DYNAMIC FREQUENCY DOMAIN GRAPH CONVOLUTIONAL NETWORK FOR TRAFFIC FORECASTING

Yujie Li[1,2], *Zezhi Shao*[1,2], *Yongjun Xu*[1,2], *Qiang Qiu*[1,2], *Zhaogang Cao*[3], *Fei Wang*[1,2*]

[1]Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China
[2]University of Chinese Academy of Sciences, Beijing, China
[3]Academy of Cyber, Beijing, China

ABSTRACT

Complex spatial dependencies in transportation networks make traffic prediction extremely challenging. Much existing work is devoted to learning dynamic graph structures among sensors, and the strategy of mining spatial dependencies from traffic data, known as data-driven, tends to be an intuitive and effective approach. However, Time-Shift of traffic patterns and noise induced by random factors hinder data-driven spatial dependence modeling. In this paper, we propose a novel dynamic frequency domain graph convolution network (DFDGCN) to capture spatial dependencies. Specifically, we mitigate the effects of time-shift by Fourier transform, and introduce the identity embedding of sensors and time embedding when capturing data for graph learning since traffic data with noise is not entirely reliable. The graph is combined with static predefined and self-adaptive graphs during graph convolution to predict future traffic data through classical causal convolutions. Extensive experiments on four real-world datasets demonstrate that our model is effective and outperforms the baselines.

Index Terms— Traffic prediction, frequency domain signal processing, multivariate time series analysis, dynamic graph learning, graph convolution.

1. INTRODUCTION

With the increase of global urbanization and population concentration, complex transportation environments have become a major challenge for urban management, leading to the urgent need for Intelligent Transportation Systems (ITS)[1] about AI[2] to enhance the population carrying capacity of cities.

Traffic prediction is an indispensable part of ITS, which aims to accurately predict future traffic data by learning historical traffic states and patterns[3] obtained from sensors. However, due to the high correlation between sensors, future information of each sensor relies not only on its own historical information but also on the historical data of other sensors. Spatio-temporal graph neural networks are proposed to capture not only the temporal dependence of traffic data inherent in each sensor, but also to mine the spatial dependence among sensors, and this class of methods has been proven to outperform general time series forecasting[4] methods such as ARIMA[5] and LSTM[6] with the main enhancement coming from mining the reliable spatial dependencies among sensors.

Mining effective spatial dependencies, also known as graph structures, has been an important challenge in traffic prediction. Early schemes such as DCRNN[7] relies on prior spatial information to calculate predefined graph structures, GWNet[8] self-adaptively learns graph structures with learnable parameters. STGCN[9] , AGCRN[10], and MTGNN[11] all apply or improve upon these static graphs to achieve outstanding enhancements.

In subsequent research, researchers find that the spatial dependence of different times is not the same due to the complex urban transportation environment. ASTGNN[12] and DMSTGCN[13] apply learnable parameters at each timestamp such as day of week or hour of day to self-adaptively capture spatial dependencies over different time periods. In contrast, the approach of applying changing historical traffic data to model dynamic graphs is called data-driven. Since traffic patterns can always change randomly, data-driven approaches that can adjust the graph structure based on historical information during the testing phase [14] tend to be more effective than self-adaptive approaches that rely purely on learnable parameters. STFGCN[15], DGCRN[16], and DSTAGNN[17] all utilize self-attention scores of time domain information or the spatial proximity between traffic data in time domain to construct the graph structures.

However, we argue that data-driven modeling of spatial dependence faces two problems: Time-Shift and Data Noise.

Time-shift is a common problem in urban transportation. For instance, during urban commuting, when people come off duty from industrial areas, the congested traffic can take half an hour or more to reach residential areas. Time-shift of traffic patterns due to similar situations is common in modern cities

---

because of urban functional zones, and it can result in tidal wave-like delay variations between interrelated traffic data as shown in Fig. 1 .
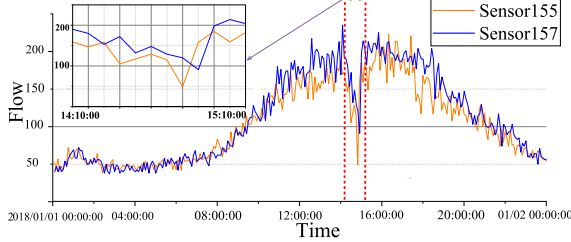
Fig. 1. Time-shift in Traffic Forecasting

However, time-shift can invalidate spatial proximity measures, which are widely applied to capture spatial dependence of traffic data. It's because schemes such as Euclidean distance and cosine similarity are computed in aligned dimensions, but changes in interrelated traffic data affected by time shifts do not occur at the same timestamp, which raises difficulties in examining proximity to capture the correlation between roads in time domain. Fig.2 represents the cosine similarity between traffic flows at the same observation time in one week, and we can observe that the similarity between traffic flows in time domain exhibits low differentiation and stacking. This is due to the misalignment of the temporal dimension caused by time-shift, which makes the traffic flows on related roads have similar similarity to those on uncorrelated roads, which makes it difficult to explore valid spatial dependencies through proximity.
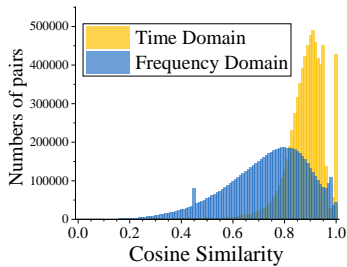
Fig. 2. Cosine similarity of traffic flow in time and frequency domains within one weak

Noise in traffic data is an unavoidable problem as in general data-driven strategies, and is often triggered by emergencies such as traffic accidents or road construction. Therefore, traffic data with noise is not always reliable, and it is a realistic challenge to introduce additional information to reduce the effect of noise while learning the spatial relationships between sensors.

To solve the above problems, we propose a dynamic frequency domain graph convolutional network(DFDGCN). To mitigate the effect of time-shift, we propose to examine the traffic patterns through the frequency domain information of traffic data. As shown in Fig.2, the frequency domain information between sensors exhibits better differentiation, it is because even if the traffic information is time-shifted, relying on the time-shift property of the Fourier transform, the frequency domain components of traffic data will still be in the same phase dimension, which will facilitate us to mine dynamic spatial dependence subsequently. In addition, we introduce identity embedding of sensors and time embedding when capturing traffic data for graph learning to minimize the effect of data noise. Specifically, we summarize the main contributions of our work as follows:

- We focus on the time-shift problem when mining spatial dependence of traffic data, and to the best of our knowledge we are the first to mitigate it by frequency domain analysis with the property of Fourier transform.
- We propose DFDGCN to learn dynamic graph structures among sensors by combining traffic data with identity embedding and time embedding to cope with data noise.
- Extensive experiments on four publicly available datasets demonstrate that our proposed model outperforms baselines, and ablation experiments witness the validity of our model.

2. METHODOLOGY

2.1. Fourier transformation

As described in the previous section, Time-Shift problem is a challenge for data-driven spatial dependence modeling, and we propose to address the problem by Fourier Transforming, which transfers traffic data to the frequency domain.

Assuming the existence of traffic data $f(t)$ captured by sensors at a particular intersection, traffic data $f(t - t_0)$ can be captured when the traffic flow is delayed by time t_0 to reach the next intersection. In the process of spatial dependence modeling, learning the relationship between $f(t)$ and $f(t - t_0)$ under the same time window is difficult or has a high computational overhead due to Time-Shift.

However, their corresponding Fourier transforms $F(\omega)$, $F_{t_0}(\omega)$ are related under the same dimension, according to the definition of Fourier transform:

$$F_{(\omega)} = \int_{-\infty}^{\infty} f(t)e^{-j\omega t}dt \qquad (1)$$

$$F_{t_0}(\omega) = \int_{-\infty}^{\infty} f(t - t_0)e^{-j\omega t}dt \qquad (2)$$

According to $t - t_0 = c$, $F_{t_0}(\omega)$ can be continued to be expressed as:

$$F_{t_0}(\omega) = \int_{-\infty}^{\infty} f(c)e^{-j\omega(t_0+c)}dc \qquad (3)$$

$$= [\int_{-\infty}^{\infty} f(c) \cdot e^{-j\omega c}dc]e^{-j\omega t_0} = F_{(\omega)} \cdot e^{-j\omega t_0}$$

Thus we prove that the traffic data in the frequency domain is represented in the same phase dimension, which solves the trouble caused by time-shift and means that learning the spatial dependencies between sensors will be more accurate and convenient.

2.2. DFDGCN

Urban traffic environments are complex and variable, and to alleviate the time-shift as well as noise problems in the data-driven dynamic graph modeling, we propose a frequency domain graph module described by Fig.3. The core of DFDGCN is to update the dynamic adjacency matrix A_D based on the traffic data observed at the current observation window.
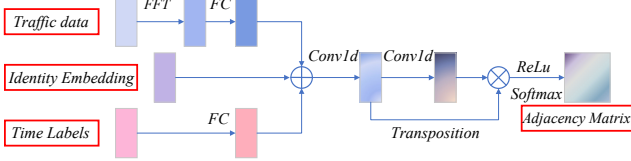
Fig. 3. The structure of proposed Frequency Domain Graph Module

According to section 2.1, traffic data in the frequency domain receives less effect of time shift than that in the time domain. As a result, we transfer the traffic data X_t at each observation window to the frequency domain by Fourier transform.

$$F_t = FFT(X_t) \tag{4}$$

In order to deal with the noise in traffic data, we introduce identity embedding[18] and time embedding[19] to increase additional information about the transportation network, so that we can easily extract the effective traffic patterns of each sensor and explore the spatial dependency among them:

$$DE_t = W_{F,t} \cdot F_t || E_t || W_{T,t} \cdot (T_t^W || T_t^D) \tag{5}$$

where E_t represents learnable identity embedding of each sensor, T_t^W T_t^D denote the labels indicating day of week and hour of day, all W are the learnable parameters for embedding the individual labels.

In addition, we apply a one-dimensional convolution layer with 1×1 convolution kernel for further embedding, in order to learn the connections between the dimensions of DE_t.

Finally, we apply a fully connected layer to learn DE_t again to obtain directionality and then matrix multiply it with the transposition of DE_t. After the activation function and $Softmax$ we obtain the final adjacency matrix:

$$A_D^t = Softmax(ReLU(DE_t W_{adj} DE_t^T)) \tag{6}$$

Consistent with previous work, we consider the dynamic frequency domain graph as a transition matrix for the hidden diffusion process. By combining it with the predefined graphs P in DCRNN and the self-adaptive graph A_{adt} in GWNet, we propose the following graph convolutional layer:

$$Z_t = \sum_{k=0}^{K} (P^k X_t W_{k,1} + A_{adt}^k X_t W_{k,2} + A_D^t X_t W_{k,3}) \tag{7}$$

For the processing of temporal information we adopt the same causal convolutional with residual network as in GWNet[8].

3. EXPERIMENTS

3.1. Datasets

We verify DFDGCN on four commonly used large real-world datasets with tens of thousands of time steps and hundreds of sensors. The statistics are summarized in Table 1. PEMS-BAY is a traffic speed dataset and the others are traffic flow datasets, and the time steps for each dataset are obtained by sampling at five-minute intervals.

Table 1. Statistics of Datasets

Type	Datasets	Time Steps	Nodes	Edges
Speed	**PEMS-BAY**	52116	325	2369
Flow	**PEMS03**	26208	358	547
Flow	**PEMS07**	28224	883	866
Flow	**PEMS08**	17856	170	295

3.2. Baselines and Metrics

We selected publicly available and classical baselines for traffic predicting, including recent traditional method HI[20], representative deep learning methods GWNet[8], DCRNN[7], AGCRN[10], STGCN[9], MTGNN[11], DGCRN[16]. Due to space limitations, we do not introduce each method in detail.We examine all baselines by three metrics, including mean absolute error (MAE), root mean square error (RMSE), and mean absolute percentage error (MAPE).

3.3. Experimental Setups

The four datasets are divided into training, validation, and test sets in the ratio of 7:1:2. We predict future traffic data of 12 time steps with historical traffic data of length 12, and compare the performances of the 3rd, 6th, 12th and the average of the 12 timestamps, which are labeled @3, @6, @12 and Avg.. In our model, The dimensions of the Fourier transformed time series embedding as well as the identity embedding are both 10, and $T_t^W T_t^D$ in the temporal embedding are both 12 dimensions. The embedding size after 1-dimensional convolution is 30. Baselines as well as training strategies can be found in the publicly available benchmark[1]. The best experimental results will be marked boldly while the second will be underlined. The source code of DFDGCN is available [2].

3.4. Experimental Results

From Tab.2, DFDGCN basically outperforms Baselines in the three metrics. Compared with GWNet, which is the base model of DFDGCN, the better results of DFDGCN indicate that the frequency domain graph actually captures the dynamic spatial dependence accurately. Compared with DGCRN which also mines dynamic spatial dependencies from traffic data, the rationality of our proposed problem and the validity of our model can be preliminarily judged.

Table 2. Results of DFDGCN and Baselines for traffic forecasting on four real datasets

Datasets		PEMS-BAY				PEMS03				PEMS07				PEMS08			
Method	Metric	@3	@6	@12	Avg.	@3	@6	@12	Avg.	@3	@6	@12	Avg.	@3	@6	@12	Avg.
HI	MAE	3.06	3.06	3.05	3.05	32.46	32.45	32.44	32.45	49.02	49.04	49.06	49.03	36.65	36.66	36.68	36.66
	RMSE	7.05	7.04	7.03	7.05	49.78	49.76	49.75	49.76	71.15	71.18	71.21	71.18	50.44	50.45	50.46	50.45
	MAPE	6.85%	6.84%	6.83%	6.84%	30.58%	30.59%	30.63%	30.60%	22.73%	22.75%	22.79%	22.75%	21.60%	21.63%	21.68%	21.63%
GWNet	MAE	**1.30**	1.65	1.99	1.59	13.38	14.45	16.00	14.42	18.90	20.48	23.01	20.47	13.50	14.41	15.81	14.40
	RMSE	2.73	3.73	4.60	3.68	23.25	25.31	27.73	25.19	30.84	33.50	37.27	33.47	21.63	23.45	25.77	23.39
	MAPE	2.71%	3.73%	4.71%	3.59%	13.82%	14.84%	16.03%	14.64%	7.89%	8.58%	9.81%	8.61%	8.63%	9.24%	10.11%	9.21%
DCRNN	MAE	1.31	1.65	1.97	1.59	14.29	15.19	17.51	15.53	19.46	21.13	24.1	21.16	14.12	15.23	16.95	15.22
	RMSE	2.76	3.75	4.60	3.69	24.75	26.55	30.34	27.18	31.22	34.14	38.46	34.14	22.12	24.18	26.95	24.17
	MAPE	2.73%	3.71%	4.68%	3.58%	14.88%	15.33%	17.31%	15.62%	8.27%	8.95%	10.37%	9.02%	9.48%	10.21%	11.37%	10.21%
AGCRN	MAE	1.35	1.67	1.94	1.61	14.22	15.34	16.86	15.29	19.25	20.58	22.68	20.57	14.26	15.24	17.03	15.32
	RMSE	2.88	3.82	4.50	3.70	25.01	26.99	29.52	26.95	31.62	34.39	38.16	34.40	22.57	24.40	26.91	24.41
	MAPE	2.91%	3.81%	4.55%	3.64%	13.88%	15.86%	15.95%	15.15%	8.18%	8.69%	9.67%	8.74%	9.39%	9.85%	11.37%	10.03%
STGCN	MAE	1.36	1.70	2.02	1.64	14.61	15.57	17.44	15.65	20.28	21.68	24.26	21.74	14.98	15.97	17.87	16.08
	RMSE	2.88	3.84	4.63	3.76	25.49	27.27	30.01	27.31	32.55	35.15	39.36	35.27	23.52	25.30	28.03	25.39
	MAPE	2.86%	3.79%	4.72%	3.67%	14.33%	15.07%	17.18%	15.39%	8.63%	9.18%	10.29%	9.24%	9.85%	10.51%	11.71%	10.60%
MTGNN	MAE	1.33	1.66	1.95	1.60	13.71	14.87	16.44	14.85	19.32	20.88	23.45	20.89	14.16	15.15	16.78	15.18
	RMSE	2.80	3.77	4.50	3.66	23.04	25.49	28.16	25.32	31.23	34.04	38.08	34.06	22.37	24.26	26.83	24.24
	MAPE	2.81%	3.75%	4.62%	3.59%	14.36%	15.12%	16.01%	14.97%	8.31%	8.95%	10.18%	9.00%	9.67%	10.81%	11.53%	10.20%
DGCRN	MAE	**1.30**	**1.61**	1.95	1.57	13.46	14.67	16.36	14.59	18.70	20.22	22.46	20.17	13.44	14.52	16.21	14.42
	RMSE	**2.72**	3.67	4.54	3.62	23.74	26.06	28.83	25.92	30.59	33.36	36.80	33.32	21.57	23.58	26.37	23.58
	MAPE	**2.69%**	**3.59%**	4.54%	3.48%	14.53%	15.44%	16.84%	15.32%	7.81%	8.39%	9.44%	8.42%	8.63%	9.40%	10.70%	9.46%
DFDGCN	MAE	**1.30**	**1.61**	1.90	1.55	13.27	14.33	15.80	14.28	18.64	20.11	22.17	20.03	13.39	14.23	15.56	14.23
	RMSE	**2.72**	3.66	4.39	3.57	23.47	25.52	27.92	25.39	30.72	33.41	36.70	33.24	21.53	23.34	25.65	23.30
	MAPE	2.70%	3.62%	4.44%	3.46%	13.41%	14.90%	15.94%	14.45%	7.74%	8.31%	9.24%	8.31%	8.52%	9.04%	9.95%	9.07%

Table 3. Ablation experiments on graph convolution on PEMS08

Graph	Metric	@3	@6	@12	Avg.
P	MAE	14.12	15.20	17.14	15.25
	RMSE	22.18	24.17	27.07	24.19
	MAPE	9.41%	9.95%	11.12%	10.09%
SA	MAE	13.54	14.47	15.88	14.45
	RMSE	21.63	23.44	25.72	23.36
	MAPE	9.02%	9.55%	10.44%	9.53%
D	MAE	13.50	14.32	15.73	14.33
	RMSE	21.84	23.71	26.04	23.61
	MAPE	8.84%	9.51%	10.42%	9.48%
T	MAE	13.65	14.58	16.07	14.56
	RMSE	22.03	24.07	26.67	24.01
	MAPE	8.64%	9.31%	10.45%	9.35%
P+SA	MAE	13.50	14.41	15.81	14.40
	RMSE	21.63	23.45	25.77	23.39
	MAPE	8.63%	9.24%	10.11%	9.21%
D+P	MAE	13.51	14.34	15.80	14.36
	RMSE	21.68	23.54	26.05	23.52
	MAPE	8.87%	9.42%	10.43%	9.46%
D+SA	MAE	**13.32**	**14.17**	**15.45**	**14.16**
	RMSE	**21.53**	**23.24**	**25.33**	**23.15**
	MAPE	8.55%	9.10%	10.07%	9.13%
D+P+SA	MAE	13.39	14.23	15.56	14.23
	RMSE	**21.53**	23.34	25.65	23.30
	MAPE	**8.52%**	**9.04%**	**9.95%**	**9.07%**

Tab. 3 presents the ablation experiment on graph convolution in our model to explore which graph plays a more important role in traffic prediction since both static predefined graphs as well as adaptive graphs are applied in DFDGCN. P refers to predefined graphs and SA refer to the self-adaptive graph, they are static graphs which preserve static spatial relationships in traffic network. T is called time-domain graph, which directly applies traffic data in the time domain for the graph construction in Section 2.2 without Fourier transform,

and D refers to our proposed frequency domain graph.

From the results of the convolution of a single graph, our dynamic frequency domain graph works the best, indicating that we effectively mine the dynamic spatial dependence. The comparison with T practically illustrates the serious impact of time-shift on data-driven spatial dependence mining and our effectiveness in analyzing spatial dependence from the frequency domain. When multiple graphs are convolved, the frequency-domain graph always brings more improvement. However, the effect always deteriorates when convolving the frequency domain graph with predefined graphs. We conjecture that it may be attributed to these factors: (1). The static priori information of predefined graphs play a negative role in traffic prediction; (2). Multiple graphs convolution may suffer from problems such as different information densities of individual graphs, inconsistency of convergence speed with the direction of hyperparameter gradient descent similar to that in multitask learning, which is a separate topic of research.

4. CONCLUSION

In this paper, we focus on spatial dependence in traffic prediction and propose a novel model DFDGCN. We argue that data-driven spatial dependence modeling suffers from time-shift and data noise, thus we propose to analyze traffic data in the frequency domain and construct an effective graph structure to mine spatial correlation in transportation network. Experiments support our argument and demonstrate that our model outperforms other baselines. In subsequent research, we will focus on exploring spatial dependence between sensor signals and optimizing multi-graph convolution.

5. REFERENCES

[1] Mohammad Reza Jabbarpour, Houman Zarrabi, Rashid Hafeez Khokhar, Shahaboddin Shamshirband, and Kim-Kwang Raymond Choo, "Applications of computational intelligence in vehicle traffic congestion problem: a survey," *Soft Comput.*, vol. 22, no. 7, pp. 2299–2320, 2018.

[2] Yongjun Xu, Xin Liu, Xin Cao, and et al., "Artificial intelligence: A powerful paradigm for scientific research," *The Innovation*, vol. 2, no. 4, 2021.

[3] Chunggi Lee, Yeonjun Kim, Seungmin Jin, Dongmin Kim, Ross Maciejewski, David S. Ebert, and Sungahn Ko, "A visual analytics system for exploring, monitoring, and forecasting road traffic congestion," *IEEE Trans. Vis. Comput. Graph.*, vol. 26, no. 11, pp. 3133–3146, 2020.

[4] Chengqing Yu, Fei Wang, Zezhi Shao, Tao Sun, Lin Wu, and Yongjun Xu, "Dsformer: A double sampling transformer for multivariate time series long-term prediction," in *Proceedings of the 32nd ACM International Conference on Information and Knowledge Management*, 2023, pp. 3062–3072.

[5] Billy M Williams and Lester A Hoel, "Modeling and forecasting vehicular traffic flow as a seasonal arima process: Theoretical basis and empirical results," *Journal of transportation engineering*, vol. 129, no. 6, pp. 664–672, 2003.

[6] Xiaolei Ma, Zhimin Tao, Yinhai Wang, Haiyang Yu, and Yunpeng Wang, "Long short-term memory neural network for traffic speed prediction using remote microwave sensor data," *Transportation Research Part C: Emerging Technologies*, vol. 54, pp. 187–197, 2015.

[7] Yaguang Li, Rose Yu, Cyrus Shahabi, and Yan Liu, "Diffusion convolutional recurrent neural network: Data-driven traffic forecasting," in *International Conference on Learning Representations*, 2018.

[8] Zonghan Wu, Shirui Pan, Guodong Long, Jing Jiang, and Chengqi Zhang, "Graph wavenet for deep spatial-temporal graph modeling," in *IJCAI*, 2019, pp. 1907–1913.

[9] Bing Yu, Haoteng Yin, and Zhanxing Zhu, "Spatio-temporal graph convolutional networks: a deep learning framework for traffic forecasting," in *IJCAI*, 2018, pp. 3634–3640.

[10] Lei Bai, Lina Yao, Can Li, Xianzhi Wang, and Can Wang, "Adaptive graph convolutional recurrent network for traffic forecasting," *Advances in neural information processing systems*, vol. 33, pp. 17804–17815, 2020.

[11] Zonghan Wu, Shirui Pan, Guodong Long, Jing Jiang, Xiaojun Chang, and Chengqi Zhang, "Connecting the dots: Multivariate time series forecasting with graph neural networks," in *Proceedings of the 26th ACM SIGKDD international conference on knowledge discovery & data mining*, 2020, pp. 753–763.

[12] Shengnan Guo, Youfang Lin, Huaiyu Wan, Xiucheng Li, and Gao Cong, "Learning dynamics and heterogeneity of spatial-temporal graph data for traffic forecasting," *IEEE Transactions on Knowledge and Data Engineering*, vol. 34, no. 11, pp. 5415–5428, 2021.

[13] Liangzhe Han, Bowen Du, Leilei Sun, Yanjie Fu, Yisheng Lv, and Hui Xiong, "Dynamic and multi-faceted spatio-temporal deep learning for traffic speed forecasting," in *Proceedings of the 27th ACM SIGKDD conference on knowledge discovery & data mining*, 2021, pp. 547–555.

[14] Yuyol Shin and Yoonjin Yoon, "Pgcn: Progressive graph convolutional networks for spatial-temporal traffic forecasting," *arXiv preprint arXiv:2202.08982*, 2022.

[15] Mengzhang Li and Zhanxing Zhu, "Spatial-temporal fusion graph neural networks for traffic flow forecasting," in *AAAI*, 2021, vol. 35, pp. 4189–4196.

[16] Fuxian Li, Jie Feng, Huan Yan, Guangyin Jin, Fan Yang, Funing Sun, Depeng Jin, and Yong Li, "Dynamic graph convolutional recurrent network for traffic prediction: Benchmark and solution," *ACM Transactions on Knowledge Discovery from Data*, vol. 17, no. 1, pp. 1–21, 2023.

[17] Shiyong Lan, Yitong Ma, Weikang Huang, Wenwu Wang, Hongyu Yang, and Pyang Li, "Dstagnn: Dynamic spatial-temporal aware graph neural network for traffic flow forecasting," in *International conference on machine learning*. PMLR, 2022, pp. 11906–11917.

[18] Zezhi Shao, Zhao Zhang, Fei Wang, Wei Wei, and Yongjun Xu, "Spatial-temporal identity: A simple yet effective baseline for multivariate time series forecasting," in *CIKM*, 2022, pp. 4454–4458.

[19] Zezhi Shao, Zhao Zhang, Wei Wei, Fei Wang, Yongjun Xu, Xin Cao, and Christian S. Jensen, "Decoupled dynamic spatial-temporal graph neural network for traffic forecasting," *Proc. VLDB Endow.*, vol. 15, no. 11, pp. 2733–2746, 2022.

[20] Yue Cui, Jiandong Xie, and Kai Zheng, "Historical inertia: A neglected but powerful baseline for long sequence time-series forecasting," in *Proceedings of the 30th ACM international conference on information & knowledge management*, 2021, pp. 2965–2969.