Hindawi
Journal of Advanced Transportation
Volume 2020, Article ID 8846715, 16 pages
https://doi.org/10.1155/2020/8846715

Research Article

Dynamic Origin-Destination Matrix Estimation Based on Urban Rail Transit AFC Data: Deep Optimization Framework with Forward Passing and Backpropagation Techniques

Yuedi Yang,[1] Jun Liu,[1] Pan Shang,[1] Xinyue Xu,[2] and Xuchao Chen[3]

[1]*School of Traffic and Transportation, Beijing Jiaotong University, Beijing 100044, China*
[2]*State Key Laboratory of Railway Traffic Control and Safety, Beijing Jiaotong University, Beijing 100044, China*
[3]*Beijing Infrastructure Investment Co., LTD, Beijing 100044, China*

Correspondence should be addressed to Jun Liu; jliu@bjtu.edu.cn and Pan Shang; shangpan@bjtu.edu.cn

Received 30 March 2020; Revised 10 October 2020; Accepted 23 November 2020; Published 7 December 2020

Academic Editor: Jose E. Naranjo

At present, the existing dynamic OD estimation methods in an urban rail transit network still need to be improved in the factors of the time-dependent characteristics of the system and the estimation accuracy of the results. This study focuses on predicting the dynamic OD demand for a time of period in the future for an urban rail transit system. We propose a nonlinear programming model to predict the dynamic OD matrix based on historic automatic fare collection (AFC) data. This model assigns the passenger flow to the hierarchical flow network, which can be calibrated by backpropagation of the first-order gradients and reassignment of the passenger flow with the updated weights between different layers. The proposed model can predict the time-varying OD matrix, the number of passengers departing at each time, and the travel time spent by passengers, of which the results are shown in the case study. Finally, the results indicate that the proposed model can effectively obtain a relatively accurate estimation result. The proposed model can integrate more traffic characteristics than traditional methods and provides an effective and hierarchical passenger flow estimation framework. This study can provide a rich set of passenger demand for advanced transit planning and management applications, for instance, passenger flow control, adaptive travel demand management, and real-time train scheduling.

1. Introduction

As an important part of passenger flow prediction, in an urban rail transit system, origin-destination (OD) matrix estimation plays an important role, which provides basic data for passenger flow assignment. Most of the developed approaches are usually applied to road traffic systems, such as freeway, highway, and road networks, in which link (section) flow can usually be obtained by detectors. However, the development of approaches for estimating dynamic traffic demand for large-scale and complex rail transit networks without observed link flows, such as the Beijing and Shanghai subway systems in China, remains a critical and challenging problem, and this issue has been attracting a significant amount of attention from transport operation researchers and managers [1].

Many approaches have been proposed for distributing trips among origins and destinations over the years. The gravity model is a typical traditional model that can predict the OD distribution of traffic zones [2]. However, conventional approaches based on economic, population, and spatial relationship data are time-consuming and are highly expensive.

Some estimation approaches have become popular in OD estimation or updating the OD matrix from traffic counts. The scope of the literature in this field is very wide. The existing works are based on the entropy maximization [3], the maximum likelihood approach [4], the generalized least squares estimator [5, 6], the Bayesian inference approach [7, 8], and other approaches to solving this problem. However, due to the complexity of calculating this parameter, mostly the developed models based on the assignment matrix are limited to simple networks like intersections, interchanges, and freeways [9].

These conventional methods for estimating origin-destination (OD) trip matrices from link traffic counts assume that route choice proportions are given constants. But this assumption does not hold in a network with realistic congestion levels [10]. A bilevel programming approach has been used for the estimation of the OD matrix in congested networks [11, 12]. This approach combined the generalized least squares estimation model and the network equilibrium model into one process. However, the bilevel approach has certain difficulties in finding an optimal solution because of nonconvex and nondifferential formulations. Sherali et al. [13] constructed a linear programming model with a user-equilibrium solution for synthesizing OD tables from traffic volume counts. Later, Toledo and Kolechkina [14] presented the methods based on use of linear approximations of the assignment matrix in the optimization iterations. Fujita et al. [15] proposed an OD modification approach formulated as a static user-equilibrium assignment with elastic demand, based on the residual demand at the end of each period. Applying the model to large-scale road network demonstrates that it efficiently improves estimation accuracy because the 24-hour time coefficients of survey data are slightly biased and may be modified properly. Unlike the gravity model, these approaches are based on traffic count data, which can be detected from links by vehicle identification or locating technologies, such as GPS floating, automatic license plate recognition (ALPR), and radio frequency identification (RFID). Tang et al. [16] proposed a new method based on the entropy-maximizing theory to model OD distribution in Harbin city using large-scale taxi GPS trajectories. The results demonstrate that the entropy-maximizing model is superior to the gravity model, which can validate the feasibility of OD distribution from taxi GPS data in the urban system. Rao et al. [17] formulated a particle filter model for vehicle trajectory reconstruction based on ALPR data. And the OD patterns are estimated by adding up the path flows, which is conducted through dividing the reconstructed complete trajectories. Liu et al. [18] predicted the OD matrix based on the historical ALPR data. Guo et al. [19] developed an optimization model based on the least squares method. The optimization model estimated the dynamic OD matrix by integrating the preliminary OD matrix, dynamic assignment matrix derived by RFID data, and link flow detected by the inductive loop detectors. However, the route choice and travel time delay issues are still difficult to deal with. So, establishing the dynamic flow equations is the first challenge to estimate dynamic passenger flow demand for rail transit systems, for the lack of observed information and complex structure of the network.

In recent decades, the application of the neural network model expands this problem into a new field. The neural network operates as a black box, model-free, and adaptive tool for capturing and learning significant structures in data [20]. Gong [21] used the Hopfield neural network (HNN) model to estimate the urban orientation-destination (OD) distribution matrix from the link volumes of the transportation network so as to promote the solving speed and precision. Yang et al. [22] proposed a dynamic model based on backpropagation (BP) learning for estimating OD flows from road entering and exiting counts. The OD flows in each short time interval are estimated through the minimization of the squared errors between the predicted and observed exiting counts. Li et al. [23] proposed a new dynamic radial basis function neural network to forecast outbound passenger volumes and improve passenger flow control. Passenger flow control was considered to improve the prediction accuracy by adding passenger flow control coefficients to their model. However, the current perceptron neural networks may not perform well in all issues due to reasons such as model nontransferability, insufficient ability to generalize, and reliance on activation functions [24].

Subsequently, the computational graph was proposed as a description language to represent mathematical expressions. It is important to understand how the underlying computational graph of a deep learning network, combined with the BP algorithm, can be used to describe the forward propagation and backward feedback processes between different levels of transportation planning and decision making [25, 26]. Wu et al. [27] proposed a multilayered hierarchical flow network representation to structurally model different levels of travel demand for road networks, including trip generation, OD matrices, path and link flows, and individual behavior parameters. However, the travel times were assumed to be observed in their study. In other words, their model was constructed in a static network rather than a time-dependent dynamic network. This issue has been improved in this paper.

In this paper, we aim to predict the dynamic OD demand for a time of period in the future based on historical observations. In most research papers, it is assumed that the OD matrix can be predicted from historical data [18, 28]. We apply the historic AFC to train a time-dependent hierarchical flow network. Then, we use it to predict the future OD demand with real-time AFC data at current as input information, which is the basic data to formulate operational and organizational strategies. Besides, the programming model proposed in this paper can also estimate a hierarchical traveling decision process for passengers in an urban rail transit system, including the departure time choice at the origin, the path choice, and the corresponding arrival time at the destination. Therefore, the proposed method in this study achieves a combination prediction of dynamic OD matrix, departure time choice, route choice, and travel time.

A nonlinear programming model is proposed to conduct real-time OD matrix estimation for an urban rail transit system based on historic automatic fare collection (AFC) data in this paper. Forward passing in the hierarchical flow network of urban rail transit sequentially assigns passengers to candidate stations, paths, and different travel time intervals. The network can be calibrated by backpropagation of the first-order gradients and reassignment of the passenger flow with the updated weights between different layers under the deep optimization framework. This model can determine the time-varying OD matrix, the number of passengers departing at each time, and the travel time spent by passengers, of which the results are shown in the case study. Finally, a comparative analysis with artificial neural networks is conducted to illustrate the effect and efficiency of the proposed model.

The potential contributions are as follows:

(1) A modeling framework using the multilayer hierarchical flow network is applied to describe the passenger transit process in an urban rail system. Based on the flow-oriented prediction formulation, this deep learning modeling approach can simultaneously estimate different levels of unobserved or partially observed passenger flow variables. This model is applicable to the estimation of the OD matrix of passenger flow with AFC data, unlike other traditional estimation methods based on traffic counts.

(2) This modeling paradigm enables us to capture the mathematical structure inside the OD matrix estimation problem by representing and decomposing complex composite functions through a graph of current states and numerical gradients. This model is constructed by the passengers' trip process, unlike the black-box model ANN. Therefore, the computational graph can express more traffic characteristics than the ANN and provides an effective and hierarchical passenger flow estimation.

(3) The layered framework provides a flexible mechanism for further expansion. In particular, the framework can easily add a new hierarchical structure to achieve OD estimation when other sensor data sources can be obtained.

(4) In this model, the departure time and travel time are considered as variables, and the additional departure time layer and travel time layer are constructed in the network. It is more reasonable to develop a dynamic hierarchical flow network to estimate the time-dependent OD passenger flow matrix.

The remainder of the paper is organized as follows. The next section presents the mathematical formulation of the time-dependent hierarchical flow network estimation model. In the following section, we present the solution framework for implementing forward and backward propagation. In Section 4, we describe a numerical experiment based on the Beijing Subway and compare the results with the ANN method. The conclusion is given in the last section.

2. Problem Statement

2.1. Problem Statement and Notation. This paper aims to design a time-dependent hierarchical flow network (TDHFN) model according to historic AFC records. The model is based on a passenger assignment network considering time variations. There is an abundance of historic AFC records that can be applied to train an optimal model. An urban rail transit network consists of a set of stations $N, N = \{1, 2, 3 \ldots ,\}$. The departure time is divided into equal time intervals composed of a set $T, T = \{1, 2, \ldots, t, \ldots\}$. The travel time set Γ ($\Gamma = \{1, 2, \ldots, \tau, \ldots\}$) can be obtained from historic AFC data. The path set P ($P = \{1, 2, \ldots, p, \ldots\}$) is the given information consisting of alternative routes for each OD. Therefore, there are 5 layers in the passenger assignment network: origin station, departure time, destination station, paths, and travel time. The passengers are assigned from the origin station to different departure time intervals, assigned to different destination stations, then assigned to different paths, and finally assigned to different travel times.

In the passenger assignment network design problem, the following inputs should be given: (1) AFC records of how many passengers enter at each station, depart at each time interval, exit at each station, and arrive at each time interval and (2) the supply network of the paths of each OD with minimum travel time and maximum travel time.

From the perspective of system-optimal passenger assignment, we can obtain (1) the number of passengers departing at each time; (2) the number of passengers arriving at each station; (3) the number of passengers arriving at each time; and (4) the number of passengers choosing each path.

There is an important assumption in this model: the same departure time interval of different origin stations will be marked as different departure time interval indices, as well as the destination station indices and travel time indices. This ensures that each path for the different destination stations in the network belongs to a different OD.

A multilayer TDHFN is adopted to describe the OD matrix estimation of the urban rail passenger flow problem. The notations used in this paper are shown in Table 1.

2.2. Physical Description. Consider a simple physical urban rail network with four nodes, as shown in Figure 1. Node 1 is the origin station where passengers enter (tap-in) the urban rail system. Nodes 2 and 4 are the destination stations where the passengers exit (tap-out) the system. Node 3 is the transfer station. Four paths belong to two different ODs in this network. We consider a time-space passenger network based on the simple physical network (from Figure 1), as shown in Figure 2.

There is a very important principle in the numbering. With different departure times but equal travel times for the same OD, the destination station, path, and travel time values should be numbered with different indices. Additionally, when different OD pairs have the same departure times and equal travel times, the path and travel time values should be numbered with different indices. Similarly, when different OD pairs have the same departure times, equal travel times, but different paths, the travel time values should be numbered with different indices. This principle ensures that the model proposed is time-dependent. In other words, the passengers departing from the origin station at different times may choose different paths and different travel times. However, in a static network, passengers are often considered to be homogenous, such as in the research of Wu et al. [27]. In this paper, the time-dependent numbering principle can be used to consider the

Table 1: Sets, indices, variables, vectors, and parameters.

	Definitions
Indices	
i	Index of nodes (origin stations), $i \in N$
j	Index of destination stations, $j \in D$
t	Index of departure time intervals, $t \in T$
τ	Index of travel time, $\tau \in \Gamma$
p	Index of paths, $p \in P$
Sets	
N	Set of stations (origin stations)
D	Set of destination station indices
D_t	Set of destination station indices for the departure time t
T	Set of departure time indices
T_i	Set of departure time indices for the origin station i
T_{ij}	Set of departure time indices from the origin station i to the destination station j
Γ	Set of travel time indices
Γp	Set of travel time indices for the path p
P	Set of path indices
P_j	Set of path indices for the destination station j
P_{ijt}	Set of path indices from origin station i to the destination station j at departure time t
Variables	
x_i	Number of passengers entering (tap-in) the system at origin station i
h_t	Number of passengers departing at time t
h_j	Number of passengers exiting (tap-out) the system at destination station j
h_p	Number of passengers choosing the path p
y_τ	Number of passengers spending the travel time τ
Vectors	
X	The vector of the input layer, $X = (x_i \mid i \in N)$
H_T	The vector of hidden layer 1, $H_T = (h_t \mid t \in T)$
H_D	The vector of hidden layer 2, $H_D = (h_j \mid j \in D)$
H_P	The vector of hidden layer 3, $H_P = (h_p \mid p \in P)$
Y	The vector of the output layer, $Y = (y_\tau \mid \tau \in T)$
Parameters	
$\alpha_{i,t}$	The proportion of passengers departing at time t to the passengers entering (tap-in) the system at node i
$\beta_{t,j}$	The proportion of passengers exiting (tap-out) the system at node j to the passengers departing at time t
$\rho_{j,p}$	The proportion of passengers choosing the path p to the passengers exiting (tap-out) the system at node j
$\omega_{p,\tau}$	The proportion of passengers with the travel time τ to the passengers choosing the path p
η	Learning rate
g_t	A gradient of hidden layer 1
g_j	A gradient of hidden layer 2
g_p	A gradient of hidden layer 3
g_τ	A gradient of the output layer
$\Delta\alpha$	The updated value of $\alpha_{i,t}$
$\Delta\beta$	The updated value of $\beta_{t,j}$
$\Delta\rho$	The updated value of $\rho_{j,p}$
$\Delta\omega$	The updated value of $\omega_{p,\tau}$
α	The matrix of the departure time layer proportion $\alpha_{i,t}$
β	The matrix of the destination station layer proportion $\beta_{t,j}$
ρ	The matrix of the path layer proportion $\rho_{j,p}$
ω	The matrix of the travel time layer proportion $\omega_{p,\tau}$
$\widehat{y}_\tau$	Number of real passengers spending the travel time τ from AFC records

characteristics of passenger heterogeneity, which is more practical. Finally, a simple example of the time-dependent numbering principle is shown in Figure 3, which is based on Figures 1 and 2. The indices are shown above the bold horizontal lines.

The numbering of all the stations, departure times, paths, and travel times, as well as the determination of the connection between the decision variables of each level, is the basis of the model. This method is a very important and complex process, especially in a large-scale complicated urban transport network, such as the Beijing Subway.

2.3. Mathematical Description. A TDHFN representation is used as a high-level modeling abstract to formulate the OD matrix estimation problem. Let a TDHFN $G = G(V, E)$ be the collection of all the elements of the traffic demand variables in different layers, where each layer controls a subset of the demand variables and receives network flows

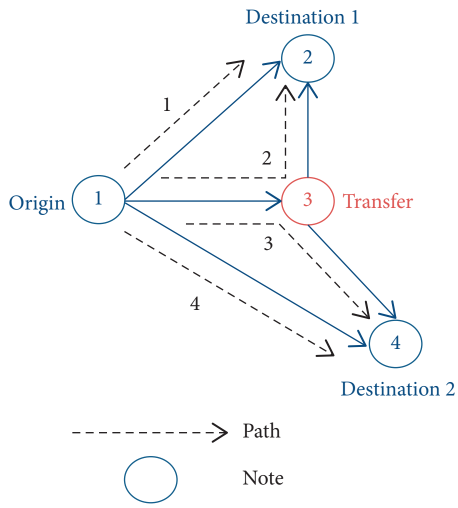


Figure 1: Illustration of the physical urban rail network.

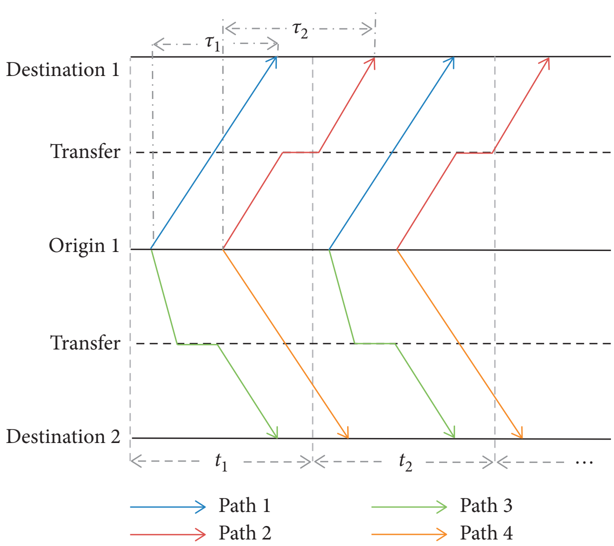


Figure 2: Illustration of the time-space passenger network.

from its upper layers. Let V = $N \cup T \cup D \cup P \cup \Gamma$ be the sets of vertexes arranged in the different layers.

Definition of vertexes (V):

(1) The first layer is the origin station layer, containing each origin station with the index i corresponding to the number of passengers x_i entering (tap-in) the system at origin station i.

(2) The second layer is the departure time layer containing each departure time interval, with the index t corresponding to the number of passengers h_t departing at time t.

(3) The third layer is the destination station layer containing each destination station, with the index j corresponding to the number of passengers h_j exiting (tap-out) the system at destination station j.

(4) The fourth layer is the path layer containing each path, with the index p corresponding to the number of passengers h_p choosing path p.

(5) The five-layer travel time layer contains each path, with the index τ corresponding to the number of passengers τ with travel time τ.

(6) The edges in the graph are defined as $E = E_{\mathrm{NT}} \cup E_{\mathrm{TD}} \cup E_{\mathrm{DP}} \cup E_{\mathrm{P\Gamma}}$ to specify the connections between vertexes.

Definition of edges (E):

(1) E_{NT} contains edges connecting the vertexes in N and T, where each edge corresponds to the proportion of passengers $\alpha_{i,t}$ departing at time t to the passengers entering (tap-in) the system at station i.

(2) E_{TD} contains edges connecting the vertexes in T and D, where each edge corresponds to the proportion of passengers $\beta_{t,j}$ exiting (tap-out) the system at station j to the passengers departing at time t.

(3) E_{DP} contains edges connecting the vertexes in D and P, where each edge corresponds to the proportion of passengers $\rho_{j,p}$ choosing the path p to the passengers exiting (tap-out) the system at station j.

(4) E_{DP} contains edges connecting the vertexes in P and Γ, where each edge corresponds to the proportion of passengers $\omega_{p,\tau}$ with the travel time τ to the passengers choosing path p.

$$H_T = \alpha \times X, \tag{1}$$

$$H_D = \beta \times H_T, \tag{2}$$

$$H_P = \rho \times H_D, \tag{3}$$

$$Y = \omega \times H_p. \tag{4}$$

Equation (1) describes the process of trip production from the origin station layer to the departure time layer. Equation (2) maps the flow from the departure time layer to the destination station layer. Equation (3) maps the flow from an OD pair to the candidate routes. Equation (4) aggregates the path flows to the travel time flows.

3. Model and Solution

We propose a nonlinear programming model with linear constraints for the studied passenger assignment problem. Forward passing in the TDHFN sequentially assigns passengers to candidate stations, paths, and different travel time windows. The network can be improved by backward propagation of the first-order gradients and reassignment of the passenger flow with the updated weights between different layers under the deep optimization framework.

3.1. Optimization Model. We propose a nonlinear programming model with linear constraints for the OD matrix estimation problem. Then, the optimization model is reformulated in the TDHFN for the urban rail system.

3.1.1. Constraints for Passenger Assignment. Assuming the total number of passengers entering the urban rail system at station i is x_i, passengers may depart at station i at each time interval t. Therefore, equation (5) formulates the assignment process, where the passengers in the urban rail system are assigned to each departure time interval t. Equation (6) assigns the passenger flow h_t in departure time interval t to the destination station j as flow h_j. Equation (7) assigns the passenger flow h_j from destination station j to path p as h_p. Equation (8) assigns the passenger flow h_p from path p to the travel time τ as y_τ.

Assigning the departure time intervals:

$$h_t = \sum_i \alpha_{i,t} \times x_i, \quad t \in T. \tag{5}$$

Assigning the destination stations:

$$h_j = \sum_t \beta_{t,j} \times h_t, \quad j \in D. \tag{6}$$

Assigning the paths:

$$h_p = \rho_{j,p} \times h_j, \quad p \in P. \tag{7}$$

Assigning the travel times:

$$y_\tau = \sum_p \omega_{p,\tau} \times h_p, \quad \tau \in \Gamma. \tag{8}$$

3.1.2. Constraints for Flow Equilibrium. The passenger flow equilibrium constraints are shown in equations (9)–(12):

$$\sum_{t\in T_i} \alpha_{i,t} = 1, \tag{9}$$

$$\sum_{j\in D_t} \beta_{t,j} = 1, \tag{10}$$

$$\sum_{p\in P_j} \rho_{j,p} = 1, \tag{11}$$

$$\sum_{\tau\in \Gamma_p} \omega_{p,\tau} = 1. \tag{12}$$

3.1.3. Objective Function. The objective function is shown in the following equation:

$$\min \text{Loss} = \sum_{\tau\in\Gamma} \frac{1}{2}\left(\hat{y}_\tau - y_\tau\right)^2. \tag{13}$$

3.2. BP of Gradient. The Lagrangian functions are as follows:

$$\begin{aligned}
\mathscr{L}\left(\omega_{p,\tau}, h_p, \lambda_\tau\right) &= \sum_{\tau\in\Gamma} \frac{1}{2}\left(\hat{y}_\tau - \sum_p \omega_{p,\tau} \times h_p\right)^2 + \lambda_\tau\left(\sum_{\tau\in\Gamma_p} \omega_{p,\tau} - 1\right),\\
\mathscr{L}\left(\rho_{j,p}, h_j, \lambda_p\right) &= \sum_{\tau\in\Gamma} \frac{1}{2}\left(\hat{y}_\tau - \sum_p \omega_{p,\tau} \times \rho_{j,p} \times h_j\right)^2 + \lambda_p\left(\sum_{p\in P_j} \rho_{j,p} - 1\right),\\
\mathscr{L}\left(\beta_{t,j}, h_t, \lambda_j\right) &= \sum_{\tau\in\Gamma} \frac{1}{2}\left(\hat{y}_\tau - \sum_p \omega_{p,\tau} \times \rho_{j,p} \times \sum_t \beta_{t,j} \times h_t\right)^2 + \lambda_j\left(\sum_{j\in D_t} \beta_{t,j} - 1\right),\\
\mathscr{L}\left(\alpha_{i,t}, x_i, \lambda_t\right) &= \sum_{\tau\in\Gamma} \frac{1}{2}\left(\hat{y}_\tau - \sum_p \omega_{p,\tau} \times \rho_{j,p} \times \sum_t \beta_{t,j} \times \sum_i \alpha_{i,t} \times x_i\right)^2 + \lambda_t\left(\sum_{t\in T_i} \alpha_{i,t} - 1\right).
\end{aligned} \tag{14}$$

Therefore, the gradient of each level based on the KKT conditions is as shown in (15)–(18):

$$g_\tau = \hat{y}_\tau - y_\tau, \tag{15}$$

$$g_p = \sum_{\tau\in\Gamma_p} \omega_{p,\tau} \times \left(\hat{y}_\tau - y_\tau\right), \tag{16}$$

$$g_j = \sum_{p\in P_j} \rho_{j,p} \times \sum_{\tau\in\Gamma_p} \omega_{p,\tau} \times \left(\hat{y}_\tau - y_\tau\right), \tag{17}$$

$$g_t = \sum_{j\in D_t} \beta_{t,j} \times \sum_{p\in P_j} \rho_{j,p} \times \sum_{\tau\in\Gamma_p} \omega_{p,\tau} \times \left(\hat{y}_\tau - y_\tau\right). \tag{18}$$

The Lagrangian multipliers λ are known as the adjoint variables. To compute the gradient, we simply read the gradient concerning $\nabla\mathscr{L} = 0$.

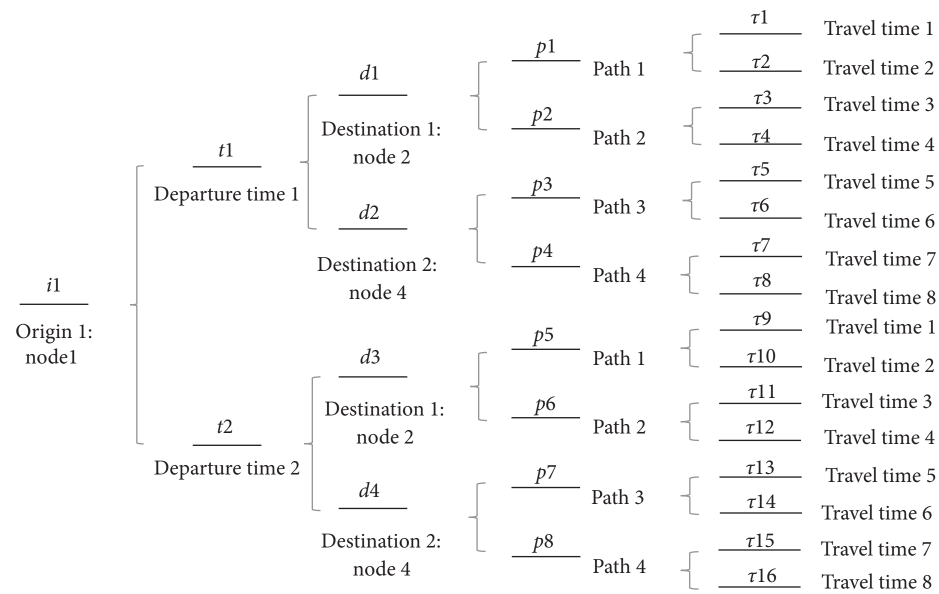


FIGURE 3: Illustration of the time-dependent hierarchical passenger network.

$$\begin{cases} \nabla_{\lambda_\tau} L = \sum_{\tau\in\Gamma} \omega_{p,\tau} - 1 = 0, \\ \nabla_{\lambda_p} L = \sum_{p\in P} \rho_{j,p} - 1 = 0, \\ \nabla_{\lambda_j} L = \sum_{j\in D} \beta_{t,j} - 1 = 0, \\ \nabla_{\lambda_t} L = \sum_{t\in T} \alpha_{i,t} - 1 = 0. \end{cases} \tag{19}$$

3.3. Reformulation in the Deep Optimization Framework. We extend the TDHFN as a computational graph to express the passenger flow assignment process of an urban rail transit system. In the TDHFN, we implement forward passing and backward propagation (BP) to update the estimation variables to approximate the objective functional relationship expressed by (13). As BP is an essential part of the procedure, we use the term BP algorithm to represent the overall procedure throughout this paper.

The model is divided into five layers. The first layer is the input layer, which represents the passenger flow entering the urban rail system by tapping in the card from the origin station; the second layer is the first hidden layer, which represents the passenger flow departing at a certain time; the third layer is the second hidden layer, which represents the passenger flow exiting the system by tapping out the card at the destination station; the fourth layer is the third hidden layer, which represents the passenger flow choosing a certain path; the fifth layer is the output layer, which represents the passenger flow arriving at a certain time. The propagation process of the passenger flow in the network is shown in Figure 4. Through the connection relationships between neurons and the weight of each layer, the passenger volumes of each OD within various time periods can be predicted precisely. At this point, the output layer y_τ represents $y(i, j, t, p, \tau)$. In this paper, to solve the problem conveniently, we proposed a numbering principle (shown in Section 2.2) so that the unique τ can represent (i, j, t, p, τ).

We can calculate many complex marginal values (update values of weights) using the chain rule in calculus, for example,

$$\Delta\omega = \eta \cdot \frac{\partial \text{Loss}}{\partial y_\tau} \cdot \frac{\partial y_\tau}{\partial \omega}, \tag{20}$$

where ω is a dimension vector of partial derivatives. We see that the marginal values consist of calculating a gradient product for each operation in the computational graph. Similarly, the updated formulas for other weights are as follows:

$$\begin{aligned} \Delta\rho &= \eta \cdot \frac{\partial \text{Loss}}{\partial y_\tau} \cdot \frac{\partial y_\tau}{\partial h_j} \cdot \frac{\partial h_j}{\partial \rho}, \\ \Delta\beta &= \eta \cdot \frac{\partial \text{Loss}}{\partial y_\tau} \cdot \frac{\partial y_\tau}{\partial h_j} \cdot \frac{\partial h_j}{\partial h_t} \cdot \frac{\partial h_t}{\partial \beta}, \\ \Delta\alpha &= \eta \cdot \frac{\partial \text{Loss}}{\partial y_\tau} \cdot \frac{\partial y_\tau}{\partial h_j} \cdot \frac{\partial h_j}{\partial h_t} \cdot \frac{\partial h_t}{\partial x_i} \cdot \frac{\partial x_i}{\partial \alpha}. \end{aligned} \tag{21}$$

3.4. Solution Framework. Table 2 shows the solution algorithm for determining the estimation results, including the following three main parts.

3.4.1. Forward Passing. The forward passing step sequentially implements trip generation, trip distribution estimation, and a route-based passenger flow assignment, which can be viewed as a process of the 3-step (from Step 2.1 to Step 2.3) approach in the area of traffic planning.

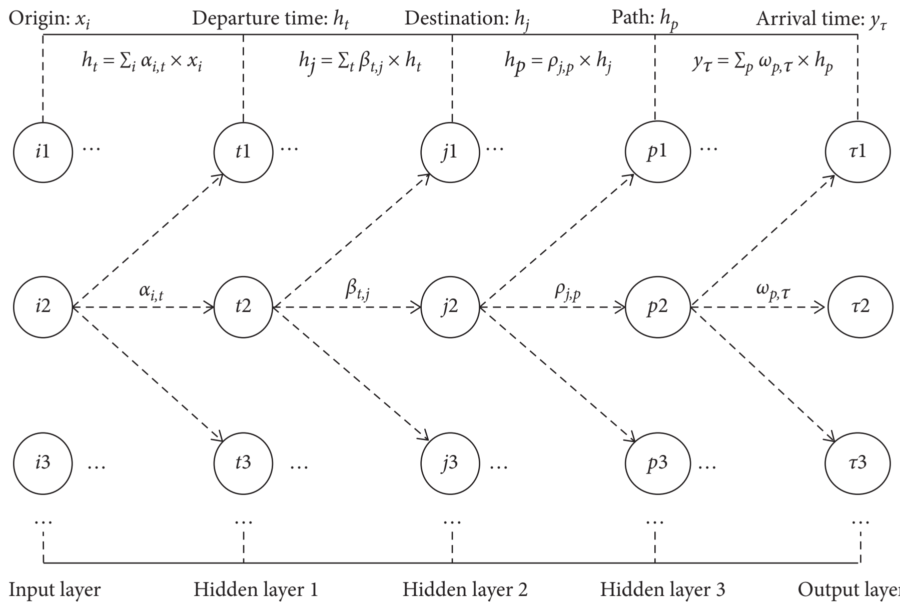


Figure 4: The forward passing process of the TDHFN.

Table 2: Algorithm steps.

Step 1: initialization
Step 2: iterative optimization process
Step 2.1: perform the forward propagations
Based on the fixed passenger flow proportion variables in the multilayer passenger flow network, assign a passenger from the origin station layer to the departure time layer, from the departure time layer to the destination station layer, from the destination station layer to the path layer, and from the path layer to the travel time layer
Step 2.2: calculate the subgradient information
Calculate the subgradient of the passenger flow in the output layer of the multilayer passenger flow network
Step 2.3: set the “error”
Set the “error” of the output layer in the multilayer passenger flow network
Step 2.4: perform the backward error propagations
Perform the backward error propagations in the multilayer passenger flow network from the travel time layer to the path layer, from the path layer to the destination station layer, from the destination station layer to the departure time layer, and from the departure time layer to the origin station layer
Step 2.5: update the auxiliary flow proportion variables
Update the auxiliary passenger flow proportion variables
Step 3: termination
Determine if all the iterations are complete; if not, go back to Step 2

3.4.2. Backward Propagation. The backpropagation step inversely implements feedback control on the forward passing process. Different layers of first-order partial derivatives or “loss errors” are aggregated to calculate the marginal gradients (as shown in Step 2.4).

3.4.3. Update. Update values of variables using gradient descent (as shown in Step 2.5).

4. Numerical Experiments

4.1. Parameter Settings. A partial network of the Beijing Subway system is adopted to verify the proposed predictive model. This portion of the network contains 12 lines (including 6 two-direction lines) and 43 stations, as shown in Figure 5. The research time ranges from 7 am to 9 am, which is the early peak period of the Beijing metro. The AFC record data collected from Sep. 3rd to 7th (from Monday to Friday) in 2018 are utilized to train the model. Then, the data of Sep. 10th (Monday) are adopted for testing. The time intervals are set as 10 min. Accordingly, the passenger flow for each station in the early peak hour is divided into 12 groups.

In this paper, we mainly focus on the OD passenger flow, not the section passenger flow in the subway network. Moreover, the congestion of the route is mainly reflected by the passengers’ travel time, so the passenger flow state of the subway section is not considered. Therefore, we only apply the AFC record of which the origin station and destination station both belong to the partial network of Beijing Subway shown in Figure 5.

In this paper, the travel time is defined as the time range between passengers entering (tap-in) and exiting (tap-out) the station. To facilitate the data statistics, the travel time in this

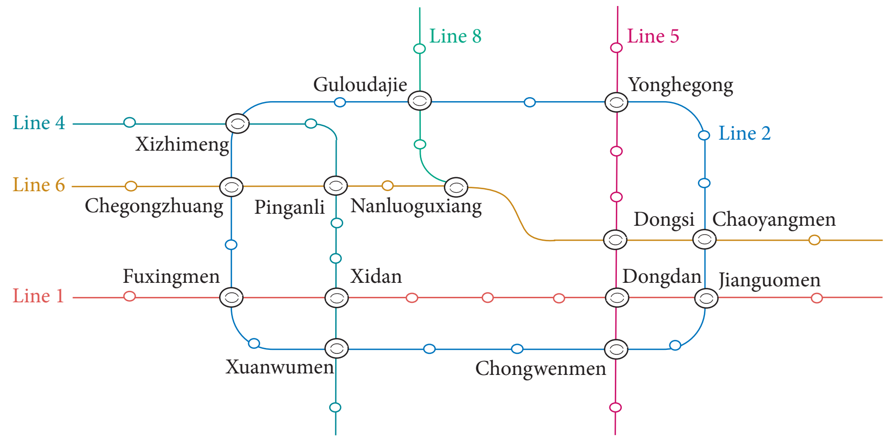


Figure 5: Topology map of the Beijing Subway.

experiment is rounded up to an integer multiple of the time interval (i.e., 10 min). Based on the AFC records, we calculate the travel time of each passenger for each OD. Then, the travel time-frequency distribution histogram of each OD can be obtained. Two examples of the travel time-frequency distributions of the OD from Dongzhimen to Dongdan and the OD from Xizhimen to Xidan are listed in Figures 6(a) and 6(b), respectively. The travel time distribution of each OD is relatively concentrated. In particular, the travel time of more than 90% of the passengers in both of the ODs ranges from 10 to 20 min. In contrast, the proportions of passengers with travel times that are longer than 30 min are less than 1% for the two ODs. Because the frequencies of some travel times are relatively small when constructing the travel time index set Γ, the travel times for which the frequency is less than a specific threshold (e.g., 5%) can be eliminated to reduce the network size. For instance, for the OD from Xizhimen to Xidan, as shown in Figure 6(b), only one index that points to the travel times of 20 min is assembled into the set Γ. The threshold can be adjusted. A smaller threshold of less than 5% can be chosen if a finer resolution is needed.

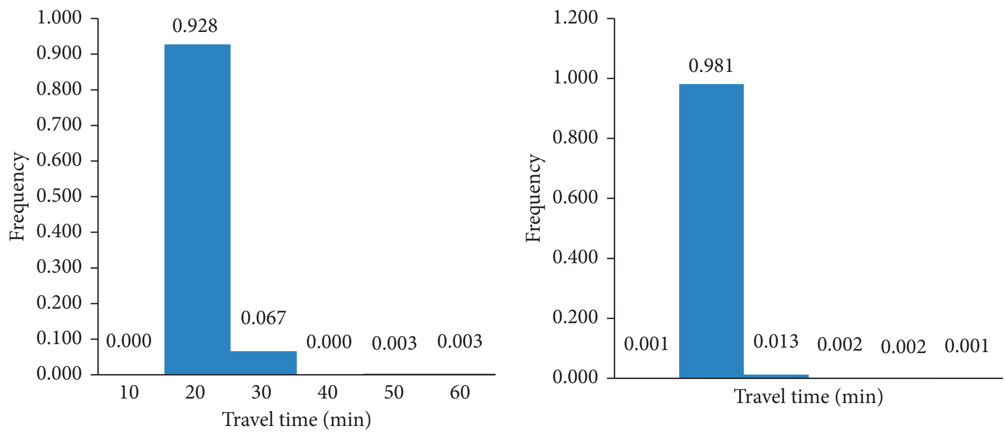
The difference in travel time of each path is due to the path's congestion and individual characteristics of passengers. If a logit model is used to describe the choice probability and behaviors of passengers, the path choice probability is only related to the path cost, which cannot reflect the difference of path's congestion and individual characteristics of passengers. Therefore, we reversely deduce the possible path for passengers based on the real travel time data from AFC and the travel time distribution of each path.

4.2. Result Analysis. We implement the TDHFN using Python 3.6.1, and a part of the Beijing Subway is selected to examine the applicability as well as the computational efficiency of our proposed model. The computational environment is an Intel(R) Core(TM) i5-45900 Processor CPU with 3.30 GHz, 8.00 GB RAM, and 64 bit OS. In addition to TensorFlow, we can use other off-the-shelf software tools, such as Theano, to construct a computation graph-based model.

Extracted from the AFC data, the origin layer has 43 nodes, the departure time layer has 516 nodes, the destination layer has 21,672 nodes, the path layer has 45,732 nodes, and the travel time layer has 39,396 nodes. In this experiment, we let the maximum iterations = 10000 and set the initial learning rate = 0.00001. The iterative curve of the case study is presented in Figure 7, which shows that the loss function can achieve convergence at the 9000th iteration.

To compare the estimated OD passenger flows with the actual passenger flows, we can apply some goodness-of-fit measures, such as the mean absolute percentage error (MAPE), the mean square error (MSE), the root mean square error (RMSE), the root mean square normalized (RMSN) [29], and R-squared. Since we adopted the time-dependent prediction errors in this article, this situation cannot be avoided when the value of OD passenger flow would be zero. Therefore, MAPE is not available because the divisor cannot be zero. RMSE and RMSN measures can be adopted because their divisors would not be zero in this study. But the value of RMSE is related to the value of variables. Therefore, we also adopted the RMSN to compare and show the accuracy of different variables.

The classical function of *RMSE* is presented in equation (22). Besides, RMSE_i in equation (23) represents the measure of the output nodes belonging to the network of which the origin station index is i. RMSE_{ij} in equation (24) represents the measure of the output nodes belonging to the network of which the origin station layer index is i and the destination station index is j. Moreover, RMSE_{ijt} in equation (25) represents the measure of the output nodes belonging to the network of which the origin station layer index is i, the destination station layer index is j, and the departure time layer index is t. In the same vein, the functions of RMSN, RMSN_i, RMSN_{ij}, and RMSN_{ijt} are reported in equations (26)–(29).



(a) (b)

Figure 6: Frequency of travel times. (a) Dongzhimen to Dongdan. (b) Xizhimen to Xidan.

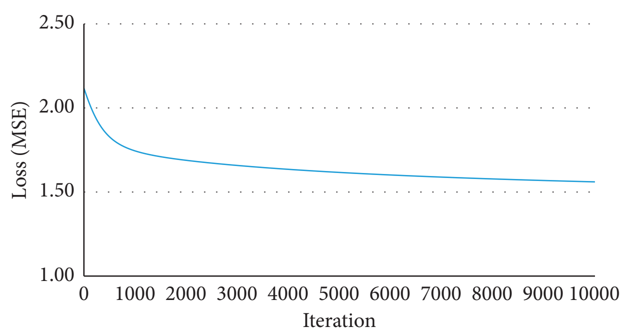


Figure 7: Iterative process.

$$\mathrm{RMSE} = \left(\frac{1}{|\Gamma|}\sum_{\tau\in\Gamma}\left(\hat{y}_\tau - y_\tau\right)^2\right)^{1/2}, \tag{22}$$

$$\mathrm{RMSE}_i = \left(\frac{1}{|T_i|\cdot|D_t|\cdot|P_j|\cdot|\Gamma_p|}\sum_{t\in T_i}\sum_{j\in D_t}\sum_{p\in P_j}\sum_{\tau\in\Gamma_p}\left(\hat{y}_\tau - y_\tau\right)^2\right)^{1/2}, \tag{23}$$

$$\mathrm{RMSE}_{ij} = \left(\frac{1}{|T_{ij}|\cdot|P_j|\cdot|\Gamma_p|}\sum_{t\in T_{ij}}\sum_{p\in P_j}\sum_{\tau\in\Gamma_p}\left(\widehat{y}_\tau - y_\tau\right)^2\right)^{1/2}, \tag{24}$$

$$\mathrm{RMSE}_{ijt} = \left(\frac{1}{\cdot|P_{ijt}|\cdot|\Gamma_p|}\sum_{p\in P_{ijt}}\sum_{\tau\in\Gamma_p}\left(\widehat{y}_\tau - y_\tau\right)^2\right)^{1/2}, \tag{25}$$

$$\mathrm{RMSN} = \frac{\mathrm{RMSE}}{(1/|\Gamma|)\sum_{\tau\in\Gamma}\widehat{y}_\tau}, \tag{26}$$

$$\mathrm{RMSN}_i = \frac{\mathrm{RMSE}_i}{\left(1/\left(|T_i|\cdot|D_t|\cdot|P_j|\cdot|\Gamma_p|\right)\right)\sum_{t\in T_i}\sum_{j\in D_t}\sum_{p\in P_j}\sum_{\tau\in\Gamma_p}\widehat{y}_\tau}, \tag{27}$$

$$\mathrm{RMSN}_{ij} = \frac{\mathrm{RMSE}_{ij}}{\left(1/\left(|T_{ij}|\cdot|P_t|\cdot|\Gamma_P|\right)\right)\sum_{t\in T_{ij}}\sum_{p\in P_j}\sum_{\tau\in\Gamma_p}\widehat{y}_\tau}, \tag{28}$$

Table 3: RMSE and RMSN of the estimation results.

Station index	Origin station	$RMSE_i$	$RMSN_i$ (%)
1	Andelibeijie	0.669	2.71
2	Andingmen	1.890	2.01
3	Beihaibei	0.541	2.57
4	Beijingzhan	2.347	2.61
5	Beixinqiao	0.871	3.07
6	Caishikou	0.831	2.78
7	Changchunjie	5.602	2.52
8	Chaoyangmen	2.002	3.24
9	Chegongzhuang	1.532	2.47
10	Chegongzhuangxi	1.076	2.36
11	Chongwenmen	3.475	2.14
12	Ciqikou	0.747	3.92
13	Dengshikou	0.805	3.42
14	Dongdan	0.864	3.05
15	Dongdaqiao	1.057	2.83
16	Dongsi	0.849	3.13
17	Dongsishitiao	1.644	2.91
18	Dongwuyuan	0.708	2.81
19	Dongzhimen	3.988	2.70
20	Fuchengmen	2.200	2.73
21	Fuxingmen	1.323	3.00
22	Guloudajie	1.738	1.94
23	Hepinglibeijie	1.337	2.19
24	Hepingmen	1.707	1.92
25	Jianguomen	1.627	2.98
26	Jishuitan	4.913	2.17
27	Lingjinghutong	0.692	3.00
28	Nanlishilu	1.111	2.46
29	Nanluoguxiang	0.624	2.66
30	Pinganli	0.905	2.36
31	Qianmen	2.716	2.62
32	Shishahai	0.431	3.23
33	Tiananmendong	0.634	3.65
34	Tiananmenxi	0.482	3.94
35	Wangfujing	0.771	3.82
36	Xidan	0.827	2.81
37	Xinjiekou	1.000	2.84
38	Xisi	0.601	3.11
39	Xizhimen	3.687	2.77
40	Xuanwumen	2.076	2.03
41	Yonganli	1.229	3.16
42	Yonghegong	1.571	2.39
43	Zhangzizhonglu	0.719	2.77
	Avg.	1.545	2.79
	Max.	5.602	3.94
	Min.	0.431	1.92

$$\mathrm{RMSN}_{ijt} = \frac{\mathrm{RMSE}_{ijt}}{\left(1/\left(\cdot\left|P_{ijt}\right|\cdot\left|\Gamma_P\right|\right)\right)\sum_{p\in P_{ijt}}\sum_{\tau\in\Gamma_p}\widehat{y}_\tau}. \tag{29}$$

Table 3 shows the results of RMSE_i and RMSN_i for each station. Several observations can be made:

(1) On the whole, the results for the RMSE_i and RMSN_i of all the stations are relatively low. The average RMSN is below 3%, which indicates that the proposed TDHFN can provide an effective estimation of passenger flow for urban rail systems.

(2) Some stations' RMSN_i are relatively poor, such as those of Tiananmendong, Tiananmenxi, Dongsi, and Ciqikou. The reason for these results may be that these stations are mainly located at famous

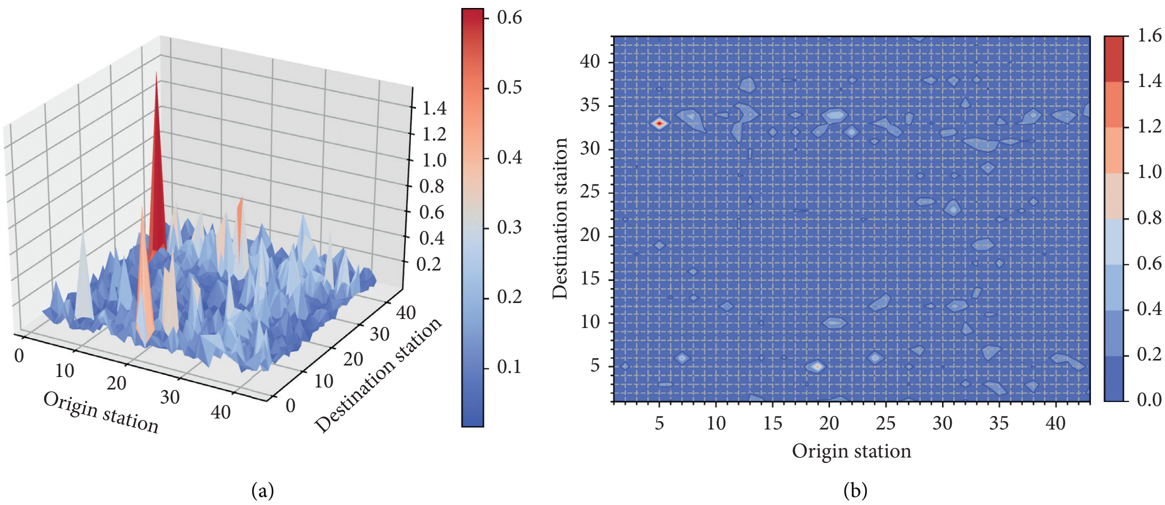


Figure 8: RMSN for the origin-destination matrix. (a) Three-dimensional surface graph of RMSN. (b) Contour graph of RMSN.

scenic spots and shopping mall areas rather than the places where residents live or work. Thus, with morning peak data on working days, the characteristics of the passenger flow in these types of stations cannot be fully captured. In the future, all-day data can be collected to improve the estimation effect.

To explore the estimation results among the passenger ODs, a 3-dimensional surface map of the $RMSN_{ij}$ matrix is shown in Figure 8(a), where the indices of the origin and destination stations are considered as the x-axis and y-axis, respectively, and the $RMSN_{ij}$ value is considered as the z-axis. Besides, the contour line of the $RMSN_{ij}$ matrix from a 2-dimensional perspective is given in Figure 8(b). Note that the station indices in Figure 8 are the same as the indices presented in Table 3.

Furthermore, we produce a 3-dimensional surface map and a contour graph, as shown in Figure 9, for the specific origin station in Chongwenmen. In Figure 9, the departure time, destination station, and $RMSN_{ij}$ values are considered as the x-axis, y-axis, and z-axis, respectively. The definitions of the departure time indices in Figure 9 are given in Table 4.

From the contour graph in Figure 8, we can see that most of the $RMSN_{ij}$ values are relatively small. This result indicates that TDHFN is effective in estimating the OD matrix of urban rail transit passenger flow. However, we can see that there is one point drawn in a dark red color that represents the value of the OD from Tiananmendong to Beijingzhan. The passenger flow between Tiananmendong and Beijingzhan is quite small during the morning peak, which results in a relatively large error.

Most of the points in Figure 9 are drawn with cool colors, which further validates the effectiveness of the proposed method in estimating the time-dependent OD matrix passenger flow. There are few points marked with warm colors, of which the destination stations include Hepingmen, Beijingzhan, etc. In terms of the time dimension, the time range of these data points is mainly concentrated between 7:50 and 8:10.

In addition to the time-dependent OD estimations, the time-dependent travel times for passengers can also be obtained based on the TDHFN method. The results for passengers from Chongwenmen to Changchunjie are illustrated in Figure 10, where the estimated and actual time-dependent travel time distributions are presented. The fluctuation trend of the estimated values is consistent with the trend of the actual values, which shows the effectiveness of the proposed method in travel time estimation.

4.3. Comparative Analysis. The estimation results of TDHFN are compared with the results of an artificial neural network (ANN). For a detailed introduction of the ANN method, we refer to the literature by Remya and Mathew [20] and Mozolin et al. [24]. The eigenvalues selected in this paper are obtained from AFC data and urban rail network topology, including the daily average passenger flow of the origin station, the daily average passenger flow of the destination station, the number of alternative paths, the average travel time, the distance (replaced by section number), the departure time, and the average transfer times. After training and adjusting, we got a well-trained ANN model. There are three layers in the network, including the input layer, the output layer, and one hidden layer. The activation function is Relu and Sigmoid, and the number of hidden layer nodes is 5.

The comparison results are illustrated in Figure 11 and Table 5, which show that the results of the model proposed in this paper are significantly better than those of the ANN. However, it should be noted that the source of the input data for ANN is the same as that of the TDHFN model. The performance of the ANN method can be improved when additional data are collected, such as commuter numbers, commuter properties, and land types. However, in a

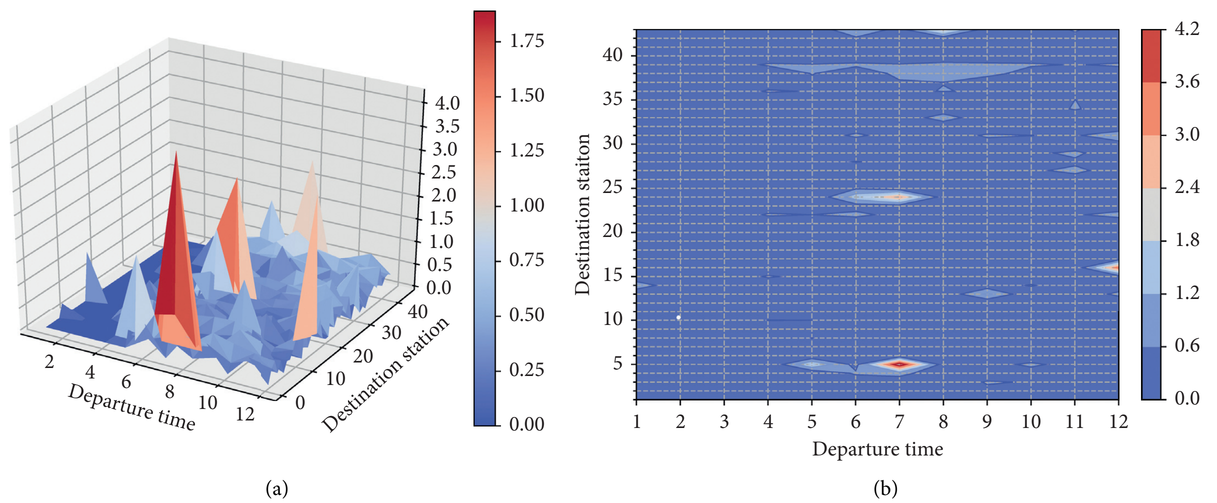


Figure 9: RMSN from Chongwenmen Station. (a) Three-dimensional surface graph of RMSN. (b) Contour graph of RMSN.

Table 4: Departure time index.

Index	Departure time
1	7:00
2	7:10
3	7:20
4	7:30
5	7:40
6	7:50
7	8:00
8	8:10
9	8:20
10	8:30
11	8:40
12	8:50

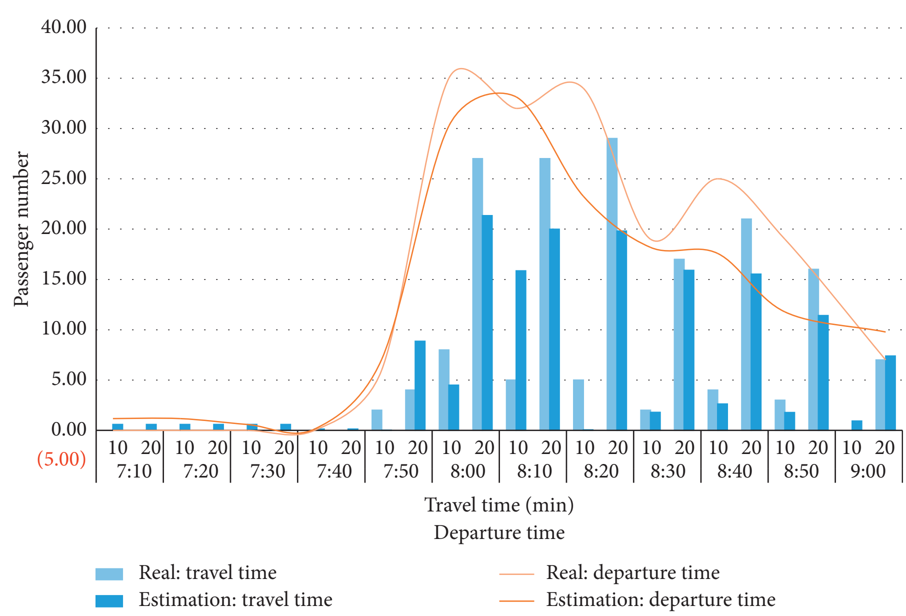


Figure 10: Estimation results from Chongwenmen to Changchunjie.

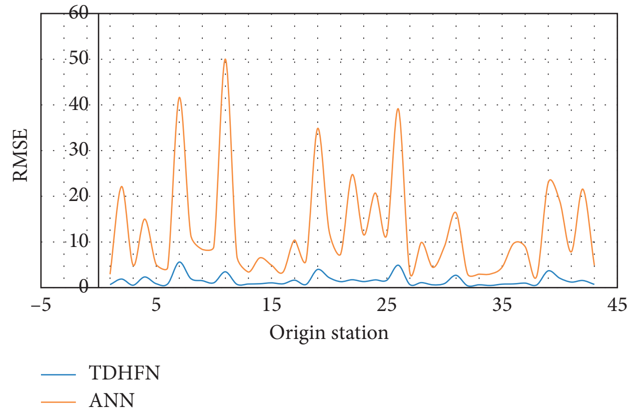


Figure 11: Comparative analysis with the ANN method.

Table 5: RMSE and RMSN of TDHFN compared with the ANN method.

Error	TDHFN	ANN
RMSE	257.49	2236.63
RMSN (%)	0.54	4.66

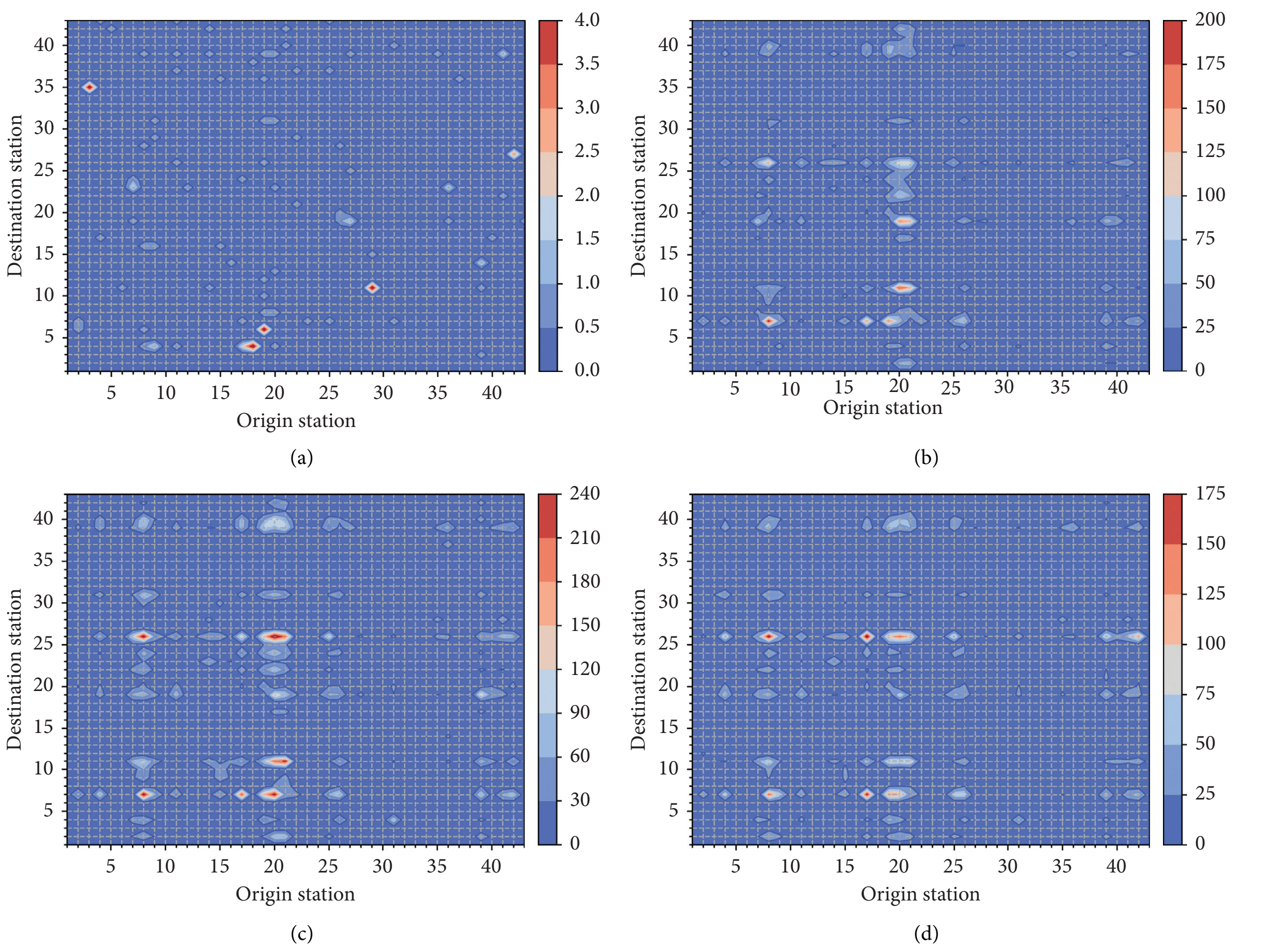


Figure 12: Dynamic OD matrix estimation of passenger flow. (a) OD passenger volume from 7:00 to 7:30. (b) OD passenger volume from 7:30 to 8:00. (c) OD passenger volume from 8:00 to 8:30. (d) OD passenger volume from 8:30 to 9:00.

practical situation, a detailed and comprehensive collection is difficult.

The difference between the ANN and computational graph algorithm is that the former neural network is a black-box model, and the number of neurons, activation functions, and neural network layers is not certain, so this method often requires continuous experiments and adjustments to find the optimal model. However, in TDHFN, the number of neurons, the form of the activation function, and the number of layers of the neural network are determined values with practical physical significance. Only the weight matrix of each layer in the network is unknown and needs to be determined through learning. Therefore, the computational graph can express more traffic characteristics than the ANN and provides an effective and hierarchical passenger flow estimation.

Finally, the dynamic OD matrix estimation of passenger flow is shown in Figure 12. It shows the passenger flow changes of each OD in different periods. The dynamic OD matrix estimation of passenger flow can provide basic data for the passenger flow control strategy of urban rail transit.

5. Conclusions

This study proposed a time-dependent hierarchical flow network for urban rail transit passengers. The OD passenger flow matrix at each time in the subway network can be obtained by inputting the incoming passenger volume of each station during the morning peak to the model. This model can be improved by backpropagation of the first-order gradients and reassignment of the passenger flow with the updated weights between different layers under the deep optimization framework. The result analysis indicates that the TDHFN can provide abundant and hierarchical passenger flow estimation results. A comparative analysis shows that the proposed model can effectively obtain relatively accurate passenger flow estimation results.

At present, the existing OD dynamic estimation methods of urban rail network passenger flow still need to be improved in the factors of timeliness and accuracy. The most important contribution of this paper is to propose a multilayer hierarchical flow network applied to urban rail with deep learning research. This method can solve the dynamic OD matrix estimation problem. This flow-oriented prediction formulation can simultaneously estimate different levels of unobserved or partially observed passenger flow variables. Furthermore, when more data sources are available, this method can achieve hierarchical expansion, making this method more flexible. To build a theoretically sound modeling framework, this paper hopes to trace back to the fundamentals or low-level representation of deep learning networks and construct a transportation-focused computational graph as a structured modeling language. This modeling paradigm enables us to capture the mathematical structure inside the OD matrix estimation problem by representing and decomposing complex composite functions through a graph of current states and numerical gradients.

However, the model proposed in this study does not apply to all stations. The model function is better when the subway stations are mainly the distribution of the places where residents live or work. By only using the data of the morning peaks over a few working days, we cannot determine the characteristics of passenger flow through training. In the future, more comprehensive data should be collected, such as GPS trajectory data [16], land-use data, or the (point of interest) POI features [30]. Tang et al. [31] applied to uncover the characteristics of travel patterns from temporal and spatial dimensions in the metro network according to the POI data. Based on their study, the stations can be clustered by node significance on the metro network or POI features of stations. Thus, the applicability of this model may be improved.

Data Availability

The numerical data used to support the findings of this study are available from the corresponding author upon request.

Disclosure

The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The authors wish to thank Jingjia Cao, Qingying Lai, Feiran Liu, Xu Xu, and Linqi Xia for helpful discussions. This research was funded in part by the National Natural Science Foundation of China under grant nos. 71871012 and 72001020, in part by the State Key Lab of Rail Traffic Control and Safety of China under grant no. RCS2020ZT003, in part by the China Postdoctoral Science Foundation under grant no. 2020M670128, and in part by the Beijing Municipal Natural Science Foundation under grant no. L181007.

References

[1] X.-M. Yao, P. Zhao, and D.-D. Yu, “Real-time origin-destination matrices estimation for urban rail transit network based on structural state-space model,” *Journal of Central South University*, vol. 22, no. 11, pp. 4498–4506, 2015.

[2] J. de D. Ortuzar and L. G. Willumsen, *Modeling Transport*, John Wiley & Sons, New York, NY, USA, 1994.

[3] S. Kikuchi and N. Kronprasert, “Constructing a transit origin-destination table using the uncertainty maximization concept,” *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2112, no. 1, pp. 43–52, 2009.

[4] M. V. Aerde, H. Rakha, and H. Paramahamsan, “Estimation of origin-destination matrices: relationship between practical and theoretical considerations,” *Transportation Research Record: Journal of the Transportation Research Board*, vol. 1831, no. 1, pp. 122–130, 2003.

[5] L. Caggiani, M. Ottomanelli, and D. Sassanelli, “A fixed point approach to origin-destination matrices estimation using

uncertain data and fuzzy programming on congested networks," *Transportation Research Part C: Emerging Technologies*, vol. 28, pp. 130–141, 2013.

[6] X.-F. Yang, Y. Lu, and W. Hao, "Origin-destination estimation using probe vehicle trajectory and link counts," *Journal of Advanced Transportation*, vol. 2017, Article ID 4341532, 2017.

[7] L. Cheng, S.-L. Zhu, Z.-M. Chu, and J.-X. Cheng, "A bayesian network model for origin-destination matrices estimation using prior and some observed link flows," *Discrete Dynamics in Nature and Society*, vol. 2014, p. 9, Article ID 192470, 2014.

[8] Y.-X. Ji, J.-Z. Zhao, Z.-M. Zhang, and Y.-C. Du, "Estimating bus loads and OD flows using location-stamped farebox and Wi-Fi signal data," *Journal of Advanced Transportation*, vol. 2017, Article ID 6374858, 2017.

[9] S. Bera and K. V. K. Rao, "Estimation of origin-destination matrix from traffic counts: the state of the art," *European Transport\ Trasporti Europei*, vol. 49, no. 49, pp. 3–23, 2011.

[10] H.-P. Lo and C.-P. Chan, "Simultaneous estimation of an origin-destination matrix and link choice proportions using traffic counts," *Transportation Research Part A: Policy and Practice*, vol. 37, no. 9, pp. 771–788, 2003.

[11] H. Yang, "Heuristic algorithms for the bilevel origin-destination matrix estimation problem," *Transportation Research Part B: Methodological*, vol. 29, no. 4, pp. 231–242, 1995.

[12] X. Zhou and H. S. Mahmassani, "Dynamic origin-destination demand estimation using automatic vehicle identification data," *IEEE Transactions on Intelligent Transportation Systems*, vol. 7, no. 1, pp. 105–114, 2006.

[13] H. D. Sherali, R. Sivanandan, and A. G. Hobeika, "A linear programming approach for synthesizing origin-destination trip tables from link traffic volumes," *Transportation Research Part B: Methodological*, vol. 28, no. 3, pp. 213–233, 1994.

[14] T. Toledo and T. Kolechkina, "Estimation of dynamic origin-destination matrices using linear assignment matrix approximations," *IEEE Transactions on Intelligent Transportation Systems*, vol. 14, no. 2, pp. 618–626, 2013.

[15] M. Fujita, S. Yamada, and S. Murakami, "Time coefficient estimation for hourly origin-destination demand from observed link flow based on semidynamic traffic assignment," *Journal of Advanced Transportation*, vol. 2017, Article ID 6495861, 2017.

[16] J. Tang, S. Zhang, X. Chen, F. Liu, and Y. Zou, "Taxi trips distribution modeling based on entropy-maximizing theory: a case study in Harbin city-China," *Physica A: Statistical Mechanics and Its Applications*, vol. 493, pp. 430–443, 2018.

[17] W. Rao, Y. Wu, J. Xia, J. Ou, and R. Kluger, "Origin-destination pattern estimation based on trajectory reconstruction using automatic license plate recognition data," *Transportation Research Part C*, vol. 95, 2017.

[18] J. Liu, F. Zheng, H. J. Van Zuylen, and J. Li, "A dynamic OD prediction approach for urban networks based on A dynamic OD prediction approach for urban networks based on automatic number plate recognition data automatic number plate recognition data," *Transportation Research Procedia*, vol. 47, 2019.

[19] J. Guo, Y. Liu, X. Li, W. Huang, J. Cao, and Y. Wei, "Enhanced least square based dynamic OD matrix estimation using Radio Frequency Identification data," *Mathematics and Computers in Simulation*, vol. 155, pp. 27–40, 2019.

[20] K. P. Remya and S. Mathew, "OD matrix estimation from link counts using artificial neural network," *International Journal of Scientific & Engineering Research*, vol. 4, no. 5, pp. 293–296, 2013.

[21] Z. Gong, "Estimating the urban o-d matrix: a neural network approach," *European Journal of Operational Research*, vol. 106, pp. 108–115, 1998.

[22] H. Yang, T. Akiyama, and T. Sasaki, "Estimation of time-varying origin-destination flows from traffic counts: a neural network approach," *Mathematical and Computer Modelling*, vol. 27, no. 9–11, pp. 323–334, 1998.

[23] H.-Y. Li, Y.-T. Wang, X.-Y. Xu, L.-Q. Qin, and H.-Y. Zhang, "Short-term passenger flow prediction under passenger flow control using a dynamic radial basis function network," *Applied Soft Computing*, vol. 83, 2019.

[24] M. Mozolin, J.-C. Thill, and E. Lynn Usery, "Trip distribution forecasting with multilayer perceptron neural networks: a critical evaluation," *Transportation Research Part B: Methodological*, vol. 34, no. 1, pp. 53–73, 2000.

[25] P. Shang, R.-M. Li, Z.-Y. Liu, and K. Xian, "Timetable synchronization and optimization considering time-dependent passenger demand in an urban subway network," *Transportation Research Record*, vol. 34, 2018.

[26] P. Shang, R. Li, J. Guo, K. Xian, and X. Zhou, "Integrating Lagrangian and Eulerian observations for passenger flow state estimation in an urban rail transit network: a space-time-state hyper network-based assignment approach," *Transportation Research Part B: Methodological*, vol. 121, pp. 135–167, 2019.

[27] X. Wu, J. Guo, K. Xian, and X. Zhou, "Hierarchical travel demand estimation using multiple data sources: a forward and backward propagation algorithmic framework on a layered computational graph," *Transportation Research Part C: Emerging Technologies*, vol. 96, pp. 321–346, 2018.

[28] K. Lu, A. Khani, and B. Han, "A trip purpose-based data driven alighting station choice model using transit smart card data," *Complexity*, vol. 2018, Article ID 3412070, 2018.

[29] R. Frederix, F. Viti, W. Himpe, and C. Tampère, "A hierarchical approach for dynamic origin-destination matrix estimation on large-scale congested networks," *International IEEE Conference on ITS*, vol. 53, 2011.

[30] Z.-Y. Du, J.-J. Tang, Y. Qi, Y.-W. Wang, C.-Y. Han, and Y.-F. Yang, "Identifying critical nodes in metro network considering topological potential: a case study in shenzhen city-China," *Physica A*, vol. 539, pp. 1–13, 2020.

[31] J.-J. Tang, X.-L. Wang, F. Zong, and Z. Hu, "Uncovering spatio-temporal travel patterns using a tensor-based model from metro smart card data in Shenzhen, China," *Sustainability*, vol. 12, no. 1475, pp. 1–16, 2020.