Transactions on Environment, Energy and Earth Sciences
ISSN: 2960-1924, eISSN: 2960-2300 | Volume 7
GTREE 2026

Passenger Flow Prediction for Urban Rail Transit Transfer Stations Based on Multi-Scale Dual-Graph Spatial-Temporal Fusion Network

Han Zhuang *, Jinbing Ha

Nanjing University of Science and Technology, Nanjing, China

* Corresponding Author Email: 13773576802@163.com

Abstract. To improve the accuracy of short-term passenger flow prediction at urban rail transit transfer stations, this paper proposes a Multi-Scale Dual-Graph Spatial-Temporal Fusion Network (MS-DSTGFN). First, AFC data and network topology are integrated to identify transfer events through travel-time decomposition and a two-stage path search method, thereby constructing inbound, outbound, and transfer passenger-flow time series. Then, a physical topology graph and a data-driven adaptive graph are jointly introduced into the forecasting model, and a dual-graph hybrid propagation mechanism is designed to capture both explicit and implicit spatial dependencies. On the temporal side, three encoders are built for recent, daily-periodic, and weekly-periodic sequences, and a cross-scale attention module is adopted to fuse multi-scale temporal patterns adaptively. Finally, a channel-aware decoder is used for differentiated joint prediction of the three passenger-flow types. Experiments on Hangzhou Metro AFC data show that MS-DSTGFN outperforms all baseline models. Compared with the best-performing baseline DCRNN, MAE, RMSE, and MAPE are reduced by 8.88%, 6.58%, and 10.06%, respectively. Ablation results further confirm the effectiveness of the dual-graph fusion, cross-scale fusion, and channel-aware decoding modules.

Keywords: Urban rail transit; transfer station; transfer passenger flow extraction; short-term passenger flow prediction; spatial-temporal graph neural network; multi-scale fusion.

1. Introduction

Urban rail transit has become the backbone of urban public transportation due to its large capacity, high punctuality, and environmental friendliness. Transfer stations, as key nodes connecting multiple lines, directly affect the operational efficiency of the entire network and passenger travel experience. Accurate short-term passenger flow prediction is of great significance for capacity allocation, station organization optimization, and emergency dispatch.

Rail transit passenger flow forecasting methods have evolved from traditional statistical models to deep learning and spatial-temporal joint modeling. Early studies mostly used time series methods such as ARIMA and SARIMA to fit passenger flow under stationarity assumptions, but their ability to capture complex nonlinear patterns is limited [1-2]. Subsequently, machine learning methods such as wavelet decomposition and support vector machines were applied to establish nonlinear mappings, improving forecasting performance to some extent [3-4]. In recent years, recurrent neural networks such as LSTM and GRU have become mainstream tools due to their ability to capture long-term and short-term dependencies [5-6]; Transformer models based on self-attention mechanisms have also begun to be applied to passenger flow forecasting, showing advantages in global dependency modeling and parallel computing [7-9].

However, modeling only from the temporal dimension cannot fully reflect the spatial dependencies formed by physical connections and passenger flow propagation in rail transit networks. To address this issue, researchers introduced graph convolution into passenger flow forecasting and proposed spatial-temporal graph neural network models such as STGCN, DCRNN, and Graph WaveNet [10-12], significantly improving the expressive power for networked passenger flow evolution patterns. On this basis, attention mechanisms and adaptive adjacency matrices further enhanced the model's ability to characterize heterogeneous spatial associations [13-14]. In recent years, dynamic hybrid

Content from this work may be used under the terms of CC BY-NC 4.0 licence (https://creativecommons.org/licenses/by-nc/4.0/).
Published by Warwick Evans Publishing.

spatial-temporal graph convolution and adaptive diffusion graph convolution have been further developed, helping to mine more complex station relationships [15-16].

Compared with network-wide station passenger flow prediction, research specifically targeting transfer stations is relatively scarce. Existing work mainly focuses on two aspects: one is indirect identification and statistical analysis of transfer passenger flow based on AFC data and network topology [17-18]; the other is forecasting modeling for transfer station passenger flow sequences, mostly using decomposition strategies such as STL, EMD, and CEEMDAN combined with models such as GRU and LSTM to improve fitting ability for non-stationary sequences [19-20].

However, most of the above studies treat transfer stations as isolated time series nodes, relying only on historical data of the station itself for modeling, without fully incorporating network-wide inter-station spatial dependencies. Although graph neural networks have achieved inter-station information propagation through physical topology graphs, when facing multi-flow forecasting tasks for transfer stations, there is a lack of exploration of potential passenger flow associations beyond fixed physical adjacency. Meanwhile, rail transit passenger flow exhibits multiple temporal patterns such as short-term fluctuations, daily periodicity, and weekly periodicity; effective integration of multi-scale information also helps improve forecasting accuracy.

To address the above issues, this paper proposes the MS-DSTGFN model for joint prediction of inbound, outbound, and transfer passenger flow at transfer stations. At the data level, AFC card data and network topology information are integrated, and transfer passenger flow is extracted through travel time decomposition and path search to construct inbound, outbound, and transfer passenger flow time series. At the model level, both a physical topology graph and a data-driven adaptive graph are introduced simultaneously, and a dual-graph hybrid propagation mechanism is designed to account for both explicit structural constraints of the network and potential spatial associations. Spatial-temporal encoders are built for three time scales: recent, daily periodic, and weekly periodic, and a cross-scale attention module is used to achieve adaptive fusion of multi-scale temporal information. Finally, multi-step future predictions of three types of passenger flow at transfer stations are performed simultaneously, improving forecasting accuracy and stability in key node scenarios.

2. Problem Definition

The problem of multi-step prediction for three categories of passenger flows at urban rail transit transfer stations is defined as follows: to simultaneously predict the inbound, outbound, and transfer passenger flows of transfer stations. Suppose the rail network contains N stations, the number of transfer stations is M, and the time granularity is $\triangle t$(15 min). The passenger flow state of transfer station v at time t is expressed as:

$$\boldsymbol{x}_v^t = [x_v^{t,\text{in}}, x_v^{t,\text{out}}, x_v^{t,\text{trans}}]^\top \tag{1}$$

Where $x_v^{t,in}, x_v^{t,out}$, $x_v^{t,trans}$ are the inbound, outbound and transfer passenger volumes of transfer station v at time t, respectively. The passenger-flow state of the whole network is denoted as$X_t \in \mathbb{R}^{N\times 3}$.

Three-scale historical sequences, namely the recent-reference sequence, daily-reference sequence and weekly-reference sequence, are denoted as $\boldsymbol{X}^{(r)}, \boldsymbol{X}^{(d)}$ and $\boldsymbol{X}^{(w)}$ respectively. The prediction task can be formalized as: given multi-scale historical passenger-flow sequences, temporal context features T, static attributes of stations S and adjacency matrices A, learn a mapping function f to realize the synchronous prediction of three types of passengers flows at transfer stations over the K time steps:

$$\widehat{\boldsymbol{Y}}_{t+1:t+K} = f\left(\boldsymbol{X}^{(r)}, \boldsymbol{X}^{(d)}, \boldsymbol{X}^{(w)}, \boldsymbol{T}, \boldsymbol{S}, \boldsymbol{A}; \Theta\right) \tag{2}$$

Where Θrepresents model parameters, and $\hat{Y}$ is the predicted value.

3. Transfer Passenger Flow Extraction Method

The automatic fare collection (AFC) system only records passengers' entry and exit transaction records and cannot directly capture in-station transfer behaviors. To derive transfer passenger flow sequences, AFC records are integrated with network topology. Transfer events are reconstructed via travel-time decomposition and path inference, and transfer passenger flows are aggregated within predefined time windows.

The metro network is abstracted as a weighted directed graph$G = (V, E)$, where V denotes the set of stations, E represents directed links between adjacent stations, and edge weights correspond to train travel durations between adjacent stations. The total travel time for a single trip is decomposed as:

$$T_{\text{obs}} = T_{\text{run}}(o, d) + t_{\text{in}} + t_{\text{out}} + \sum_{k \in \Phi} P_k \tag{3}$$

Where $T_{\text{run}}(o, d)$ is the theoretical train running time between an origin-destination (OD) pair, which is computed using the Dijkstra algorithm; t_{in} and t_{out} represent the comprehensive entry time at the origin station (including post-swipe walking and waiting durations) and the comprehensive exit time at the destination station, respectively; P_k is the transfer impedance at transfer station k (encompassing transfer-corridor walking time and waiting time at intermediate stations); Φdenotes the set of transfer stations along the corresponding path.

OD samples without transfers are selected to construct an over-determined system of equations. The least-squares method is employed to estimate the comprehensive entry and exit time parameters for each station. For OD records passing through exactly one transfer station, the transfer impedance is estimated using the median of residual travel durations after subtracting theoretical running time and calibrated entry-exit time. Upon parameter calibration, direct paths are prioritized for intra-line OD pairs. For inter-line OD pairs, global path search is conducted with the sum of segment travel time and transfer impedance treated as path cost.

The arrival time at a transfer station is inferred as follows:

$$t_{\text{arr}}(k) = t_{\text{in}} + t_{\text{in}} + \lambda_k \cdot T_{\text{net}} \tag{4}$$

Where t_{in} is the card-swiping entry time; t_{in} is the comprehensive entry time at the origin station; λ_k is the ratio of the cumulative elapsed time up to the transfer station along the theoretical path to the total theoretical running time of the whole path; T_{net} is the net running time within the network after eliminating entry-exit time consumption.

Transfer events are assigned to 15-minute time windows according to the inferred arrival time to obtain the transfer passenger flow sequence of each transfer station. The reconstruction results are further examined through spatiotemporal consistency and Bootstrap resampling to ensure path reachability, temporal rationality and statistical stability.

4. Model Construction

Figure 1 shows the overall framework of MS-DSTGFN. The model first constructs historical inputs on the recent, daily-reference, and weekly-reference time scales, and extracts multi-scale spatiotemporal representations through a spatiotemporal encoder with shared parameters. It then fuses features of different time scales through a cross-scale attention module, and finally realizes multi-step joint prediction of three types of passenger flows at transfer stations by means of a channel-aware decoder.

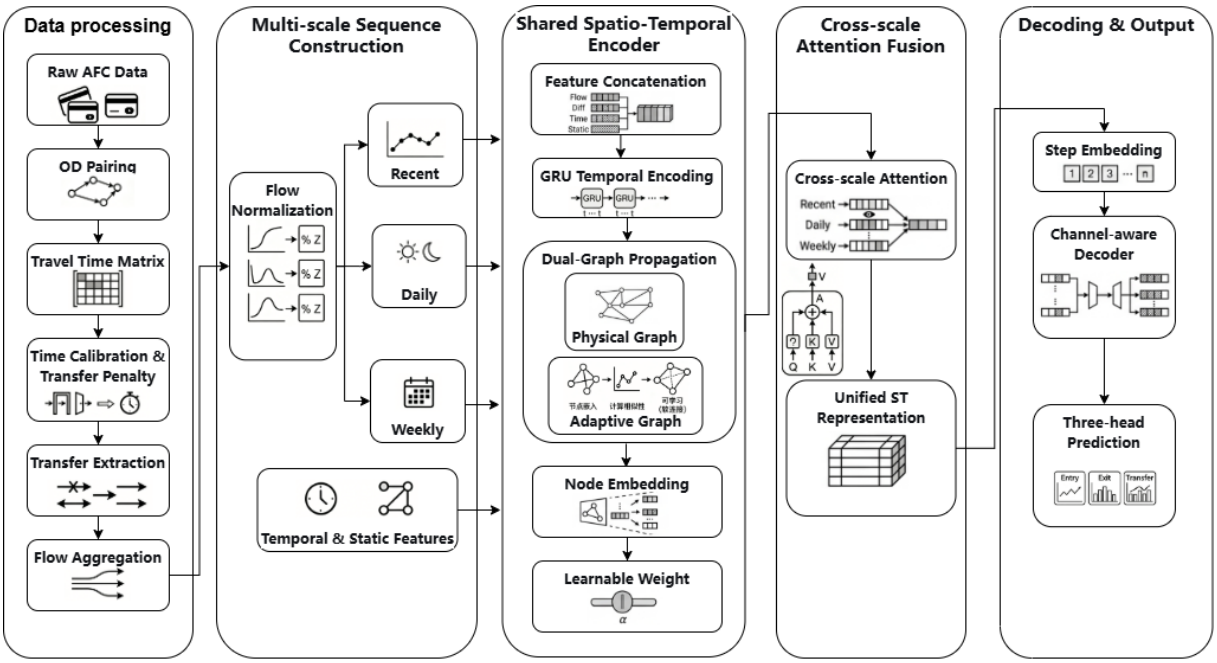


Fig 1. Figure of the MS-DSTGFN model framework.

4.1. Dual-Graph Spatial Propagation Module

The physical topology graph $\boldsymbol{A}_{\text{topo}} \in \mathbb{R}^{N\times N}$ is constructed based on the network connection relationships to reflect the actual connectivity structure among stations. The adaptive graph $\boldsymbol{A}_{\text{adp}} \in \mathbb{R}^{N\times N}$ calculates the correlation strength between stations through learnable node embeddings$\boldsymbol{e}_i$:

$$\boldsymbol{A}_{\text{adp}}(i,j) = \frac{exp(-\|\boldsymbol{e}_i - \boldsymbol{e}_j\|^2/\tau)}{\sum_{k=1}^{N\Sigma_{ik}{}^2} exp} \tag{5}$$

Where τ denotes the temperature coefficient controlling the smoothness of attention distribution.

Let the input node feature be $\boldsymbol{H} \in \mathbb{R}^{B\times N\times d}$ (B represents the batch size). For an arbitrary graph A, the spatial propagation process is formulated as:

$$\tilde{H}(\boldsymbol{A}) = \phi([\boldsymbol{H}, \boldsymbol{A}\boldsymbol{H}]) + \boldsymbol{H} \tag{6}$$

Where $[\cdot,\cdot]$ refers to concatenation along the feature dimension; $\phi(\cdot)$ stands for a two-layer nonlinear mapping consisting of fully connected layers, GELU activation function and Dropout.

Let the propagation outputs derived from the physical topology graph and adaptive graph be $\boldsymbol{H}_{\text{topo}}$ and $\boldsymbol{H}_{\text{adp}}$, respectively. A learnable coefficient is introduced ,by the sigmoid function (initialized to 0.5), for weighted fusion:

$$\boldsymbol{H}_{\text{spatial}} = \alpha\boldsymbol{H}_{\text{topo}} + (1-\alpha)\boldsymbol{H}_{\text{adp}} \tag{7}$$

4.2. Spatiotemporal Joint Encoder

Encoders with shared parameters are adopted for the three temporal scales (recent, daily reference, and weekly reference). Let the input of a certain scale be $\boldsymbol{X} \in \mathbb{R}^{B\times L\times N\times 3}$ (T denotes the length of time window), the temporal feature be $\boldsymbol{T} \in \mathbb{R}^{B\times L\times 5}$, and the static attribute of stations be $\boldsymbol{S} \in \mathbb{R}^{N\times 2}$.

Historical passenger flow, differential features, temporal embeddings and static attributes are concatenated and projected into the latent space:

$$\widetilde{\boldsymbol{X}} \in \mathbb{R}^{B\times L\times N\times d} \tag{8}$$

Subsequently, $\tilde{X}$ is rearranged along the temporal dimension into $[B \times N, L, d]$. Taking each station as an independent sequence, it is fed into a single-layer GRU to extract temporal representations. On this basis, the temporal representations are fed into the dual-graph spatial propagation module, and the spatiotemporal encoding outputs H_r, H_d and H_w under this scale are obtained for the cross-scale feature fusion module.

4.3. Cross-Scale Feature Fusion Module

Learnable scale embedding vectors $\boldsymbol{e}_r, \boldsymbol{e}_d, \boldsymbol{e}_w \in \mathbb{R}^d$ are introduced for different scales and applied uniformly:

$$\widetilde{\boldsymbol{H}}_s = \boldsymbol{H}_s + \boldsymbol{e}_s, s \in \{r, d, w\} \tag{9}$$

The embedded representation of the recent branch is taken as the query vector, while representations of all three scales serve as keys and values. The multi-head attention mechanism (the number of heads is set to 4) is utilized to calculate the cross-scale aggregated output. Residual connection, layer normalization and feed-forward network are further adopted to obtain the fused representation $H_{fused} \in \mathbb{R}^{B\times N\times d}$, which acts as the unified spatiotemporal representation for all stations across the network.

4.4. Channel-Aware Multi-Step Decoding

Target-node representations are extracted from the fused representations$\boldsymbol{H}_{\text{fused}}$:

$$\boldsymbol{H}_M = \boldsymbol{H}_{\text{fused}}[:, \mathcal{I}_M, :] \in \mathbb{R}^{B\times M\times d} \tag{10}$$

Where $M = 5$ denotes the number of transfer stations, and $\mathcal{I}_M$ is the index set of transfer stations. To distinguish different prediction steps, learnable step-wise embedding vectors $\boldsymbol{p}_k \in \mathbb{R}^d$ are assigned for each of the future K time steps. The hidden state for the *k*-th step is given by:

$$\boldsymbol{H}^{(k)} = \boldsymbol{H}_M + \boldsymbol{p}_k \tag{11}$$

Further, a modulation mechanism is adopted on the channel dimension to generate differentiated representations for inbound, outbound and transfer passenger flows:

$$\boldsymbol{H}_c^{(k)} = \gamma_c^{(k)} \odot \boldsymbol{H}^{(k)} + \beta_c^{(k)}, c \in \{\text{in}, out, trans\} \tag{12}$$

Where $\odot$ denotes the element-wise product; $\gamma_c^{(k)}$ and $\beta_c^{(k)}$ are learnable scaling and shift parameters, respectively, which are used to adjust the representation scales for different passenger-flow types. Finally, independent prediction heads are configured for each step and each channel to output the predicted values $\hat{Y} \in \mathbb{R}^{B\times K\times M\times 3}$ over the future K time steps.

5. Experiments and Analysis

5.1. Dataset and Experimental Setup

The AFC smart card data of Hangzhou Metro from January 1 to January 26, 2019 is adopted, covering a total of 79 stations on Line 1, Line 2 and Line 4. After passenger flow extraction, three types of passenger flow time series are constructed for 5 transfer stations, with a temporal granularity of 15 min and 72 time slices per day. The dataset is chronologically divided into a training set (the first 70%), a validation set (the middle 15%) and a test set (the last 15%). Through sensitivity analysis on the validation set, the step windows of the daily and weekly reference time scales are determined as $T_d = 4$, $\mathrm{T_w} = 3$ .LSTM, GRU, STGCN [4], DCRNN [6] and AGCRN [10] are selected as comparison models.

Main parameter settings of MS-DSTGFN: hidden dimension is 128, node embedding dimension is 8, number of multi-head attention heads is 4, Dropout=0.1, batch size is 32, and maximum training epochs is 150. Adam is adopted as the optimizer, with an initial learning rate of 1×10^{-3} and a weight decay of 1×10^{-4}. The loss function is 0.5MSE + 0.5MAE, and the early stopping patience is set to 20. The evaluation metrics include Mean Absolute Error (MAE), Root Mean Square Error (RMSE) and Mean Absolute Percentage Error (MAPE), all calculated on the original passenger flow scale after denormalization.

5.2. Overall Prediction Performance Comparison

Table 1 presents the prediction results of different models. Traditional time series models only utilize the historical passenger flow of a single station, which makes it difficult to effectively describe the propagation relationship of passenger flow across the network, resulting in relatively large prediction errors. Spatiotemporal graph models outperform traditional deep learning models on the whole by introducing spatial correlations between stations.

Table 1. Average Prediction Results of Each Model at Different Step Lengths.

Model	Single-step prediction (15min)			Two-step prediction (30min)			Three-step prediction (45min)			Four-step prediction (60min)		
	MAE	RMSE	MAPE	MAE	RMSE	MAPE	MAE	RMSE	MAPE	MAE	RMSE	MAPE
LSTM	78.61	123.56	18.25	83.34	130.30	19.35	96.94	152.40	22.82	112.07	178.07	29.11
GRU	72.81	116.03	15.84	77.84	119.74	18.16	91.27	139.69	22.92	107.58	169.66	28.35
STGCN	64.07	102.29	11.98	64.93	105.02	12.13	68.21	110.60	13.05	72.54	118.31	14.62
AGCRN	55.28	89.49	10.70	56.30	91.23	10.93	57.25	93.31	11.47	59.14	97.13	12.47
DCRNN	54.92	85.17	10.45	55.33	86.80	10.77	55.90	88.33	10.82	57.24	91.07	11.28
MS-DSTGFN	47.30	75.26	8.70	50.09	80.75	9.65	52.04	83.76	10.03	54.14	88.05	10.57

MS-DSTGFN achieves optimal results on all metrics of MAE, RMSE and MAPE. Compared with the strongest baseline DCRNN, the proposed model reduces MAE, RMSE and MAPE by 8.88%, 6.58% and 10.06%, respectively. This indicates that the physical topology graph can reflect explicit connection relationships between stations, while the adaptive graph can supplement implicit passenger flow correlations. The joint modeling of recent, daily reference and weekly reference sequences helps improve the model's adaptability to passenger flow patterns at different time periods, and the channel-aware decoding can further explore the coupling relationships among inbound, outbound and transfer passenger flows.

5.3. Per-Channel Performance Analysis

To further investigate the adaptability of the model to different types of passenger flow, Table 2 presents the detailed prediction results of MS-DSTGFN and DCRNN across the three channels.

Table 2. Performance comparison of different models by passenger flow category.

Model	Channel	MAE	RMSE	MAPE(%)	R2
DCRNN	in	57.36	99.02	12.50	0.9630
	out	46.10	74.26	11.15	0.9828
	trans	64.08	90.36	8.84	0.9902
MS-DSTGFN	in	51.17	90.25	10.84	0.9693
	out	43.95	73.62	10.02	0.9831
	trans	57.55	82.44	8.35	0.9918

The results show that the error of outbound passenger flow is the smallest overall, reflecting that its variation is relatively stable; the error of transfer passenger flow is the largest, indicating that it is more difficult to predict due to its large passenger flow base.

Compared with DCRNN, MS-DSTGFN achieves more prominent improvements in inbound and transfer passenger flows. This verifies that the model has better modeling capability for inbound and transfer passenger flows with stronger volatility.

5.4. Ablation Study

To verify the independent contribution of each core module of MS-DSTGFN, 6 groups of ablation variants are designed for comparison with the full model, and the results are shown in Table 3.

Table 3. Results of Ablation Experiments.

Variant	InMAE	OutMAE	TransMAE	AvgMAE(Δ)
MS-DSTGFN	51.17	43.95	57.55	50.89 (-)
w/o_Weekly	51.35	43.74	59.85	51.64(+0.75)
w/o_ChannelAware	51.85	45.29	58.58	51.91(+1.02)
w/o_AdaptiveGraph	51.93	44.27	61.02	52.41(+1.52)
w/o_Daily	54.54	44.53	60.54	53.20(+2.31)
w/o_MultiScale	54.93	44.76	61.41	53.70(+2.81)
w/o_CrossScale	56.19	47.30	66.78	56.75(+5.86)

The experimental results show that the prediction error of the model increases after any module is removed. Among them, removing the cross-scale attention fusion module leads to the largest increase in AvgMAE, indicating that cross-scale fusion can improve prediction performance by integrating multi-scale information. The daily reference time scale ranks second in terms of contribution, which demonstrates that the daily periodic pattern is more stable within the data span. After removing the adaptive graph, the MAE of transfer passenger flow rises, with a growth rate significantly higher than that of inbound and outbound passenger flows, which verifies the unique value of the data-driven graph for transfer station modeling.

5.5. Interpretability Analysis

To further understand the spatial dependencies learned by the model, analysis is carried out from two perspectives: the dual-graph fusion weights and the connection patterns of the adaptive graph, as shown in Figure 2.

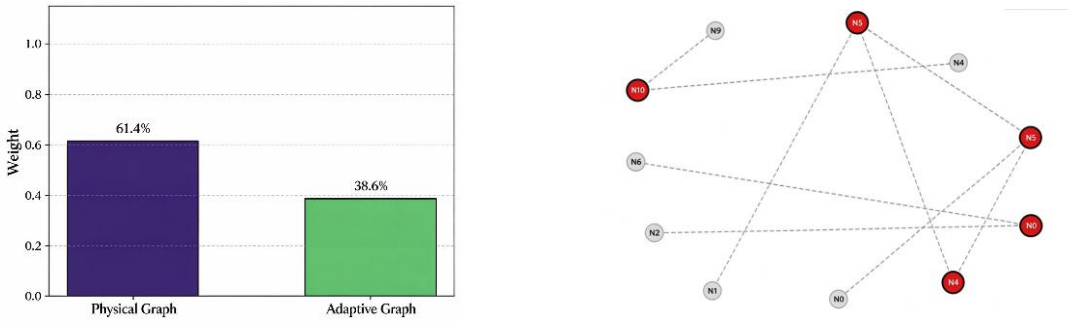


Fig 2. Dual-graph fusion weights and Top-2 correlated stations in the adaptive graph.

The fusion weights of the physical topology graph and the adaptive graph are 61.4% and 38.6% respectively, indicating that while retaining the physical connectivity constraints, the model captures potential correlations between non-physically adjacent stations. Further observation reveals that most high-weight connections of transfer stations in the adaptive graph are non-directly adjacent stations, which effectively supplements the implicit inter-station relationships that are difficult to characterize by physical topology.

5.6. Typical Station Analysis

To further analyze the prediction performance of MS-DSTGFN at different types of transfer stations, Fengqi Road Station (comprehensive service type in the central urban area) and Pengbu Station (residential commuting output type) are selected as typical cases for comparative analysis, as shown in Figure 3.

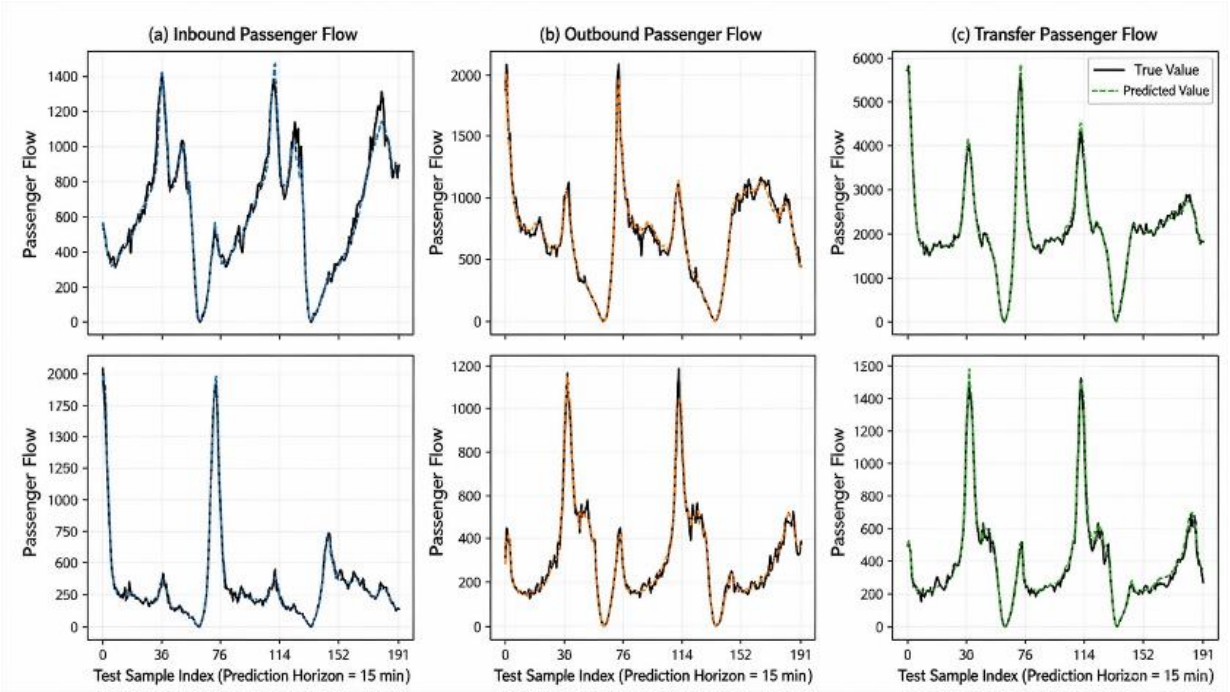


Fig 3. Single-step prediction curves of three types of passenger flow at typical stations.

Fengqi Road Station is located in the core business district of Hangzhou, with multiple functions including commercial services, employment attraction and transfer hub. At this station, MS-DSTGFN can effectively capture the rise, fall and local fluctuations of passenger flow during peak hours. Pengbu Station is a residential commuting output-type station, and the absolute errors of the three types of passenger flow predicted by the model at this station are lower than those at Fengqi Road Station, which is mainly attributed to the difference in passenger flow scale.

Overall, MS-DSTGFN exhibits favorable fitting capability for transfer stations of different functional types.

6. Conclusion

To address the joint prediction problem of inbound, outbound and transfer passenger flows at rail transit transfer stations, a multi-scale dual-graph spatiotemporal fusion network (MS-DSTGFN) is proposed. The transfer events are reconstructed through travel time decomposition and path search, a dual-graph propagation mechanism combining physical topology and adaptive graph is constructed to integrate explicit structural dependencies and implicit passenger flow correlations, and cross-scale attention is introduced to integrate recent, diurnal and weekly periodic patterns. Validation on Hangzhou Metro data demonstrates that the proposed method outperforms all comparison baselines on all evaluation metrics. Compared with the strongest baseline DCRNN, it reduces MAE, RMSE and MAPE by 8.88%, 6.58% and 10.06% respectively. The effectiveness of the dual-graph fusion and cross-scale attention modules is further verified by ablation experiments. The research results can provide technical support for passenger flow early warning and dynamic capacity allocation at transfer stations.

For future research, AFC data with a longer time span can be introduced to verify the seasonal generalization performance of the model, and external factors such as large-scale events and abnormal weather can be integrated to further improve the prediction robustness under complex scenarios.

References

[1] Fu R, Zhang Z, Li L. Using LSTM and GRU neural network methods for traffic flow prediction[C]//2016 31st Youth Academic Annual Conference of Chinese Association of Automation (YAC). 2016.

[2] ZHAO Lixin, JIN Hui. AN IMPROVED TRANSFORMER RAIL TRANSIT SHORT-TERM PASSENGER FLOW FORECASTING METHOD CONSIDERING FUSION ATTENTION MECHANISM [J]. Computer Applications and Software, 2025, 42(12): 98-104.

[3] ZHANG Wenjuan, YANG Haozhe, ZHANG Bin, et al. Short-time Passenger Flow Prediction Model of Urban Rail Transit Considering Multi-timescale Features [J]. Journal of Transportation Systems Engineering and Information Technology, 2022, 22(6): 212-223.

[4] Yu B, Yin H, Zhu Z. Spatio-temporal graph convolutional networks: A deep learning framework for traffic forecasting [C] //Proceedings of the 27th International Joint Conference on Artificial Intelligence. 2018: 3634-3640.

[5] WANG Runqi, HAO Yanxi, HU Hua, et al. Inbound Passenger Flow Prediction of Spatio-temporal Graph Convolutional Neural Net-work for Urban Rail Transit Based on Spatio-temporal Correlation [J]. Urban Mass Transit, 2024, 27(9): 91-96.

[6] Li Y, Yu R, Shahabi C, et al. Diffusion convolutional recurrent neural network: Data-driven traffic forecasting [C] //International Conference on Learning Representations. 2018.

[7] Wu Z, Pan S, Long G, et al. Graph WaveNet for deep spatial-temporal graph modeling [C] //Proceedings of the 28th International Joint Conference on Artificial Intelligence. 2019: 1907-1913.

[8] Guo S, Lin Y, Feng N, et al. Attention based spatial-temporal graph convolutional networks for traffic flow forecasting [C] //Proceedings of the AAAI Conference on Artificial Intelligence, 2019, 33(1): 922-929.

[9] TANG Zhengyi, HUANG Jiahuan, WANG Jinshui, et al. Metro passenger flow prediction based on the adaptive diffusion graph convolution attention network [J]. Journal of Railway Science and Engineering, 2024, 21(12): 4910-4923.

[10] Bai L, Yao L, Li C, et al. Adaptive graph convolutional recurrent network for traffic forecasting[J]. Advances in Neural Information Processing Systems, 2020, 33: 17804-17815.

[11] Wu Y, Liu T, Gong L, et al. Mining smart card data to estimate transfer passenger flow in a metro network[J]. IET Intelligent Transport Systems, 2024, 18(10): 1830-1846.

[12] ZHAO Jiandong, ZHU Dan, LIU Jiaxin. Metro Transfer Passenger Flow Prediction Based on STL-GRU [J]. Journal of South China University of Technology (Natural Science Edition), 2022, 50(5): 22-31.