Graph Neural Network-Enhanced Multivariate Time Series Forecasting with Series-Core Fusion

1st Yuntian Hou*
Department of AI and Advanced Computing
Xi'an Jiaotong-Liverpool University
Suzhou, China
Yuntian.Hou20@alumni.xjtlu.edu.cn

2nd Di Zhang*
Department of AI and Advanced Computing
Xi'an Jiaotong-Liverpool University
Suzhou, China
Di.Zhang@xjtlu.edu.cn

2025 8th International Conference on Advanced Algorithms and Control Engineering (ICAACE) | 979-8-3315-3508-7/25/$31.00 ©2025 IEEE | DOI: 10.1109/ICAACE65325.2025.11019045

***Abstract*—This study introduces Starformer, a hybrid model combining Graph Neural Networks (GNNs) with a novel Series-Core Fusion (SC-Fusion) mechanism for urban traffic prediction. By leveraging GNNs for spatial modeling and SC-Fusion for efficient temporal dependency capture, the model effectively addresses complex spatio-temporal dynamics in traffic systems. Evaluated on six widely used traffic datasets—METR-LA, PEMS-BAY, PEMS03, PEMS04, PEMS07, and PEMS08—Starformer demonstrates consistent and robust performance across diverse traffic conditions and regions. The results highlight its ability to model both short-term and long-term dependencies, making it well-suited for real-world applications. These findings emphasize the potential of integrating advanced neural network architectures for intelligent traffic management, contributing to smarter, more sustainable urban transportation systems.**

Index Terms*—*Traffic Prediction, Graph Neural Networks, Series-Core Fusion, Spatio-temporal Embeddings

I. Introduction

Urban traffic prediction is pivotal for alleviating congestion, optimizing public transit, and managing incidents. Accurate forecasts enhance navigation efficiency, reduce emissions, and support infrastructure planning. Despite its significance, traffic prediction is challenging due to the complex spatio-temporal dynamics influenced by road networks, weather, and fluctuating demand.

Traditional methods like ARIMA fail to capture the non-linear, non-stationary nature of traffic data. Deep learning models address these limitations but struggle with the intricacies of dynamic traffic systems. Traffic patterns involve intertwined spatio-temporal dependencies, complicating their joint modeling. Additionally, effective prediction must handle both short-term fluctuations and long-term trends, which remains an open problem.

Transformer-based architectures excel at capturing long-term temporal dependencies [1] but often neglect critical spatial interactions. Conversely, Graph Neural Networks (GNNs) effectively model spatial dynamics by representing road networks as graphs [2] but face limitations in temporal modeling. These shortcomings highlight the need for a hybrid approach that integrates spatial and temporal components.

We propose **Starformer**, a novel model combining GNNs for spatial modeling with Series-Core Fusion for enhanced temporal dependency capture. By replacing traditional attention mechanisms with Series-Core Fusion, **Starformer** achieves robust spatio-temporal representations. Our experiments demonstrate **Starformer**'s superiority over state-of-the-art models, including **SOTAformer** [3] and **Autoformer** [9], on real-world traffic datasets. Inspired by **SOFTS** [11], **Starformer** balances computational efficiency with predictive accuracy, setting a new benchmark in traffic forecasting(Fig. 1).

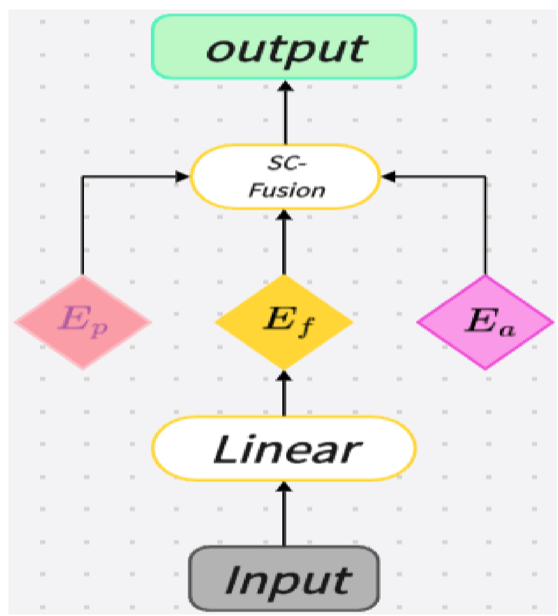

Fig. 1. Input Embeddings in Different Traffic Models.

This paper reviews related work, defines the problem, introduces the proposed model, and evaluates its performance.

II. Related Work

A. Graph Neural Networks in Traffic Prediction

Graph Neural Networks (GNNs) are ideal for modeling spatial dependencies in road networks [6]. Scarselli et al. (2009) introduced the first GNN model [4], which Kipf and Welling (2017) improved into Graph Convolutional Networks (GCNs) [5] for large-scale data processing. Li et al. (2018) proposed the Diffusion Convolutional Recurrent Neural Network

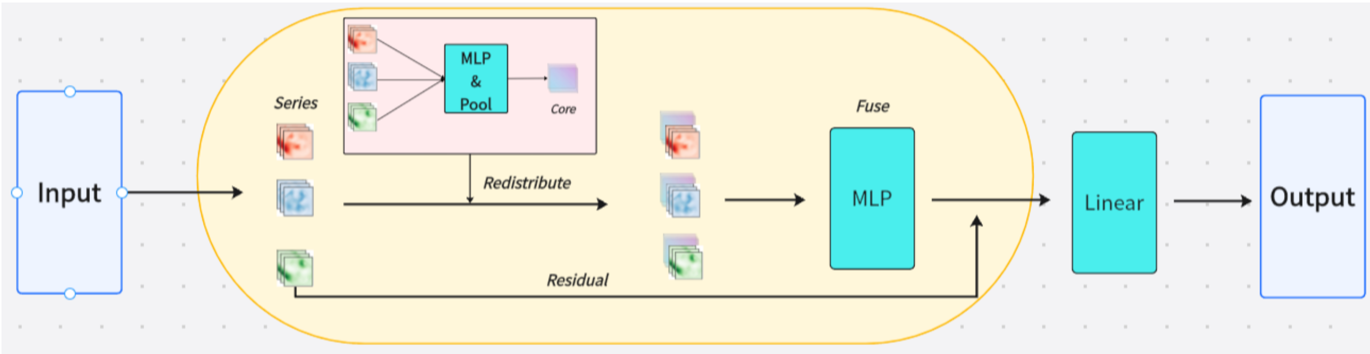

Fig. 2. The architecture of the proposed PreSTAEFormer model.

(DCRNN), combining GNNs with recurrent layers to model both spatial and temporal dynamics [7]. Other advancements, like those by Bai et al. (2020), introduced adaptive methods for real-time traffic prediction [8]. For Starformer, GNNs are used to capture spatial dependencies, while Series-Core Fusion replaces traditional attention mechanisms to enhance model efficiency and robustness.

B. Time Series Prediction and Channel Independence

Multivariate time series forecasting plays a significant role in decision-making for traffic prediction. While Transformer models have demonstrated effectiveness in sequence modeling, they struggle with complexity and scalability, especially when dealing with long-term and multivariate time series. To mitigate this, some models focus on reducing computational demands by simplifying the attention mechanism, such as by leveraging autocorrelation or frequency-domain strategies [9], [10]. A key challenge in multivariate time series is how to handle inter-channel dependencies. Traditional models often mix channels together, leading to reduced robustness when distribution shifts occur [12]. Recent research has shifted towards channel-independent methods to improve model stability, but these can neglect useful inter-channel interactions [13], [14]. Starformer addresses these issues by using Series-Core Fusion, which efficiently captures both temporal dependencies and maintains channel independence, resulting in a more robust and scalable solution for traffic prediction tasks.

C. Starformer: Integrating GNNs with Series-Core Fusion

Starformer is a novel approach that combines the spatial modeling capabilities of GNNs with the efficient time-series modeling of Series-Core Fusion. Unlike traditional Transformer-based models, which rely on attention mechanisms that can be computationally expensive, Starformer replaces attention with Series-Core Fusion, enhancing both scalability and accuracy. This fusion mechanism enables the model to capture complex temporal dependencies while maintaining the independence of different channels, offering superior robustness to distribution shifts and better generalization in large-scale traffic forecasting.

III. Problem Definition

Urban traffic prediction forecasts future traffic patterns $\mathcal{Y} \in \mathbb{R}^{F \times N}$ using historical data $\mathcal{X} \in \mathbb{R}^{P \times N}$, where P and F denote past and future time steps, and N represents the number of traffic sensors. The data encompasses temporal and spatial dynamics, modeled as a graph $G = (V, E)$, with V as nodes (traffic sensors) and E as edges (road network connections).

The goal is to learn a function $f(\mathcal{X}, G) \to \mathcal{Y}$ that captures spatio-temporal interactions. The **Starformer** framework integrates Graph Neural Networks (GNNs) for spatial dependency modeling and Series-Core Fusion (SC-Fusion) for temporal dynamics. Spatial features are derived via GNN embeddings, while SC-Fusion replaces traditional attention mechanisms to process temporal patterns more efficiently, improving scalability.

$$\hat{\mathcal{Y}} = \text{SC-Fusion}(Z_t, H_t) \tag{1}$$

The model is trained to minimize the Mean Absolute Error (MAE) to improve the accuracy of the traffic prediction.

IV. Methodology

Starformer integrates Graph Neural Networks (GNNs) for spatial dependency modeling and Series-Core Fusion (SC-Fusion) for efficient temporal dependency modeling. The architecture utilizes spatial embeddings from the GNN layers and temporal embeddings from SC-Fusion to predict traffic conditions (Fig. 2).

A. Graph Neural Networks (GNNs)

The GNN module captures spatial dependencies by processing sensor data $X_t \in \mathbb{R}^{N \times d}$ over the road network graph $G = (V, E)$, where N represents the number of sensors and d is the feature dimension. A extbfdynamic adjacency matrix updates edges based on real-time correlations, enabling the model to capture evolving spatial relationships. Additionally, an extbfadaptive embedding mechanism expands node-specific features across spatial and temporal dimensions, improving context-aware representations.

The GNN employs a extbfdual attention mechanism, where temporal attention layers capture sequence-level interactions and spatial layers focus on node-level dependencies. This

TABLE I
COMPARISON OF FORECASTING PERFORMANCE ACROSS DIFFERENT MODELS ON METR-LA AND PEMS-BAY

Dataset	Horizon	Metric	HI	DGCRN	DLinear	DSFormer	Autoformer	GTS	FEDformer	Informer	STAEformer	Starformer
METR-LA	Horizon 3 (15 min)	MAE	6.80	3.43	3.80	4.06	7.87	3.65	3.98	4.97	3.53	**3.36**
		RMSE	14.21	6.71	8.10	7.15	11.75	**5.72**	7.27	9.50	9.30	7.13
		MAPE	16.72%	8.44%	9.30%	19.63%	11.75%	7.84%	11.35%	15.72%	7.94%	**7.55%**
	Horizon 6 (30 min)	MAE	6.80	4.71	4.80	4.63	7.91	4.45	4.38	4.91	4.44	**4.32**
		RMSE	14.21	8.75	10.02	8.36	11.85	**7.62**	7.73	9.35	11.62	9.37
		MAPE	16.72%	11.94%	12.26%	19.83%	19.83%	11.04%	11.60%	15.50%	9.63%	**9.33%**
	Horizon 12 (60 min)	MAE	6.80	7.86	6.16	5.65	8.01	7.88	5.67	8.85	5.68	**5.58**
		RMSE	14.20	11.92	11.76	9.81	12.05	11.41	**8.36**	9.30	14.35	10.36
		MAPE	12.15%	19.48%	16.86%	14.97%	20.29%	18.19%	12.60%	15.48%	11.92%	**11.48%**
PEMS-BAY	Horizon 3 (15 min)	MAE	3.05	1.97	2.15	2.42	4.15	1.98	2.30	3.01	2.11	**1.92**
		RMSE	7.03	4.15	4.55	4.63	6.30	**3.62**	4.00	5.03	5.10	4.25
		MAPE	6.83%	4.12%	4.60%	9.12%	7.55%	3.85%	5.35%	6.81%	3.98%	**3.75%**
	Horizon 6 (30 min)	MAE	3.06	2.65	2.70	2.58	4.19	2.46	2.53	2.88	2.53	**2.42**
		RMSE	7.04	4.60	5.10	4.36	6.40	**4.22**	4.25	5.07	5.92	4.77
		MAPE	6.84%	5.12%	5.26%	9.15%	9.24%	4.95%	5.15%	7.35%	4.15%	**4.03%**
	Horizon 12 (60 min)	MAE	**3.06**	3.98	3.65	3.33	4.25	3.96	3.38	4.15	3.45	3.15
		RMSE	7.03	5.32	5.40	**4.88**	6.51	5.16	4.89	5.40	6.00	5.45
		MAPE	6.95%	9.25%	8.76%	7.98%	9.42%	8.10%	6.72%	7.50%	5.42%	**5.23%**

ensures comprehensive spatio-temporal modeling while optimizing information propagation between input and output layers for efficient prediction.

B. Series-Core Fusion for Temporal Dependencies

Temporal dependencies are captured using Series-Core Fusion (SC-Fusion), replacing traditional attention mechanisms. SC-Fusion aggregates temporal information into a global core representation and redistributes it back to individual series. The temporal embeddings are computed as:

$$H_t = \text{SC-Fusion}(Z_t) \tag{2}$$

C. Star Aggregate-Redistribute (STAR) Module

The STAR module captures inter-channel dependencies by aggregating series embeddings into a global core representation:

$$o = f(s_1, s_2, \ldots, s_C) \tag{3}$$

where s_i represents the series embedding of the i-th channel. This core representation is fused with individual series embeddings:

$$S_i = \text{MLP}(\text{Repeat\_Concat}(S_{i-1}, o)) + S_{i-1} \tag{4}$$

This fusion integrates local and global information efficiently, enhancing predictive performance.

V. EXPERIMENT

In this section, we present the experimental setup and results for evaluating the performance of the proposed **Starformer** model. The experiments focus on urban traffic prediction, specifically using the METR-LA dataset. We compare the performance of **Starformer** against several state-of-the-art models, evaluating accuracy and efficiency using a variety of performance metrics.

A. Dataset Description

We assess the performance of **Starformer** on six widely used traffic prediction datasets: **METR-LA**, **PEMS-BAY**, **PEMS03**, **PEMS04**, **PEMS07**, and **PEMS08**. These datasets encompass diverse traffic conditions, including both vehicle speed and flow data, collected at 5-minute intervals over periods ranging from two to six months. Detailed statistics of each dataset are provided in Table II.

TABLE II
STATISTICS OF DATASETS.

Type	Dataset	# Node	# Edge	# Time Step
Speed	METR-LA	207	1722	34272
Speed	PEMS-BAY	325	2369	52116
Flow	PEMS03	358	768	26208
Flow	PEMS04	307	680	16992
Flow	PEMS07	883	1512	28224
Flow	PEMS08	170	548	17856

These datasets exhibit unique characteristics that test the model's adaptability. **METR-LA** and **PEMS-BAY**, for example, primarily feature urban traffic speed data with strong spatiotemporal correlations, reflecting dense metropolitan networks. Conversely, **PEMS03**, **PEMS04**, **PEMS07**, and **PEMS08** focus on vehicle flow, representing regional traffic patterns with varying densities and temporal dynamics. The differences in node counts and edge connections across these datasets provide diverse scenarios for spatiotemporal modeling, testing the model's ability to handle both sparse and complex networks.

B. Data Preprocessing

To ensure compatibility with **Starformer**, the raw traffic data undergoes the following preprocessing steps:

Normalization: Traffic speed and flow values are scaled to [0, 1] using Min-Max normalization to standardize features. **Handling Missing Data:** Missing values caused by sensor malfunctions are reconstructed using interpolation methods.

Data Splitting: The datasets are partitioned into training (70%), validation (15%), and testing (15%) sets. A sliding window approach generates input-output sequences, where P past time steps form the input and F future time steps are the prediction targets.

These streamlined steps prepare the data for efficient model training, enabling **Starformer** to adapt effectively to diverse traffic scenarios.

C. Evaluation Metrics

The performance of Starformer is assessed using the following metrics:

- **Mean Absolute Error (MAE)**:

$$\mathrm{MAE} = \frac{1}{NF}\sum_{i=1}^{N}\sum_{j=1}^{F}|\hat{\mathcal{Y}}_{i,j} - \mathcal{Y}_{i,j}| \tag{5}$$

- **Root Mean Squared Error (RMSE)**:

$$\mathrm{RMSE} = \sqrt{\frac{1}{NF}\sum_{i=1}^{N}\sum_{j=1}^{F}(\hat{\mathcal{Y}}_{i,j} - \mathcal{Y}_{i,j})^2} \tag{6}$$

- **Mean Absolute Percentage Error (MAPE)**:

$$\mathrm{MAPE} = \frac{1}{NF}\sum_{i=1}^{N}\sum_{j=1}^{F}\left|\frac{\hat{\mathcal{Y}}_{i,j} - \mathcal{Y}_{i,j}}{\mathcal{Y}_{i,j}}\right| \tag{7}$$

D. Efficiency Analysis

Training Efficiency: Among tested models, **Informer** trains fastest, completing an epoch in **161 seconds** due to its sparse attention. **Autoformer** and **FEDformer** require **390** and **471 seconds/epoch**, respectively, due to complex architectures. **STAEFormer** and **StarFormer** are competitive, taking **360** and **365 seconds/epoch**, with **StarFormer** benefiting from optimized spatio-temporal embeddings.

Testing Efficiency: **FEDformer** tests fastest at **31 seconds**, leveraging frequency-domain operations. **STAEFormer** and **StarFormer** follow at **40** and **41 seconds**, balancing speed and accuracy. **Autoformer** and **Informer** require **64** and **81 seconds** due to computational overhead.

Key Insights: **StarFormer** balances training and testing efficiency, ideal for scenarios requiring high accuracy and moderate computational demands. While **FEDformer** excels in real-time inference, **StarFormer** offers versatility across broader applications(Fig. 3).

E. Results Analysis and Comparative Performance

In Table I and Table III, **StarFormer** demonstrates superior performance across six widely used datasets, including **METR-LA**, **PEMS-BAY**, **PEMS03**, **PEMS04**, **PEMS07**, and **PEMS08**. On **METR-LA** and **PEMS-BAY**, it achieves the lowest Mean Absolute Error (MAE) values (**3.36** and **1.92**) for short-term forecasting horizons, outperforming advanced models such as **FEDformer** and **Informer**. Additionally, it maintains strong accuracy over extended horizons, highlighting its robust spatio-temporal modeling capabilities.

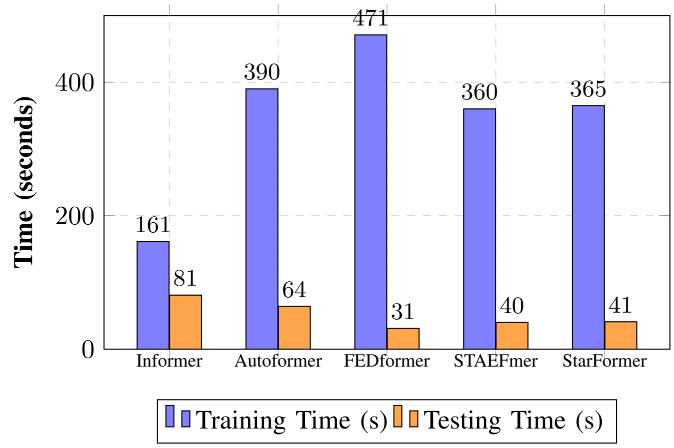

Fig. 3. Training and Testing Times of Models on the METR-LA Dataset.

For datasets like **PEMS04** and **PEMS08**, **StarFormer** achieves remarkable MAE values (**18.68** and **13.98**), demonstrating its adaptability to diverse traffic conditions in Table IV. By integrating Graph Neural Networks (GNNs) and Series-Core Fusion, **StarFormer** consistently surpasses both traditional models (e.g., **DCRNN**) and contemporary Transformer-based approaches (**FEDformer**, **Informer**), establishing its effectiveness and generalizability across different scenarios.

F. Performance Under Diverse Traffic Conditions

The datasets include both peak and non-peak hours, showcasing diverse traffic patterns. **Starformer** effectively handles these variations by using GNNs for localized congestion modeling during peak hours and SC-Fusion for stability during steady flows. While extreme scenarios like accidents are not explicitly labeled, the model's robust spatio-temporal architecture suggests adaptability, warranting further evaluation under such conditions.

G. Key Insights and Real-World Potential

The effectiveness of **Starformer** stems from its innovative design:

- **Dynamic Spatial Modeling**: GNNs capture evolving spatial dependencies, surpassing purely temporal approaches.
- **Efficient Temporal Fusion**: SC-Fusion improves long-term forecasting with reduced computational costs.
- **Integrated Spatio-Temporal Framework**: The fusion of spatial and temporal features ensures consistent performance across various horizons.

These capabilities position **Starformer** as a robust tool for traffic management and urban planning, offering accurate and reliable spatio-temporal predictions.

VI. Conclusion

This study introduces **Starformer**, a hybrid model combining GNNs with a novel Series-Core Fusion (SC-Fusion) mechanism to address the challenges of urban traffic prediction. By integrating spatio-temporal representations, dynamic

TABLE III
PERFORMANCE COMPARISON ACROSS PEMS03, PEMS04, PEMS07, AND PEMS08 DATASETS

Methods	PEMS03			PEMS04			PEMS07			PEMS08		
	MAE	RMSE	WAPE	MAE	RMSE	WAPE	MAE	RMSE	WAPE	MAE	RMSE	WAPE
GWNet	20.45	32.68	13.55%	19.36	31.72	13.31%	22.10	34.95	9.42%	15.07	23.85	9.51%
DCRNN	21.10	33.91	14.02%	21.22	33.44	14.17%	23.35	36.44	10.11%	16.82	26.36	10.92%
DGCRN	19.80	31.85	12.80%	18.84	30.48	12.92%	21.08	33.72	9.38%	14.77	23.81	9.77%
Linear	39.02	64.50	17.91%	37.42	62.14	17.22%	37.25	60.18	15.22%	34.04	57.07	14.71%
DLinear	39.15	64.72	17.94%	37.51	62.21	17.26%	37.40	60.45	15.25%	34.51	57.18	14.76%
NLinear	39.22	64.88	17.99%	37.62	62.38	17.31%	37.50	60.65	15.30%	34.54	57.58	14.74%
Informer	28.25	45.02	13.02%	27.94	44.74	12.84%	27.98	44.50	11.84%	26.92	43.79	11.63%
Autoformer	35.14	51.22	15.02%	34.72	50.33	14.81%	34.50	53.15	14.22%	33.75	52.31	14.13%
FEDformer	27.30	42.15	12.60%	26.89	41.16	**12.39%**	26.10	41.15	10.90%	25.14	39.17	10.87%
STAEFormer	19.55	31.20	**12.50%**	18.76	30.77	12.72%	20.55	33.25	9.22%	14.05	**23.27**	9.93%
Starformer	**19.42**	**30.88**	12.51%	**18.68**	**30.62**	12.83%	**20.38**	**33.06**	**9.12%**	**13.98**	23.43	**9.45%**

TABLE IV
PERFORMANCE COMPARISON ACROSS METR-LA, PEMS04, AND PEMS08 DATASETS

Model	METR-LA			PEMS04			PEMS08		
	MAE	RMSE	MAPE (%)	MAE	RMSE	MAPE (%)	MAE	RMSE	MAPE (%)
Informer	6.243	9.383	15.566%	27.94	44.74	12.84%	26.92	43.79	11.63%
Autoformer	7.930	11.880	17.290%	34.72	50.33	14.81%	33.75	52.31	14.13%
FEDformer	4.670	**7.780**	11.850%	26.89	41.16	12.39%	25.14	39.17	10.87%
STAEformer	4.550	11.090	9.163%	18.76	30.77	**12.72%**	14.05	**23.27**	9.93%
Starformer	**4.158**	10.226	**9.049%**	**18.68**	**30.62**	12.83%	**13.98**	23.43	**9.45%**

embeddings, and efficient fusion techniques, **Starformer** effectively captures intricate spatial and temporal dependencies. Experimental results on six datasets highlight its superior accuracy, scalability, and efficiency, surpassing existing state-of-the-art models. These findings underscore the potential of combining GNNs with innovative temporal fusion mechanisms for spatio-temporal prediction tasks. Future work will explore extending this framework to broader real-world applications and other domains requiring spatio-temporal forecasting.

ACKNOWLEDGMENTS

Partial of the work is supported by the XJTLU REF (18-01-04) and KSF-E-32

REFERENCES

[1] H. Zhou, S. Zhang, J. Peng, S. Zhang, J. Li, H. Xiong, and W. Zhang, "Informer: Beyond efficient transformer for long sequence time-series forecasting," in *Proc. AAAI Conf. Artif. Intell.*, vol. 35, no. 11, pp. 11106–11115, 2021.

[2] B. Yu, H. Yin, and Z. Zhu, "Spatio-temporal graph convolutional networks: A deep learning framework for traffic forecasting," in *Proc. 27th Int. Joint Conf. Artif. Intell. (IJCAI)*, 2018.

[3] H. Liu, Z. Dong, and R. Jiang, "Spatio-temporal adaptive embedding makes vanilla transformer SOTA for traffic forecasting," in *Proc. CIKM*, 2023.

[4] F. Scarselli, M. Gori, A. C. Tsoi, M. Hagenbuchner, and G. Monfardini, "The graph neural network model," *IEEE Trans. Neural Netw.*, vol. 20, no. 1, pp. 61–80, 2009.

[5] T. N. Kipf and M. Welling, "Semi-supervised classification with graph convolutional networks," in *Proc. Int. Conf. Learn. Representations (ICLR)*, 2017.

[6] W. L. Hamilton, R. Ying, and J. Leskovec, "Inductive representation learning on large graphs," in *Proc. Advances Neural Inf. Process. Syst. (NeurIPS)*, vol. 30, 2017.

[7] Y. Li, R. Yu, C. Shahabi, and Y. Liu, "Diffusion convolutional recurrent neural network: Data-driven traffic forecasting," in *Proc. Int. Conf. Learn. Representations (ICLR)*, 2018.

[8] L. Bai, L. Yao, S. Kanhere, X. Wang, and W. Liu, "Adaptive graph convolutional neural networks," in *Proc. Advances Neural Inf. Process. Syst. (NeurIPS)*, vol. 33, 2020.

[9] H. Wu, J. Xu, J. Wang, and M. Long, "Autoformer: Decomposition transformers with auto-correlation for long-term series forecasting," in *Proceedings of NeurIPS*, 2021.

[10] T. Zhou, Z. Ma, Q. Wen, X. Wang, L. Sun, and R. Jin, "FEDformer: Frequency enhanced decomposed transformer for long-term series forecasting," in *Proceedings of the 39th International Conference on Machine Learning*, PMLR, 2022.

[11] Han, L., Chen, X.-Y., Ye, H.-J., and Zhan, D.-C. (2024). SOFTS: Efficient multivariate time series forecasting with series-core fusion. arXiv preprint arXiv:2404.14197. https://arxiv.org/abs/2404.14197

[12] Ailing Zeng, Muxi Chen, Lei Zhang, and Qiang Xu. Are transformers effective for time series forecasting? In AAAI, pages 11121–11128, 2023.

[13] Shengsheng Lin, Weiwei Lin, Wentai Wu, Feiyu Zhao, Ruichao Mo, and Haotong Zhang.Segrnn: Segment recurrent neural network for long-term time series forecasting. CoRR, abs/2308.11200, 2023.

[14] Yuqi Nie, Nam H. Nguyen, Phanwadee Sinthong, and Jayant Kalagnanam. A time series is worth 64 words: Long-term forecasting with transformers. In ICLR, 2023.