023-08-13 are used as the training set for natural days, and the outbound passenger flows of 2023-08-19, 2023-08-26, 2023-09-02, 2023-09-09, 2023-09-16, 2023-09-23, and 2023-09-30 are used as the test set for natural days. The outbound passenger flow prediction results of the LSTM model are shown in Fig. 5. (a) and (b) represent the prediction results of inbound passenger flow on weekdays and natural days, respectively.The prediction results of the LSTM model basically match with the actual outbound passenger flow, the predicted passenger flow data are well fitted, the cycle pattern is obvious, both have 7 wave peaks, and the maximum daily passenger flow all appear at 08:30-09:00.



(a) Working day

(b) Natural day

Fig. 5. Departure rate forecast results

4.3. Analysis of the effect of frequency setting of optimal electric buses

4.3.1. Departure Timetable Solution Analysis. The predicted passenger flow of electric buses, passengers and the minimum total cost of electric bus companies based on LSTM neural network were substituted into a two-layer planning model, and the model was solved using genetic algorithm. The operating period of the 5th station of the electric bus of the 363rd road in Q city is divided into 930 time slots (1 time slot per minute), and the optimal frequency of 62 departures is solved accurately.

In order to show the optimal electric bus frequency method designed in this paper and the accelerated column generation algorithm [20] (IPC) of integer planning class respectively with the adaptive status of the capacity demand, the whole day operation period is divided into 34 segments according to every 30 min, calculate the capacity supply of station 5 from the beginning to the end of the period of operation, and count the passenger flow (i.e., the demand for capacity) of the fifth station every 30 min, the more the number of segments of the time period, the more balanced the demand and supply of capacity is. The higher the number of time segments, the more balanced the capacity supply and demand. The capacity supply and demand of station 5 is shown in Figure 6. The Pearson's correlation coefficient is used to measure the demand and supply of capacity, and when the correlation coefficient is in the range of 0.80~1.00, the result of the demand and supply fitness is a strong correlation. The correlation coefficient is 0.9635, which reflects the solution accuracy of the departure schedule, indicating that the capacity supply of electric buses fluctuates around the capacity demand, with small fluctuation amplitude and timely response of scheduling interval, which verifies the feasibility of the frequency setting method of electric buses.

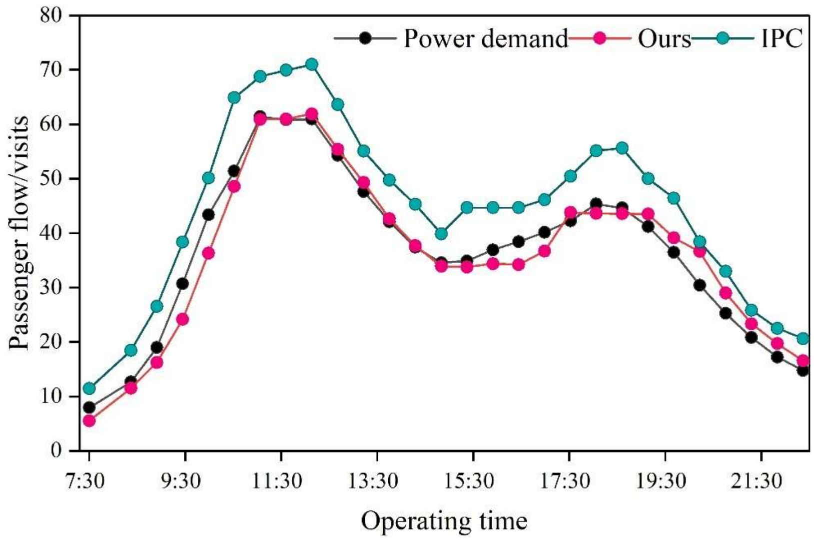


Fig. 6. Balance of transportation capacity supply and demand

4.3.2. Algorithm Stability Analysis. By using the program to calculate the numerical changes of four quantities such as the actual demand fleet size, the theoretical number of electric buses purchased under the uniform departure interval, the capacity supply, and the capacity demand in the process of changing the daily passenger flow from 300 to 3,000, the stability verification of the algorithm is shown in Fig. 7. With the increase of daily passenger flow, the actual capacity supply and actual capacity demand curves always maintain a basic coincidence, and the maximum deviation rate of the two is only 0.09%. The electric bus frequency setting model is adapted to the scenario of thousands of passenger flows, and the larger the single-day passenger flow, the higher the model accuracy. The actual number of electric buses purchased did not deviate from the trend of the theoretical minimum number of vehicles purchased, indicating that the genetic algorithm has good solution stability. The spacing between the actual number of vehicles purchased and the capacity demand is gradually shortened in Figure 7, which also proves that the larger the single-day passenger flow is, the higher the model accuracy is. This shows that the optimal setting method of electric bus frequency based on LSTM prediction and two-layer planning is suitable for the situation of large passenger flow and complex laws.

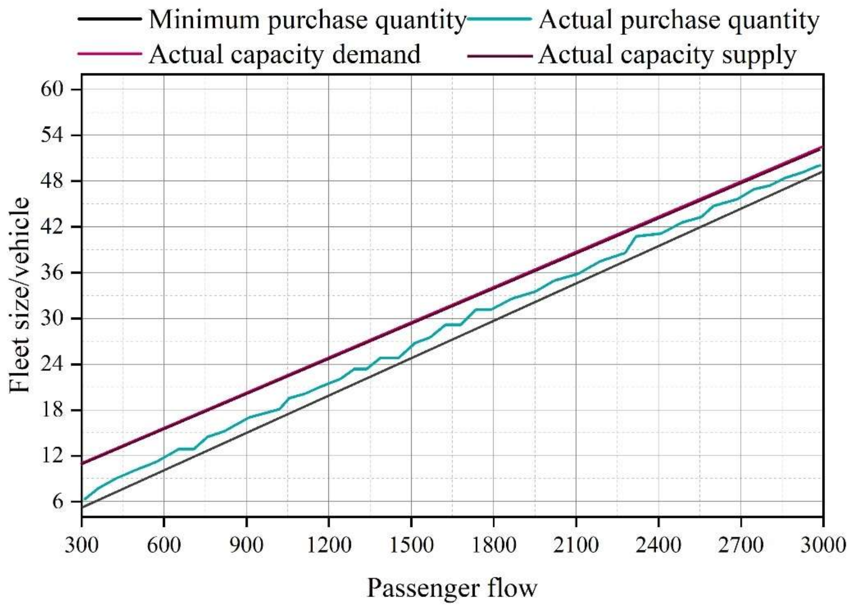


Fig. 7. Algorithm stability verification chart

5. Conclusion

In order to optimize the departure frequency of electric buses, this paper proposes a method combining LSTM prediction and two-layer planning. The experiment is based on the passenger flow data of electric bus No. 363 entering and leaving the station at the 5th station in Q city. The feasibility and applicability of the LSTM model are verified. On both weekdays and natural days, the model's prediction results of inbound and outbound passenger flow are not much different from the actual values. Use genetic algorithm to solve the optimal frequency of 62 departures at the 5th station of 363 electric buses. The Pearson correlation coefficient is used to measure the capacity supply and demand, and the correlation coefficient is obtained as 0.9635, which indicates the feasibility of the method in this paper. The maximum deviation of the actual capacity supply from the actual capacity demand curve is only 0.09% for the scenario of thousands of passengers. This shows that the method of designing optimal electric bus frequency in this paper has practical application value.

About the Author

Yixing Bao was born in Jiangxi in 2001. He obtained a Bachelor's degree from Tianjin University in China and a Master's degree from The University of Hong Kong. His main research direction is transportation optimization, industrial engineering and supply chain management.

References

[1] Muzaffar, S., & Afshari, A. (2019). Short-term load forecasts using LSTM networks. Energy Procedia, 158, 2922-2927.

[2] Zhao, Z., Chen, W., Wu, X., Chen, P. C., & Liu, J. (2017). LSTM network: a deep learning approach for short-term traffic forecast. IET intelligent transport systems, 11(2), 68-75.

[3] Shahid, F., Zameer, A., & Muneeb, M. (2021). A novel genetic LSTM model for wind power forecast. Energy, 223, 120069.

[4] Siami-Namini, S., Tavakoli, N., & Namin, A. S. (2018, December). A comparison of ARIMA and LSTM in forecasting time series. In 2018 17th IEEE international conference on machine learning and applications

(ICMLA) (pp. 1394-1401). Ieee.

[5] Merta, J., & Brandejský, T. (2022). Two-layer genetic programming. Neural Network World, 32(4), 215.

[6] Li, Y., & Hu, B. (2019). An iterative two-layer optimization charging and discharging trading scheme for electric vehicle using consortium blockchain. IEEE Transactions on Smart Grid, 11(3), 2627-2637.

[7] Xue, Z., Liu, Y., Han, G., Ayaz, F., Sheng, Z., & Wang, Y. (2021). Two-layer distributed content caching for infotainment applications in VANETs. IEEE Internet of Things Journal, 9(3), 1696-1711.

[8] Rahmatabadi, D., Aberoumand, M., Soltanmohammadi, K., Soleyman, E., Ghasemi, I., Baniassadi, M., ... & Baghani, M. (2022). A new strategy for achieving shape memory effects in 4D printed two-layer composite structures. Polymers, 14(24), 5446.

[9] Hong, J., Wang, H., Wu, Q., Hu, Y., Chen, X., Zhou, Y., ... & Zhao, J. (2023, August). Two-layer Programming of Compound Energy Storage System in New Power System. In 2023 IEEE 18th Conference on Industrial Electronics and Applications (ICIEA) (pp. 1770-1775). IEEE.

[10] Al Ali, Y., Hrairi, M., & Al Kattan, I. (2012). Potential for improving vehicle fuel efficiency and reducing the environmental pollution via fuel ionization. International Journal of Environmental Science and Technology, 9, 495-502.

[11] Ding, N., Prasad, K., & Lie, T. T. (2017). The electric vehicle: a review. International Journal of Electric and Hybrid Vehicles, 9(1), 49-66.

[12] Zhao, J., Xi, X. I., Na, Q. I., Wang, S., Kadry, S. N., & Kumar, P. M. (2021). The technological innovation of hybrid and plug-in electric vehicles for environment carbon pollution control. Environmental Impact Assessment Review, 86, 106506.

[13] Xing, J., Leard, B., & Li, S. (2021). What does an electric vehicle replace?. Journal of Environmental Economics and Management, 107, 102432.

[14] Crabtree, G. (2019). The coming electric vehicle transformation. Science, 366(6464), 422-424.

[15] Kumar, R. R., & Alok, K. (2020). Adoption of electric vehicle: A literature review and prospects for sustainability. Journal of Cleaner Production, 253, 119911.

[16] Jape, S. R., & Thosar, A. (2017). Comparison of electric motors for electric vehicle application. international Journal of Research in Engineering and Technology, 6(09), 12-17.

[17] Un-Noor, F., Padmanaban, S., Mihet-Popa, L., Mollah, M. N., & Hossain, E. (2017). A comprehensive study of key electric vehicle (EV) components, technologies, challenges, impacts, and future direction of development. Energies, 10(8), 1217.

[18] Rong Zhu,Jie Tang,Xuansen He,Xianlai Zhou,Xiaohui Huang,Fengyun Wu & Songli Chen. (2024). SE-MAConvLSTM: A deep learning framework for short-term traffic flow prediction combining Squeeze-and-Excitation Network and Multi-Attention Convolutional LSTM Network. PloS one(12),e0312601.

[19] Liu Xiaoyu,Dabiri Azita,Xun Jing & De Schutter Bart. (2023). Bi-level model predictive control for metro networks: Integration of timetables, passenger flows, and train speed profiles. Transportation Research Part E.

[20] Shengdong Li,Dajie Zuo,Wenqing Li,Yongxiang Zhang & Li Shi. (2024). Freight train line planning for large-scale high-speed rail network: An integer Benders decomposition-based branch-and-cut algorithm. Transportation Research Part E103750-103750.