IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received July 4, 2019, accepted July 30, 2019, date of publication August 2, 2019, date of current version August 15, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2932801

A Novel Bus-Dispatching Model Based on Passenger Flow and Arrival Time Prediction

ZHAO HUANG[1,2], QINGQUAN LI[1], FAN LI[1,3], AND JIZHE XIA[1]
[1]Shenzhen Key Laboratory of Spatial Smart Sensing and Services, Shenzhen University, Shenzhen 518060, China
[2]College of Information Engineering, Shenzhen University, Shenzhen 518060, China
[3]College of Computer Science and Software Engineering, Shenzhen University, Shenzhen 518060, China

Corresponding author: Jizhe Xia (xiajizhe@szu.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 41701444, and in part by the Shenzhen Scientific Research and Development Funding Program under Grant Number JCYJ20170818101704025.

ABSTRACT Public transport is vital to people's daily travel, and bus dispatching plays a significant role in the public transport system. With deep learning having been widely applied and achieved great success in many fields, bus dispatching methods based on deep learning are proposed in succession. Currently, many bus dispatching models assume that the bus departure timetable is fixed and optimize the bus departure timetable interval according to passenger flow. However, the bus departure timetable is variable in general, only considering that the bus arrival time is insufficient. Targeting the above challenges, we propose a novel dynamic bus dispatching model based on arrival time and passenger flow prediction (D-ATPF). First, the historical origin–destination (OD) data and the transfer data are obtained by processing the bus trajectory data and the passenger card-swiping records, and the bus arrival time is extracted by analyzing the GPS trajectory. Second, the components of bus arrival time and passenger flow prediction based on long short-term memory (P-LSTM) are adopted to predict the future passenger flow and bus arrival time. Finally, the genetic algorithm-based bus dispatching model (GABD model) searches the minimum waiting time for passengers by using stay strategy. By using data of five lines with 124 bus stations and a total of 9 02 509 records in Guangzhou city, China, our experimental results show that: 1) the average mean absolute percentage error ($\overline{MAPE}$) and root mean square error ($\overline{RMSE}$) of passenger prediction are 14% and 7.5, respectively; 2) the average $\overline{MAPE}$ and $\overline{RMSE}$ of bus arrival time are 7.5% and 13.5, respectively; 3) regarding the passenger flow and arrival time prediction, the proposed D-ATPF model reduced waiting time by 829.68 min, accounting for 25.19% of the total waiting time; and 4) compared with the real-time stay strategy, the reduced waiting time of this method increased by 5.94%. Therefore, the D-ATPF model provided a more practical model for buses dispatching.

INDEX TERMS Bus dispatching, LSTM, passenger flow prediction, arrive time prediction, genetic algorithm.

I. INTRODUCTION

Public transport is one of the most important means of transportation for people to travel in modern cities. To improve road safety, a strong pseudonym-based authentication (SPATA) framework is used to preserve the real identity of vehicles [1] and a novel two-layer vehicle type classification framework based on the vehicle's 3D parameters and its local features is provided to control traffic and road code violations [2]. Compared with private cars and taxis, public transport has the advantages of low fare and large passenger capacity. Meanwhile, traveling by public transport can effectively reduce road traffic flow, alleviate traffic congestion and air pollution. However, occasionally unbalanced passenger distribution and small passenger demand have historically caused low attendance rate and significant financial losses of bus companies in some towns. Targeting these issues, a variety of dynamic bus scheduling methods have been proposed [34]. However, some deficiencies still exist. First, most of the dynamic bus dispatching models are based on fixed and predictable departure timetables, which is not the case in all areas. Second, some research only addressed the

The associate editor coordinating the review of this manuscript and approving it for publication was Martin Gonzalez-Rodriguez.

frequency of bus departures according to predicted passenger flow regardless of the time of bus arrivals, which caused waste of traffic resources, increased the losses of public transport companies, and aggravated traffic congestion.

In this paper, we propose a novel bus dispatching optimization model based on arrival time and passenger flow prediction (D-ATPF model), which takes the passenger flow and bus arrival times into account to adopt a "stay" strategy at the transfer station. Stay strategy is a widely used dynamic bus dispatching strategy, which incorporates travel intervals and vehicle timetable [5], [7]–[11]. The D-ATPF model: a) extracts card-swiping records from smart IC cards, captures the bus-travel trajectory from smart terminal devices installed on buses, and obtains bus station and line-vector data; b) utilizes the methods of station matching applied to the destination and transfer station speculation to infer the passenger flow and bus arrival time from the extracted data; c) adopts a Long Short-Term Memory (LSTM) neural network to predict the bus arrival time and passenger flow in the future (P-LSTM), which outperforms SVR, DBN-SVR and ARIMA approach in predicting events with relatively long intervals and delays in time series, such as passenger flow and arrival time prediction [12]; d) establishes the objective function of minimizing passenger waiting time based on the staying strategy of travel intervals by treating stay time as a decision variable, and seeks the optimal solution of the objective function using a genetic algorithm (GABD).

The main contributions of this study include two aspects: (1) a novel optimized dynamic bus dispatching model, which combines passenger flow and bus arrival-time prediction models based on LSTM; (2) an objective function is established to minimize passenger waiting time, the optimal solution of this nonlinear objective function is determined by using a genetic algorithm.

The remaining structure of this paper is as follows. Section 2 reviews related work regarding bus dispatching. Section 3 describes the framework, the P-LSTM model, and the GABD model. Results of the experiments are provided in Section 4. Conclusions are shown in Section 5.

II. LITERATURE REVIEW

A. PASSENGER FLOW AND ARRIVAL TIME PREDICTION

Currently, models to predict passenger flow and bus arrival time are categorized into traditional, regressive, machine learning-based, and Hybrid models.

1) TRADITIONAL MODELS

Zhang et al. searched highly correlated passenger flow patterns in the historical estimations and predicted the future bus passenger flow using an Extended Kalman Filter (EKF) [13]. Moreira-Matias et al. integrated the information of streaming data into a histogram time series and predicted the spatial distribution of taxi-passengers for a short-term time horizon with a three time-series forecasting methodology [14]. The RTIS algorithm for estimating the current and instantaneous travel times using automatic vehicle identification (AVI) data was proposed by Tam *et al.* [15].

2) REGRESSIVE MODELS

Chen *et al.* [16] analyzed a time series of historical data to predict the passenger flow of the Line 16 Lingang Avenue Station using an ARIMA model, based on judging and identifying the parameters with full or partial autocorrelation. Ma *et al.* [17] proposed a short-term bus passenger demand prediction method using an interactive multiple model-based pattern hybrid (IMMPH). Suwardo *et al.* [18] proposed an Autoregressive Integrated Moving Average (ARIMA) to predict bus travel time using travel time series data.

3) MACHINE LEARNING-BASED MODELS

Yang *et al.* [19] proposed a prediction model of bus arrival times based on a Support Vector Machine with a genetic algorithm (GA-SVM) in Shenyang. A passenger-flow prediction model based on a spatial weighted least-squares support vector machine (SW+LS+SVM) was used to predict passengers of urban rail transit stations by Zhou and Zhang [20]. A short-term railway passenger demand forecasting model based on multiple temporal unit neural network (MTUNN) and parallel ensemble neural network (PENN) was proposed by Tsai *et al.* [21].

4) HYBRID MODELS

Ding et al. provided an ARIMA-SVM model combining the advantages of Regressive models with Machine learning-based models to effectively predict the bus dwell time [22]. A model of mixed support vector machine (SVM) and Kalman filtering is exploited to predict bus arrival times in Dalian by Yu *et al.* [23]. Luo et al. proposed a spatiotemporal traffic flow prediction method, which combined with KNN and LSTM [24]. Zhang et al. provided the PSR-LSTM model by the phase space reconstruction method to recover the hidden trajectory in the passenger flow [25]. Petersen et al. present bus travel time prediction system by using a combination of convolutional and long short-term memory (LSTM) layers, which leverages the non-static spatio-temporal correlations of urban bus networks [26].

B. BUS DISPATCHING

The unbalanced spatial-temporal distribution of buses is a pervasive problem. Considering the spatial dynamic of travel demand, the spatial-temporal optimization framework modeling the interactions of vehicles and demands have been developed [27], [28]. In terms of bus, Luo et al. provided an optimization model for dynamic bus dispatching to minimize the overall waiting time of passengers in a transit system by considering multiple types of real-time information [12]. A bus dispatching model that accounts for passenger occupancy rate and the profit of public transport companies to minimize costs has been developed by Li *et al.* [29]. Pang et al. designed a scheduling system of intelligent public transport with mobile internet technology to relieve unbalanced spatial and temporal distributions in urban, suburban, and rural areas [30]. A prediction model based on bus arrival times with support vector machines (SVM) has been proposed to serve bus stay dispatching by Yu *et al.* [7]. S. J. Berrebi et al. proposed a real-time holding mechanism to minimize passenger waiting time while maintaining the highest possible frequency on a loop-shaped route [8]. Zhang and Liu proposed an adaptive fleet size adjustment mechanism, which adopted the doubly dynamical system to adjust the size of the dispatched bus fleet and accommodate day-to-day variations of mode choices and traffic patterns [31]. Forbes et al. introduced a software system to dispatch buses of different types and use the simulated annealing algorithm to improve the desirability of the total allocation [32]. Strathman et al. designed a framework for an assignment focuses on documenting service reliability and passenger activity at pre-operational (baseline), initial and full implementation period [33]. An and Zhang provided a mixed integer programming model to improve transit service with a minimum cost by using the Lagrangian relaxation algorithm, which integrated a bus-holding and stop-skipping strategy [9]. Du et al. provided a mathematical model to maximize the average satisfaction of passengers and loading rate and minimize the average bus departure frequency [34]. Ting and Schonfeld developed a bus-stay strategy in transfer stations and used the initiating algorithm to seek the optimal solution of passenger wait times [35]. The change of passenger wait times in different situations was explored using specific examples.

In summary, two kinds of bus control decisions are applied in recent research work, i.e., control theory and optimization. The three main limitations of these researches work include the following: (1) most studies only considered passenger flow in the bus station and ignored the real-time information such as traffic congestion, weather factors, etc.; (2) some studies only considered simple single lines, avoiding the transfer problem of multiple lines; (3) some studies only increased departure frequency and ignored the prediction of passenger flow and the arrival time of buses. Targeting these issues, this paper introduced a dynamic bus dispatching model considering both passenger flow and bus arrival time.

III. THE BD-PFATP MODEL

A. FRAMEWORK

The purpose of this research is to optimize bus dispatching strategies and minimize the total waiting time of passengers using the D-ATPF model. The framework of D-ATPF is depicted in Fig. 1. The bus station, bus trajectory, and IC-card records are captured to match bus stations and infer alighting and transfer stations. The first component is the P-LSTM model, which consists of passenger flow and arrival time prediction. The passenger flows of the original, terminal, and transfer stations are predicted accounting for varying loads on holidays and different days of the week. The historical passenger flow features and bus arrival times are estimated using the historical arrival time and state information. LSTM is adopted to complete passenger-flow forecasting and bus arrival time prediction tasks, which performs well in dealing with sequential data.

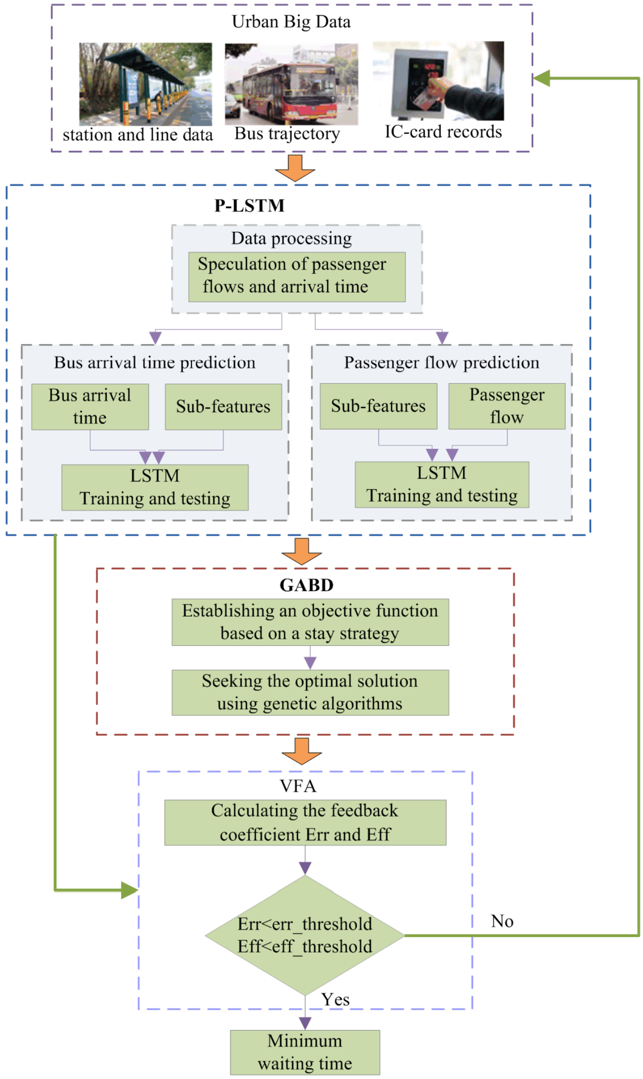


FIGURE 1. The framework of the D-ATPF model.

The second component is the GABD model, which includes the establishment of an objective function and seeking the optimum using a genetic algorithm [36]–[38]. The P-LSTM model can predict the average arrival rate of each station in the future, and the time interval of adjacent buses arriving at the same station. Therefore, an objective function is established with the stay time as the decision variable to minimize the passenger waiting time, and the optimal solution of the objective function is determined using a genetic algorithm, which can effectively deal with complex nonlinear functions.

The third component is the validation, feedback and adjustment model (VFA model), which calculates the feedback coefficient according to the P-LSTM prediction error and the GABD model convergence efficiency. The values of feedback coefficients are used to judge the necessity of data updates and LSTM model retraining and adjust the probability of crossover and variation.

B. P-LSTM MODEL

The structure of the P-LSTM model has two stages, as shown in Fig. 2. The first stage is the speculation of bus arrival times and passenger flows at stations, which includes inferring the number of boarding, alighting, and transferring passengers along with the bus arrival interval. The second stage is the prediction of passenger flow [39]–[43] and arrival times [44]–[47], which predicts the arrival time, boarding, alighting and transfer passengers using a Multilayer LSTM.

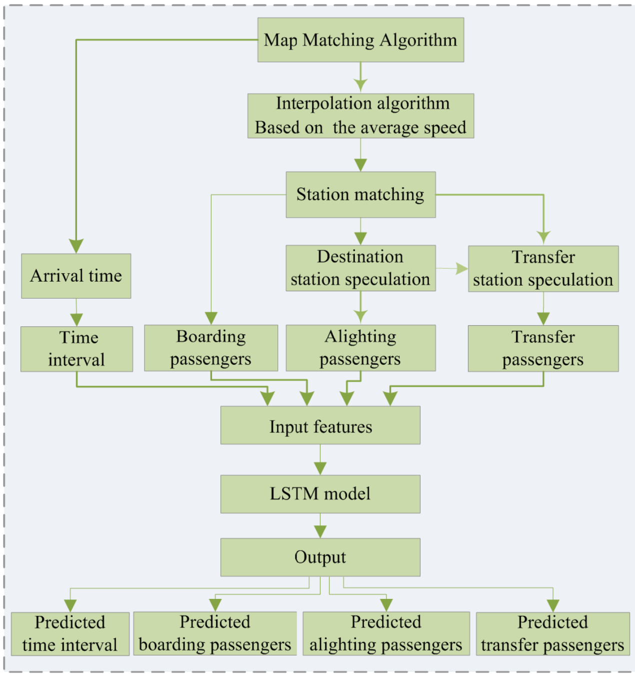


FIGURE 2. **The framework of P-LSTM model.**

1) BUS ARRIVAL TIME AND PASSENGER FLOW SPECULATION

The speculation of passenger flows consists of three components: Station matching, Inference of the Destination Station, and Inference of the Transfer Station.

1) Station matching. A map-matching method for low-frequency floating buses is adopted to restore the space-time trajectory of buses [48], and an average speed interpolation algorithm is employed to interpolate the bus trajectories uniformly in space every second while matching passenger instantaneous card-swiping positions to the trajectory in chronological order. Then, the nearest station to the passenger's instantaneous card-swiping position is regarded as a boarding station using a nearest-neighbor strategy.
2) Destination station speculation. The existing method cannot infer the passengers' destination station without continuous bus trip chains. Therefore, traveling in a continuous bus-trip chain is assumed and the last destination of the previous day is estimated to be the same as the first station of the next day. A continuous bus-trip chain indicates that we can only go out and return by bus, other means of transportation are not allowed. Therefore, the card swiping times is more than once. There are two scenarios of destination station speculation in continuous bus trip chains [49]. First, if the passenger swipes their card on the same line twice continuously, the destination of the first trip is the same as the origin station of the second travel. Second, if the line of continuous twice swipes is different bus lines, according to the nearest neighbor rule, the station on the first swipe line has the nearest distance to the second card swipe station, which is the destination station.
3) Transfer station speculation. OD data is arranged in chronological order before transfer station speculation. if the time interval between first alighting and second boarding is less than 30 min, the second trip is defined as a transfer [50].

2) PASSENGER FLOW AND BUS ARRIVAL TIME PREDICTION

The architecture of the LSTM used in this research is shown in Fig. 3.

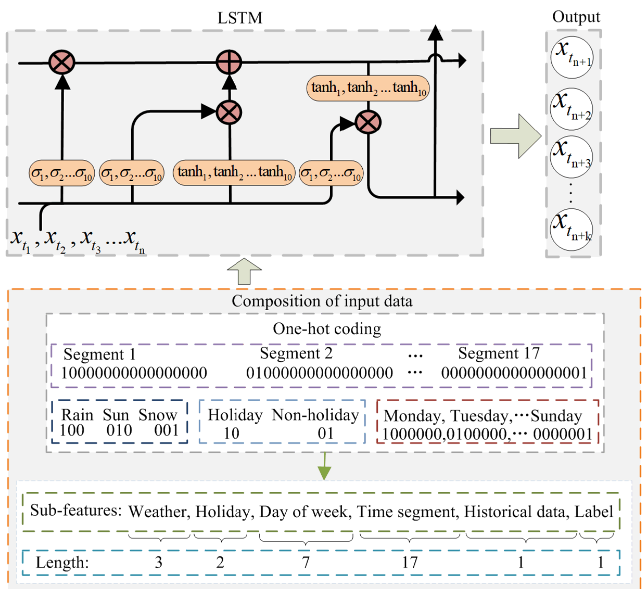


FIGURE 3. **The physical architecture of LSTM.**

The time is divided into 17 segments on average from 6:00 am to 11:00 pm. Passenger flows for each segment is counted as unique features. Meanwhile, the day of the week, holiday occurrence, and weather of each time segment are regarded as common features. The number of passengers at the next segment is taken as the label of this segment. Common and unique features and labels are integrated into an array, which is imported to the input layer, and the predicted results are output from a three-layer LSTM. Each row of features in the input data is converted into 1×30 dimensions after one hot coding. Therefore, the input data dimension is n × 30, and the hidden layer contains ten neurons. Finally, k × 1 dimensional output data represents the predicted future k results.

The prediction of the boarding, alighting and transfer passenger are described in detail as follows.

1) Prediction of boarding passenger quantity. The number of boarding passengers at different stations in each time segment can be obtained more accurately by the station matching method in the previous section, which is taken as the unique feature of prediction, and the number of boarding passengers at the next segment is regarded as the label of this segment. The common features, unique features, and labels are combined into input data, which is input to the prediction model. If the value of the loss function is lower than a threshold after enough training using a three-layer LSTM neural network, the training stops and the prediction results are output from the output layer.
2) Prediction of alighting passenger quantity. The number of alighting passengers can be captured by destination station speculation, which is taken as the unique feature of prediction. Similarly, this label is the number of alighting passengers at the next time segment. Prediction results can be obtained by inputting common and unique features and labels into the prediction model.
3) Prediction of transfer passenger quantity. The method of transfer station speculation can effectively infer the number of boarding passengers combined with the time interval between the next boarding and the last boarding, and the label is also the number of transfer passengers at the next time segment. The future number of transfer passengers is predicted by the model by inputting the integrated features consisting of common features and the historical number of transfer passengers.
4) Prediction of bus arrival time. Bus arrival and departure time are extracted using a map-matching algorithm to process GPS trajectory and station information. The time interval between arrival and departure times are captured to predict the future time interval of buses arriving at the same stations separately. First, we translate the time of arrival and departure into seconds and arrange them in chronological order within each day. Second, the time interval between the departure time of the last bus k and the arrival time of the current bus m can be calculated, denoted as $T_{\mathrm{k}\to m}$, as shown in Fig. 4.

The time intervals of n buses at station A are defined as $T_{k_1\to m_1}$, $T_{k_2\to m_2}$, $T_{k_3\to m_3}\ldots T_{k_n\to m_n}$ respectively. Similarly, the time intervals at station B are defined as $T_{j_1\to i_1}$, $T_{j_2\to i_2}$, $T_{j_3\to i_3}\ldots T_{j_n\to i_n}$. The latter time intervals are regarded as the labels of the former features, and the historical time interval as a unique feature. The common features, unique features, and labels are input into the prediction model. Future arrival time is obtained after training and prediction.

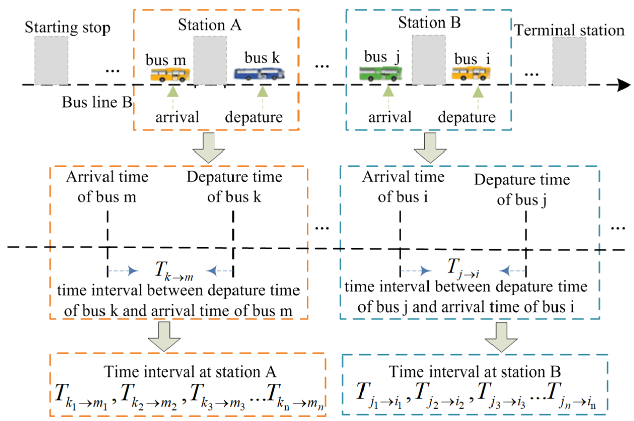


FIGURE 4. The time intervals at different stations.

C. GABD MODEL

The aim of the GABD model is to establish an objective function to minimize the waiting time of passengers based on a stay strategy at the transfer station, and to seek the optimal solution of the objective function using a genetic algorithm. The construction algorithm of GABD is shown in algorithm I.

This algorithm mainly includes establishing the objective function and using the genetic algorithm to find the optimal solution of the objective function.

1) OBJECTIVE FUNCTION

In this research, the stay strategy is applied to multi-line bus dispatching, and Table 1 shows the variables used in the objective function.

TABLE 1. The definition of variables.

Variable	Definition
n^L	the label of the bus line
Δ, Δ'	The direction of bus running
m	bus label
k	the label of the bus station
$R_{k,n^L,\Delta,i}$	passenger arrival rate at the station k in Δ direction on line n^L within the time segment i
$N^{transfering}_{k,n^L,\Delta,m}$	number of passengers transferring to bus m at the stop k
$T^{holding}_{k,n^L,\Delta,m}$	stay time of the bus m at the station k in Δ
$T'_{m+1\to m,n^L,\mathrm{k},\Delta}$	the time interval between bus m and m+1
$T^{u}_{k',n^L,\Delta,\mathrm{m}}$	boarding time of bus m at station k in Δ direction
$T^{d}_{k',n^L,\Delta,\mathrm{m}}$	alighting time of bus m at station k in Δ direction.

Based on the above notations, the variation of passenger waiting time can be divided into the following four parts.

1) Reduced time at transfer stations. Some passengers caught up with the bus m and didn't have to wait for the

Algorithm 1 Construction Algorithm of Genetic Algorithm-Based Bus Dispatching

```
Input: staying time set: T= {t1, t2, t3...tn}.
Output: reduced_maximum_time
function Build_GABD (T)
  create Tmax = f(t1, t2, t3, ...t50)
  population_size=50, generations=1000,
  chrom_length=500, pc=0.6,
  pm=0.01, genetic_population = NULL, population =
  NULL
  for each i in population _size do
    for each bus_code in
    length_of_bus_staying_time_set do
      add randint to genetic_population
    end for
  end for
  for each step in generations do
    value=genetic_population*(2**(length−1 − k))
    dec= tmin+value*(tmax − tmin)/
    (2** (chrom_length) − 1)
    for each i in population_size do
      if reduced _time> reduced_ maximum _time then
        reduced_ maximum _time = reduced _time
      end if
      if probability_c<pc then
        crossover
      end if
      if probability_m< pm then
        mutation
      end if
    end for
  end for
  return reduced_ maximum _time
end function
```

bus m + 1 at the transfer station due to bus detention. Therefore, they saved waiting time, which is denoted as S_1, which is calculated by the following formula.

$$S_1 = N_{m,n,\Delta}^{hu+} \times T'_{m+1\rightarrow m,n,\Delta} \tag{1}$$

and the number of these passengers $N_{m,n,\Delta}^{hu+}$ is defined as:

$$N_{m,n,\Delta}^{hu+} = T_{k,n^L,\Delta,m}^{holding} \times R_{k,n^L,\Delta,i} + \sum_{n',\Delta'} \sum_{i} N_{i,n^L,\Delta'\rightarrow n^L,\Delta'}^{transfering} \tag{2}$$

The average arrival rate of the station k within time fragment i is defined as:

$$R_{k,n^L,\Delta,i} = N_{k,n^L,\Delta,i} \times \frac{1}{T} \tag{3}$$

2) Reduced waiting time for the downstream station. Similarly, some passengers took the bus m and didn't have to wait for the bus m+1. Therefore, the saved time S_2 is shown by the following calculation equation.

$$S_2 = \sum_{k'} R_{k',n,\Delta,i} \times T_{k,n^L,\Delta,m}^{holding} \times T''_{k',m+1\rightarrow m,n,\Delta} \tag{4}$$

and T'' indicates the driving time interval between bus m+1 and bus m, the calculation equation is as follows,

$$T''_{k',m+1\rightarrow m,n,\Delta} = T'_{k',m+1\rightarrow m,n,\Delta} - H(t) - T_{k,n^L,\Delta,m}^{holding} \tag{5}$$

$$H(t) = \left(R_{k',n^L,\Delta,i} + R_{k,n^L,\Delta,i}\right) \times \delta \times T_{k,n^L,\Delta,m}^{holding} \times \bar{u} \tag{6}$$

Assume that only when the alighting time is longer than the boarding time, the subsequent buses will not be affected. Therefore, we define

$$\delta = \begin{cases} 1 & T_{k',n,\Delta}^{u} \geq T_{k',n,\Delta}^{d} \\ 0 & T_{k',n,\Delta}^{u} \geq T_{k',n,\Delta}^{d} \end{cases} \tag{7}$$

the extra stay time can be calculated by the formula below.

$$h(t) = R_{k',n^L,\Delta,i} \times T_{k,n^L,\Delta,m}^{holding} \times \bar{u} \tag{8}$$

3) Increased waiting time at transfer stations. Some passengers did not get off at transfer stations. Therefore, their waiting time increased due to the stay strategy of bus m at the transfer station, which is denoted as A_1 the calculating equation is as follows.

$$A_1 = N_{k,n,\Delta,m-1\rightarrow m} \times T_{k,n^L,\Delta,m}^{holding} \tag{9}$$

In this formula, $N_{k,n,\Delta,m-1\rightarrow m}$ denotes the number of passengers did not get off on the bus m at the transfer station k.

4) Increased waiting time for the downstream station. Similarly, some passengers had to wait for bus m due to miss bus m − 1 at the downstream station k''. Therefore, their increased waiting time A_2 is calculated as follows.

$$A_2 = \sum_{\mathbf{k}} R_{k,n^2,\Delta,i} \times T'_{k',m\rightarrow m-1,n^2,\Delta} \times T_{k,n^2,\Delta,m}^{hoding} \tag{10}$$

Therefore, the objective function for minimizing passenger waiting time is established as follows:

$$y_{\max}\left(T_{k,n^L,\Delta,m}^{holding}\right) = S_1 + S_2 - A_1 - A_2 \tag{11}$$

To avoid congestions at the bus station, we give the following constraints:

$$0 \leq T_{k,n^L,\Delta,m}^{holding} \leq 120\ (\mathrm{s}) \tag{12}$$

2) GENETIC ALGORITHM

Genetic algorithms are computational optimization models by simulating the natural evolution process, which continuously searches and updates the optimal solution after repeated iterations and calculating the value of the fitness function. More optimal configurations survive to the next "generation". This algorithm is effective in solving nonlinear multivariate, and the optimal solution is arrived in short times, which is adopted in path planning problems, and bus scheduling problems [51]–[54]. In this research, a genetic algorithm is adopted to seek the minimum waiting time. The genetic algorithm workflow is shown in Fig. 5. A "gene" represents a bus staying time at the transfer station, and the staying time is encoded to binary codes with length 10.

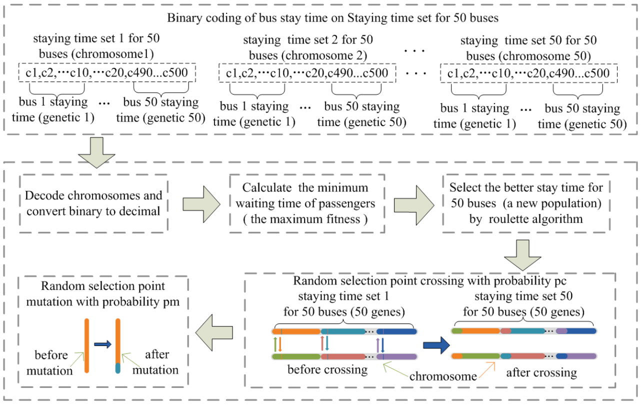


FIGURE 5. The genetic algorithm workflow.

There are three steps as follows.

1) Coding strategy. The stay time is regarded as the decision variable of dynamic dispatching. Suppose that the variable t'' indicates the stay time of the bus m at the transfer station, which must satisfy the constraints of equation (12), and n indicates the number of buses. Then, the stay time is encoded with a decimal coding, as follows:

$$[t'_1, t'_2, t'_3 \ldots t'_m \ldots t'_n] \tag{13}$$

2) Survival of the fittest. In this paper, a roulette method is adopted to select a suitable parental chromosome for breeding the next generation from the group. To ensure that the best gene can be saved to the next generation and avoid the degradation of the performance of the genetic algorithm, this paper adopts a strategy selecting the best parental chromosome (the minimum total waiting time of passengers), and automatically copying the fittest to the next generation.
3) Convergence condition. The simple single point crossover and simple Mutation methods are used for genetic operations. Crossover indicates interchanging genes (bus staying time) with probabilistic pc and remains elite individuals (best staying time). Mutation represents mutating individual genes (bus staying time) with probability pm, and obtain the best staying time. If the obtained result satisfies the convergence condition or reaches the maximum number of generations, or reaches the number of preselected settings, it will terminate the iterations. Otherwise, it will return to the survival of the fittest and continue to search.

D. VFA MODEL

This model is used to adjust the prediction effect and convergence efficiency. The evaluation criteria of this model include two parameters: Err and Eff. The Err is the error of prediction and the Eff is the maximum deviation of each of the 100 consecutive best values from the previous one, and the computational equation of deviation are shown in (14).

$$D = \max \sum_{z=1}^{s} \frac{|v_{best,z-1} - v_z|}{v_{best,z-1}} \tag{14}$$

$v_{best,z-1}$ indicates the $z-1^{th}$ best value, v_z represents the z^{th} best value. Meanwhile, the thresholds of Err and Eff can be defined by the user and be adjusted according to different requirements. In this paper, we set thresholds of Err and Eff to 0.2 and 0.05 according to the long-term transportation operation experience of the Guangzhou transportation department.

$$Err < 0.2 \tag{15}$$

$$Eff < 0.05 \tag{16}$$

IV. EXPERIMENT AND RESULTS

A D-ATPF model has been developed to conduct the experiments in this research, in which the bus station and line vector data, bus trajectory data, and smart IC-card transaction records data of Guangzhou city is used.

A. DATA DESCRIPTION

The dataset used in this research covered 27 days from April 24 to May 20, 2018 in Guangzhou city, China, which contains three vacation days and six weekend days. Yile Village Station is selected as the transfer station to adopt the stay strategy, which is the intersection station of the five lines. The reason why choose the five lines is that these five lines have the characteristics: large passenger flow, great changes during peak and off-peak period, passing through schools, crossing each other and more attention from the relevant government departments. The bus station and line-vector data, bus trajectory data, and smart IC-card transaction record data are included. More details about the dataset are listed in Table 2.

1) Bus station and line vector data: The data are collected from the Baidu Map through Open API, which contains station id, name, latitude and longitude, and line label. The line-vector data captures the detailed geographic information of 5 lines with 124 bus stations. Each pair of lines includes two directions: the forward and reverse direction, as shown in Fig. 6.
2) Bus trajectory data: The data was collected by the GPS terminal devices installed on buses and sampled at low frequency every 60 seconds. The trajectory information includes the bus plate number, time of data acquisition, instantaneous speed, direction, latitude, and longitude.
3) Smart IC-card transaction records: The data includes passenger's anonymized ID, time of card swipe, travel mode, bus IDs, consumer prices and other information,

TABLE 2. A summary of experiment data.

Data Type	Description	Feature
Station and line data	5 lines with 124 bus stations	station id, station latitude and longitude, station name and bus line of the station
Bus trajectory data	Sampling at low frequency every 60 seconds	bus plate number, instantaneous speed, direction angle, time of data acquisition instantaneous latitude and longitude,
Smart IC-card transaction records	information of smart IC-card recorded and a total of 902,509 data records from 6:00 to 23:00 every day	passenger's anonymous ID, time of swiping card, buses ID

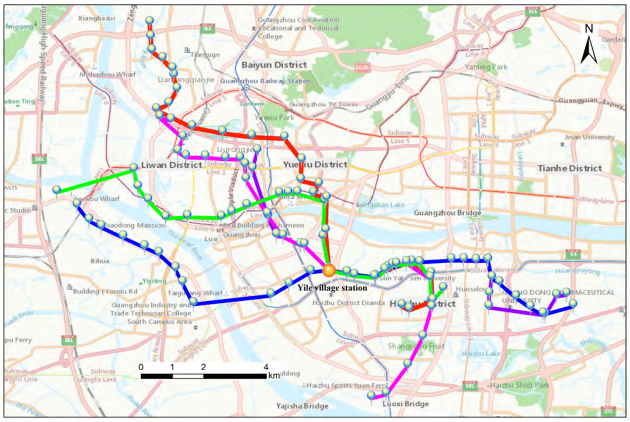


FIGURE 6. The bus line structure.

and a total of 902,509 smart IC-card transaction records from 6:00 am to 11:00 pm.

B. SPECULATION OF BUS ARRIVAL TIME AND PASSENGER FLOW RESULTS

The passenger flow and arrival time distributions are obtained using station matching as well as the inference of boarding and transfer stations. Fig. 7 depicts the passenger flow distribution from 6:00 am to 11:00 pm. During the period from 7:00 am to 9:00 am, and from 5:30 pm to 7:30 pm, the number of passengers boarding and alighting is the largest in a given day. The number of boarding passengers is larger than the number of alighting passengers before 15:00, and the alighting passenger quantity is larger than the boarding passenger quantity.

Fig. 8 illustrates the distribution of passenger arrival rate in 17-time segments. The passenger arrival rate in peak periods is higher than off-peak periods. Fig. 9 illustrates the daily passenger transfers from April 24th to May 20th. The April 29th, 30th and 1st are holidays, and the weekends include May 5th, 6th, 12th, 13th, 19th and 20th, the transfer passengers are the minimum on weekends and holidays. Fig. 10 shows boarding passenger and alighting passenger distribution from

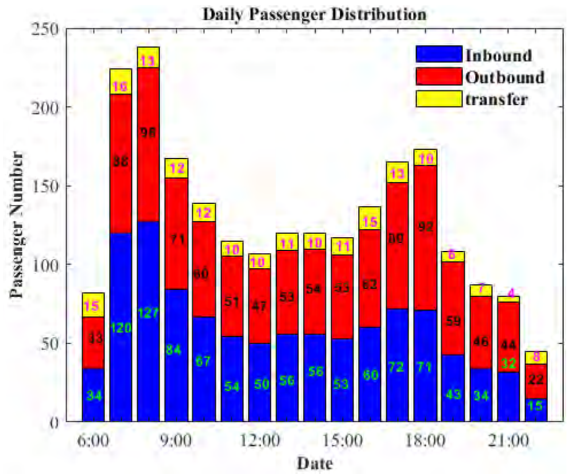


FIGURE 7. The average daily passenger distribution.

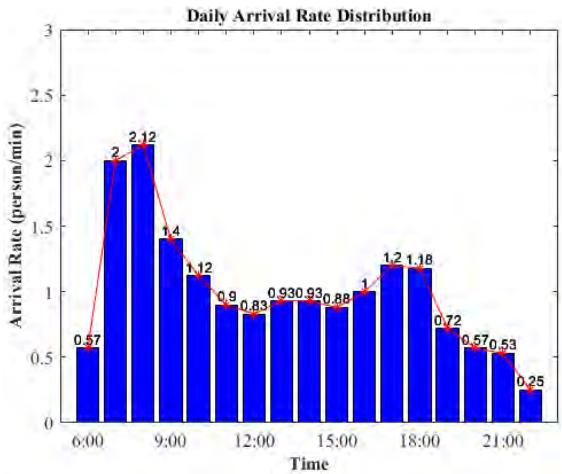


FIGURE 8. The daily arrival rate distribution.

April 24th to May 20th on the No. 14 bus. The boarding passengers and alighting passengers are the minimum on weekends and holidays.

The number of boarding, alighting and transfer passengers' statistics from Fig. 7 to Fig. 10 are used as inputs for the P-LSTM model to predict the future passenger flow and eventually used in the bus-dispatching model.

C. PREDICTION RESULTS

In our predictive model, 80% of the data in the dataset is used for training; the remaining 20% of the data is used for testing. The parametric adjustment process includes four steps [55]: 1) setting an acceptable predicting results, 2) preliminarily setting our parameter values based on previous researches, training and observing the loss changes, then, determining the range of each parameter, 3) adjusting parameters by control variable method; 4) iterating and training until loss drops, finally, showing a stabilized trend, then, saving this parameter. The training parameters are shown in Table 3.

In the prediction process, the mean absolute percentage error ($\overline{MAPE}$) and root mean square error ($\overline{RMSE}$) are used

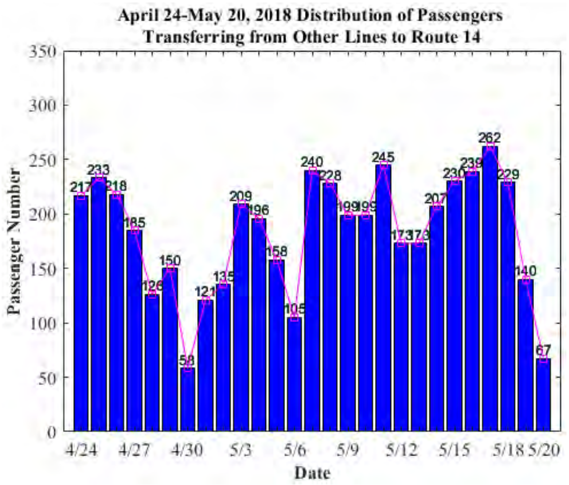


FIGURE 9. The total passenger transfer distribution.

TABLE 3. The detailed parameter setting of LSTM model.

Parameter	Description	Value
rnn_unit	Number of Hidden Layer Neurons	10
lstm_layers	Number of Hidden Layers	3
learning_rate	Learning rate in the training process	0.0006
keep_prob	Probability of retained neurons in Dropout Layer	0.5
batch_size	size of batch training	40
time_step	time step	17

as indicators to measure the performance of the prediction model. The calculation formulae of $\overline{MAPE}$ and $\overline{RMSE}$ are given in equations (17) and (18), respectively.

$$\overline{MAPE} = \frac{1}{n}\sum_{i=1}^{n}\frac{|x_h(i) - \tilde{x}_h(i)|}{x_h(i)} \times 100\% \tag{17}$$

$$\overline{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}|x_h(i) - \tilde{x}_h(i)|^2}{n}} \tag{18}$$

In the formulae, $x_h(i)$ represents the true hourly passenger numbers obtained from the IC card record, x'' indicates the predicted average hourly passenger numbers based on our proposed model, and i denotes sequence number of the time segment.

1) PASSENGER PREDICTION

In this research, four stations across five lines are selected to exhibit typical passenger flow prediction results.

Table 4 shows the $\overline{RMSE}$ and $\overline{MAPE}$ of boarding and alighting passenger prediction on the four stations (results of all 21 stations are attached in appendix). The lowest $\overline{MAPE}$ and $\overline{RMSE}$ are 8.9% and 1.43 respectively; the highest $\overline{MAPE}$ and $\overline{RMSE}$ are 22.6% and 22.16 respectively; The average $\overline{MAPE}$ and $\overline{RMSE}$ is 13.4% and 7.52 respectively.

TABLE 4. Passenger flow prediction results of the four stations.

Station	Mode	MAPE	RMSE
Kangle village	inbound	14.15%	4.8273
	outbound	13.25%	4.7810
shengrong military hospital	inbound	9.59%	6.2603
	outbound	10.27%	6.7247
Sun Yat-sen University	inbound	15.91%	7.6450
	outbound	11.47%	6.8183
Yile Village	inbound	12.26%	7.2314
	outbound	11.83%	7.7460

The predicted and real-time passenger flows of the four stations are shown in Fig. 11. The red curve represents the real-time passenger flow, and the blue curve indicates the predicted passenger flow. The prediction results show that the distribution characteristics of the passengers vary from weekends to non-weekends at different stations, which are close to the real ones. At the same time, through comparative analysis, we can conclude that the passenger flow of the workday fluctuates considerably from 7:00 am to 9:00 am and 5:30 pm to 7:00 pm, and the weekend fluctuations are relatively small.

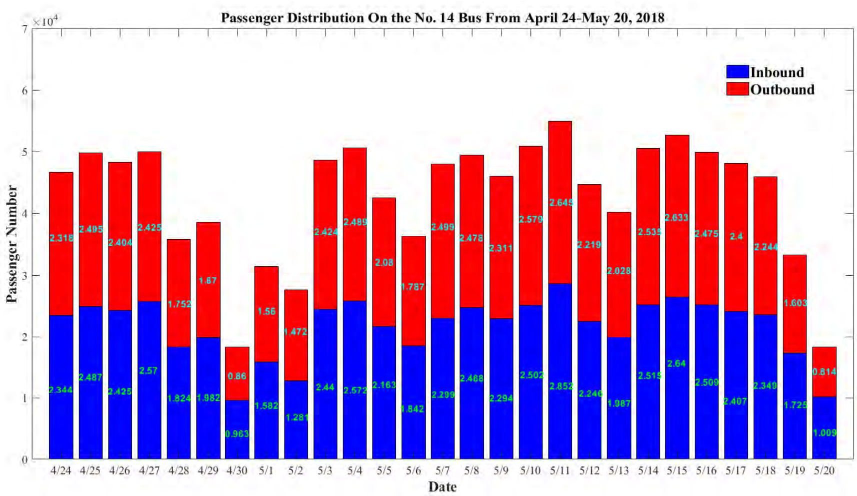


FIGURE 10. The total boarding passenger and alighting distribution.

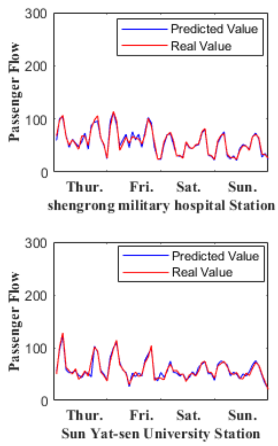


FIGURE 11. Distributions of real-time and predicted passenger flow.

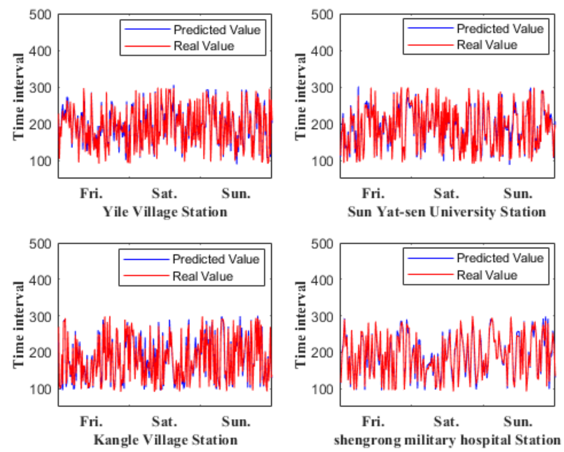


FIGURE 12. Distributions of real-time and predicted arrival time.

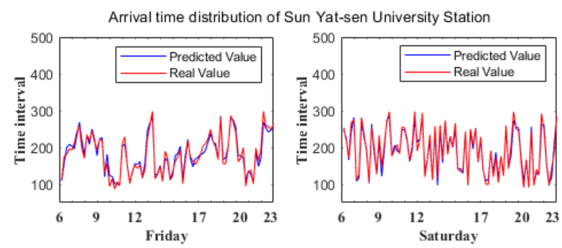


FIGURE 13. Arrival time distributions of Sun Yat-sen UNIVersity Station.

Therefore, through passenger prediction, we can keep track of future passenger flows at all stations.

In this article, only the transfer at Yile Village station is considered. Therefore, the $\overline{RMSE}$ and $\overline{MAPE}$ of transfer passenger prediction are 15% and 1.66.

2) BUS ARRIVAL TIME PREDICTION

Bus arrival time prediction results at four stations are shown in Table 5 (results of all 21 stations are attached in appendix). The lowest $\overline{MAPE}$ and $\overline{RMSE}$ are 4.7% and 10.5 respectively; the highest $\overline{MAPE}$ and $\overline{RMSE}$ are 12.6% and 16.9 respectively; The average $\overline{MAPE}$ and $\overline{RMSE}$ is 7.5% and 13.5 respectively.

TABLE 5. Arrival time prediction results of the four stations.

Station	MAPE	RMSE
Kangle village	8.33%	14.2427
shengrong military hospital	8.91%	14.5016
Sun Yat-sen University	7.96%	14.4057
Yile village	6.65%	13.1379

The real-time and predicted arrival time are shown in Fig. 12, including one working day and two days at the weekend. The blue curve represents the predicted arrival time, and the red curve indicates the real-time arrival time, they have similar trends, which shows that the predicted arrival time can better depict the real arrival time changes. As shown in Fig. 13, compared with weekends and non-weekends, the average weekends waiting time is longer than non-weekends. Meanwhile, peak waiting time is longer than the off-peak waiting time. Therefore, by predicting the future arrival time of buses, a lot of arrival time information is extracted at different stations in advance, and it provided convenience for subsequent bus dispatching.

D. GABD BUS DISPATCHING

After predicting the passenger flow and bus arrival time through the P-LSTM model, we took out 50 buses as a group from real-time data and predicted data, respectively, and encoded the chromosome with a length of 500 using a 10-bit binary encoding. According to the survey, the average boarding time for each person is 4 seconds, and the average alighting time is 2 seconds. During the experiment, the crossover probability is 0.6, the mutation probability is 0.01, and the number of iterations is 1000, the parameters used in the experiment is shown in Table 6.

TABLE 6. The detailed parameter setting of genetic algorithm.

Parameter	Description	Value
population_size	Number of selected populations	50
generations	The Number of Iterations of Genetic Algorithms	1000
chrom_length	The decimal coding length of each chromosome	500
pc	Cross probability set in crossover operation	0.6
pm	Probability of gene mutation	0.01

The experimental results are shown in Fig. 14 and Fig. 15, which represent the iteration results of predicted and real-time data, respectively. We have done 9 experiments in succession with the same parameters, represented by nine curves. As can be seen from these curves in Fig. 14 and Fig. 15,

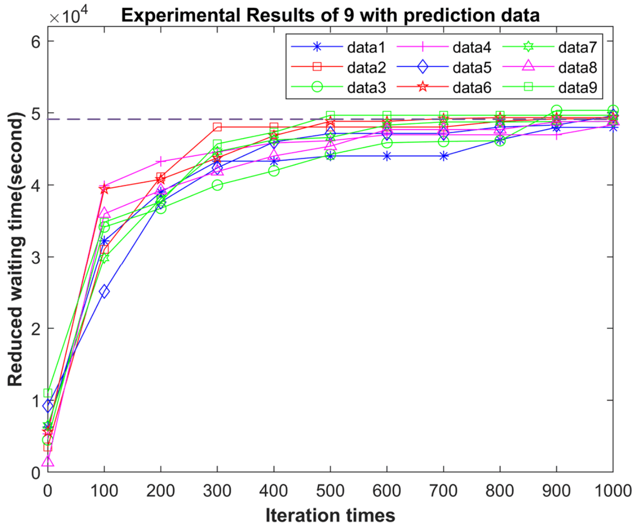


FIGURE 14. Test results of genetic algorithm with prediction data.

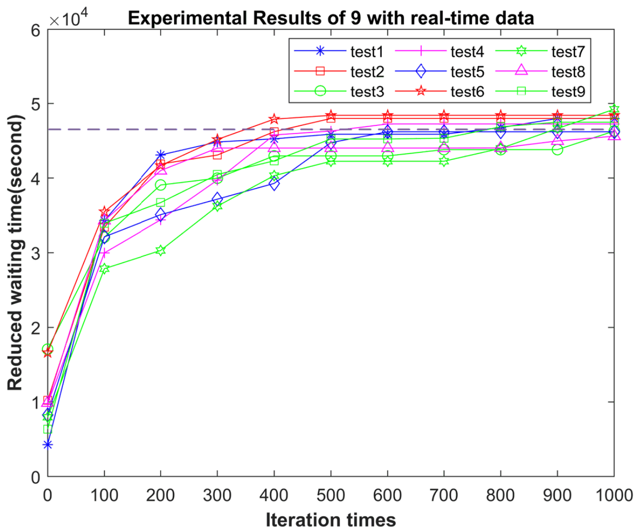


FIGURE 15. Test results of genetic algorithm with real-time data.

the waiting time converges after 500 iterations. The average convergence value of predicted and real-time data are 829.68 minutes and 783.19 minutes, which indicates that the stay strategy based on passenger flow and bus arrival time prediction is better than the stay strategy based on real-time passenger flow and bus arrival time, and the reduced time is increased by 5.94%.

The experimental results show that the D-ATPF model can accurately predict the passenger flow of all stations on the bus line and the arrival time of different buses using the LSTM network, which allows us to effectively use the predicted information to dispatch buses in advance. Meanwhile, the stay strategy based on passenger flow and bus arrival time prediction outperforms the real-time-based stay strategy. The D-ATPF model significantly reduces passenger waiting time and provides an effective tool for the stay strategy dispatching at transfer stations.

V. CONCLUSION

This article has presented a novel bus dispatching model based on dynamic arrival times and passenger flow predictions. By predicting future passenger flow and bus arrival times, future passenger flow and bus arrival time information was extracted to plan bus dispatching in advance. Two benefits are manifested in this work. Firstly, the presented method ensures that more passengers can catch up with stranded buses at transfer and subsequent stations, thus reducing their waiting time. Secondly, a bus stay strategy is proposed for transfer stations to avoid long delays waiting for other buses. In this way, the total waiting time of passengers is minimized. Then, the optimal solution of the nonlinear function is found using a genetic algorithm. From the experimental results, the D-ATPF model significantly reduces the waiting time for passengers. Compared with the stay strategy of real-time data, the reduced waiting time of this method is increased by 5.94%. However, the limitation is that 1) bus capacity spillovers rarely occur during the limited detention time. Therefore, bus capacity is not considered, which is one of the limitations of this method. 2) this method requires massive computing resources due to a large amount of data. 3) The current model does not include the dynamic of bus speed, which also provides great potential to improve bus operation.

Our future work includes adjusting the speed of the bus, reducing the stay time in the transfer stations and considering the capacity of buses. Meanwhile, we will add a road traffic accident prediction model to provide the necessary emergency strategy for our scheduling strategy.

REFERENCES

[1] H. Derrouz, A. Elbouziady, H. A. Abdelali, R. O. H. Thami, S. El Fkihi, and F. Bourzeix, "Moroccan video intelligent transport system: Vehicle type classification based on three-dimensional and two-dimensional features," *IEEE Access*, vol. 7, pp. 72528–72537, 2019.

[2] Q. E. Ali, N. Ahmad, A. H. Malik, G. Ali, M. Asif, M. Khalid, and Y. Cao, "SPATA: Strong pseudonym-based AuthenTicAtion in intelligent transport system," *IEEE Access*, vol. 6, pp. 79114–79128, 2018.

[3] X. Luo, X. Zhao, L. Sun, K. Ma, and J. Tang, "Dynamic bus dispatching under the environment of Internet of Things," in *Proc. 11th World Congr. Intell. Control Automat.*, Shenyang, China, Jun./Jul. 2014, pp. 449–454.

[4] B. Yu, S. Wu, B. Yao, and Z. Yang, "Dynamic vehicle dispatching at a transfer station in public transportation system," *J. Transp. Eng.*, vol. 138, no. 2, pp. 191–201, Feb. 2012.

[5] Y. Xuan, J. Argote, and C. F. Daganzo, "Dynamic bus holding strategies for schedule reliability: Optimal linear control and performance analysis," *Transp. Res. B, Methodol.*, vol. 45, no. 10, pp. 1831–1845, Dec. 2011.

[6] X.-G. Luo, H.-B. Zhang, Z.-L. Zhang, Y. Yu, and K. Li, "A new framework of intelligent public transportation system based on the Internet of Things," *IEEE Access*, vol. 7, pp. 55290–55304, 2019.

[7] B. Yu, Y. Cui, and W. J. Cai, "Multiline transit coordination at a hub based on bus-arrival time prediction," *Oper. Res. Manage. Sci.*, vol. 24, no. 4, pp. 65–70, Aug. 2015.

[8] S. J. Berrebi, K. E. Watkins, and J. A. Laval, "A real-time bus dispatching policy to minimize passenger wait on a high frequency route," *Transp. Res. B, Methodol.*, vol. 81, pp. 377–389, Nov. 2015.

[9] S. An and X. Zhang, "Real-time hybrid in-station bus dispatching strategy based on mixed integer programming," *Information*, vol. 7, no. 3, p. 43, 2016.

[10] K. Vanitchakornpong, N. Indra-Payoong, A. Sumalee, and W. Sriborrirux, "Performance evaluation of a RFID-enabled real time bus dispatching system: Case study of the Bangkok bus system," *ECTI Trans. Comput. Inf. Technol.*, vol. 6, no. 2, pp. 144–155, 2012.

[11] Q. Chen, E. Adida, and J. Lin, "Implementation of an iterative headway-based bus holding strategy with real-time information," *Public Transp.*, vol. 4, no. 3, pp. 165–186, 2013.

[12] X. Luo, Y. Liu, Y. Yu, J. Tang, and W. Li, "Dynamic bus dispatching using multiple types of real-time information," *Transportmetrica B, Transp. Dyn.*, vol. 7, no. 1, pp. 519–545, Mar. 2019.

[13] J. Zhang, D. Shen, L. Tu, F. Zhang, C. Xu, Y. Wang, C. Tian, X. Li, B. Huang, and Z. Li, "A real-time passenger flow estimation and prediction method for urban bus transit systems," *IEEE Trans. Intell. Transp. Syst.*, vol. 18, no. 11, pp. 3168–3178, Nov. 2017.

[14] L. Moreira-Matias, J. Gama, M. Ferreira, J. Mendes-Moreira, and L. Damas, "Predicting taxi-passenger demand using streaming data," *IEEE Trans. Intell. Transp. Syst.*, vol. 14, no. 3, pp. 1393–1402, Sep. 2013.

[15] M. L. Tam, W. H. K. Lam, B. Chen, and K. S. Chan, "Using automatic vehicle identification data for investigation of rain effects on vehicular travel speeds and travel choice behaviour," in *Proc. Conf. Traffic Transp. Stud. (ICTTS)*, Nanning, China, 2008, pp. 944–955.

[16] Y. Chen, Y. Sha, and X. Zhu, "Prediction of shanghai metro line 16 passenger flow based on time series analysis-with Lingang avenue station as a study case," *Oper. Res. Fuzzy*, vol. 6, no. 1, pp. 15–26, Jan. 2016.

[17] Z. Ma, J. Xing, M. Mesbah, and L. Ferreira, "Predicting short-term bus passenger demand using a pattern hybrid approach," *Transp. Res. C, Emerg. Technol.*, vol. 39, pp. 148–163, Feb. 2014.

[18] W. Suwardo, M. Napiah, and I. Kamaruddin, "ARIMA models for bus travel time prediction," *J. Inst. Eng. Malaysia*, vol. 71, no. 1, pp. 49–58, Jun. 2010.

[19] M. Yang, C. Chen, L. Wang, X. Yan, and L. Zhou, "Bus arrival time prediction using support vector machine with genetic algorithm," *Neural Netw. World*, vol. 26, no. 3, pp. 205–217, Jan. 2016.

[20] J. Z. Zhou and D. Y. Zhang, "Direct ridership forecast model of urban rail transit stations based on spatial weighted LS-SVM," *J. China Railway Soc.*, vol. 36, no. 1, pp. 1–7, Jan. 2014.

[21] T.-H. Tsai, C.-K. Lee, and C.-H. Wei, "Neural network based temporal feature models for short-term railway passenger demand forecasting," *Expert Syst. Appl.*, vol. 36, no. 2, pp. 3728–3736, Mar. 2009.

[22] J. Ding, M. Yang, Y. Cao, and S. Li Kong, "Dwell time prediction of bus rapid transit using ARIMA-SVM hybrid model," *Appl. Mech. Mater.*, vols. 587–589, pp. 1993–1997, Jul. 2014.

[23] B. Yu, Z.-Z. Yang, K. Chen, and B. Yu, "Hybrid model for prediction of bus arrival times at next station," *J. Adv. Transport. Banner*, vol. 44, no. 3, pp. 193–204, 2010.

[24] X. Luo, D. Li, Y. Yang, and S. Zhang, "Spatiotemporal traffic flow prediction with KNN and LSTM," *J. Adv. Transp.*, vol. 2019, pp. 1–10, Feb. 2019.

[25] Y. Zhang, J. Zhu, and J. Zhang, "Short-term passenger flow forecasting based on phase space reconstruction and LSTM," in *Proc. Int. Conf. Elect. Inf. Technol. Rail Transp.*, Beijing, China, vol. 482, 2017, pp. 679–688.

[26] N. C. Petersen, F. Rodrigues, and F. C. Pereira, "Multi-output bus travel time prediction with convolutional LSTM neural network," *Expert Syst. Appl.*, vol. 120, pp. 426–435, Apr. 2019.

[27] W. Tu, Q. Li, Z. Fang, S.-L. Shaw, B. Zhou, and X. Chang, "Optimizing the locations of electric taxi charging stations: A spatial–temporal demand coverage approach," *Transp. Res. C, Emerg. Technol.*, vol. 65, pp. 172–189, Apr. 2016.

[28] W. Tu, Q. Li, Z. Fang, and B. Zhou, "A novel spatial-temporal Voronoi diagram-based heuristic approach for large-scale vehicle routing optimization with time constraints," *ISPRS Int. J. Geo-Inf.*, vol. 4, no. 4, pp. 2019–2044, Apr. 2015.

[29] H. Li, L. Lei, and X. Feng, "Research on public transportation scheduling optimization based on particle swarm optimization (PSO)—Taking Guangzhou as an example," *J. Guangzhou Univ. (Natural Sci. Edition)*, vol. 98, no. 2, pp. 65–70, 2018.

[30] M. Pang, M. Chen, and N. Zhang, "Scheduling optimization of intelligent public transport system based on MAST," *J. Transp. Syst. Eng. Inf. Technol.*, vol. 17, no. 1, pp. 1009–6744, Feb. 2017.

[31] F. Zhang and W. Liu, "Responsive bus dispatching strategy in a multi-modal and multi-directional transportation system: A doubly dynamical approach," *Transp. Res. C, Emerg. Technol.*, to be published.

[32] M. A. Forbes, J. N. Holt, P. J. Kilby, and A. M. Watts, "BUDI: A software system for bus dispatching," *J. Oper. Res. Soc.*, vol. 45, no. 5, pp. 497–508, 1994.

[33] J. G. Strathman, K. J. Dueker, and T. J. Kimpel, "Automated bus dispatching, operations control, and service reliability: Baseline analysis," *Transp. Res. Rec.*, vol. 1666, no. 1, pp. 28–36, 1999.

[34] J. Du, C. Wang, and F. Zhang, "Multi-objective optimization of bus dispatching based on improved genetic algorithm," in *Proc. 7th Int. Conf. Comput. Intell. Secur.*, Sanya, China, Dec. 2011, pp. 106–109.

[35] C.-J. Ting and P. Schonfeld, "Dispatching control at transfer stations in multi-hub transit networks," *J. Adv. Transp.*, vol. 41, no. 3, pp. 217–243, Jun. 2007.

[36] X. Zuo, C. Chen, W. Tan, and M. Zhou, "Vehicle scheduling of an urban bus line via an improved multiobjective genetic algorithm," *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 2, pp. 1030–1041, Apr. 2015.

[37] M. H. Moradi and M. Abedini, "A combination of genetic algorithm and particle swarm optimization for optimal DG location and sizing in distribution systems," *Int. J. Elect. Power Energy Syst.*, vol. 34, no. 1, pp. 66–74, Jan. 2012.

[38] H. Bai, Z.-H. Xia, Y.-D. Zhao, and K.-L. Song, "A departure frequency model for double-line BRT with genetic algorithm," in *Proc. 10th Int. Conf. Chin. Transp. Prof. North Amer. Chin. Overseas Transp. Assoc. Beijing Univ. Technol. Amer. Soc. Civil Eng. Transp. Res. Board Nat. Natural Sci. Found. China*, Beijing, China, 2010, pp. 1114–1121.

[39] L. Liu and R.-C. Chen, "A novel passenger flow prediction model using deep learning methods," *Transp. Res. C, Emerg. Technol.*, vol. 84, pp. 74–91, Nov. 2017.

[40] T. Xu, M. Yang, J. Wu, D. Lei, and Y. Wu, "Passenger flow prediction of urban public transport hubs based on a deep learning approach: A case study in Changzhou," in *Proc. CICTP*. Reston, VA, USA: American Society of Civil Engineers, 2018, pp. 755–765.

[41] N. G. Polson and V. O. Sokolov, "Deep learning for short-term traffic flow prediction," *Transp. Res. C, Emerg. Technol.*, vol. 79, pp. 1–17, Jun. 2017.

[42] Y. Sun, B. Leng, and W. Guan, "A novel wavelet-SVM short-time passenger flow prediction in Beijing subway system," *Neurocomputing*, vol. 166, pp. 109–121, Oct. 2015.

[43] Y. Lv, Y. Duan, W. Kang, Z. Li, and F.-Y. Wang, "Traffic flow prediction with big data: A deep learning approach," *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 2, pp. 865–873, Apr. 2015.

[44] R. A. Wahab, M. N. Borhan, and R. A. Rahmat, "Prediction of bus arrival times at bus stop," *Int. J. Technol.*, vol. 8, no. 1, pp. 158–165, Jan. 2017.

[45] C. Bai, Z.-R. Peng, Q.-C. Lu, and J. Sun, "Dynamic bus travel time prediction models on road with multiple bus routes," *Comput. Intell. Neurosci.*, vol. 2015, Jul. 2015, Art. no. 63.

[46] S. I.-J. Chien, Y. Ding, and C. Wei, "Dynamic bus arrival time prediction with artificial neural networks," *J. Transp. Eng.*, vol. 128, no. 5, pp. 429–438, 2002.

[47] A. Gal, A. Mandelbaum, F. Schnitzler, A. Senderovich, and M. Weidlich, "Traveling time prediction in scheduled transportation with journey segments," *Inf. Syst.*, vol. 64, pp. 266–280, Mar. 2017.

[48] B. Y. Chen, H. Yuan, Q. Li, W. H. K. Lam, S.-L. Shaw, and K. Yan, "Map-matching algorithm for large-scale low-frequency floating car data," *Int. J. Geograph. Inf. Sci.*, vol. 28, no. 1, pp. 22–38, 2014.

[49] S. Wei, Z. Yu, and T. Wang, "Model research on passenger on and off site based on bus IC card and GPS data," *Open J. Transp. Technol.*, vol. 7, no. 3, pp. 140–146, 2018.

[50] M. A. Munizaga and C. Palma, "Estimation of a disaggregate multimodal public transport origin–destination matrix from passive smartcard data from Santiago, Chile," *Transp. Res. C, Emerg. Technol.*, vol. 24, pp. 9–18, Oct. 2012.

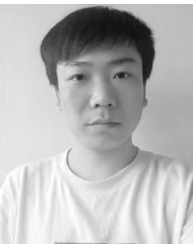
[51] M. Beamurgia, R. Basagoiti, I. Rodríguez, and V. Rodriguez, "A modified genetic algorithm applied to the elevator dispatching problem," *Soft Comput.*, vol. 20, no. 9, pp. 3595–3609, Sep. 2016.

[52] H. Li, D.-M. Luo, X.-P. Luo, C. Kun, Z.-L. Guo, and G. Tao, "The application of foreground theory and genetic algorithm in bus dispatching," in *Proc. 24th Int. Conf. Ind. Eng. Eng. Manage.* Singapore: Springer, 2019, pp. 378–387.

[53] M. Qiu, Z. Ming, J. Li, K. Gai, and Z. Zong, "Phase-change memory optimization for green cloud with genetic algorithm," *IEEE Trans. Comput.*, vol. 64, no. 12, pp. 3528–3540, Dec. 2015.

[54] J. Xia, K. M. Curtin, J. Huang, D. Wu, W. Xiu, and Z. Huang, "A carpool matching model with both social and route networks," *Comput., Environ. Urban Syst.*, vol. 75, pp. 90–102, May 2019.

[55] N. Reimers and I. Gurevych, "Optimal hyperparameters for deep LSTM-networks for sequence labeling tasks," 2017, *arXiv:1707.06799*. [Online]. Available: https://arxiv.org/abs/1707.06799



ZHAO HUANG is currently pursuing the master's degree with the College of Information Engineering, Shenzhen University, China. His research interests include spatiotemporal data mining and intelligent transportation systems.

QINGQUAN LI received the M.S. degree in engineering and the Ph.D. degree in photogrammetry and remote sensing from Wuhan University, Wuhan, China, in 1988 and 1998, respectively. From 1988 to 1996, he was an Assistant Professor with Wuhan University, where he became an Associate professor, from 1996 to 1998, and has been a Professor with the State Key Laboratory of Information Engineering in Surveying, Mapping, and Remote Sensing, since 1998. He is currently the President of Shenzhen University, Shenzhen, China, where he is also the Director of the Shenzhen Key Laboratory of Spatial Information Smart Sensing and Services. His research interests include photogrammetry, remote sensing, and intelligent transportation systems. He is an Expert in modern traffic with the National 863 Plan and an Editorial Board Member of the *Surveying and Mapping Journal* and the *Wuhan University Journal*—Information Science Edition.

FAN LI is currently pursuing the master's degree with the College of Computer Science and Software Engineering, Shenzhen University, China. His research interests include spatiotemporal data mining and intelligent transportation systems.

JIZHE XIA received the Ph.D. degree in earth systems and geoinformation sciences from George Mason University, in 2015. From 2015 to 2017, he was a Research Assistant Professor with the National Science Foundation (NSF) Spatiotemporal Invitation Center. In 2017, he joined the Department of Urban Informatics, Shenzhen University, China, as an Assistant Professor. His research interests include spatial data indexing, spatiotemporal optimization, and geographic information applications.