2023 AAAI

PDFormer

Propagation **D**elay-aware Dynamic Long-range Trans**Former** for Traffic Flow Prediction

23.3.2

Presented by Yyyq

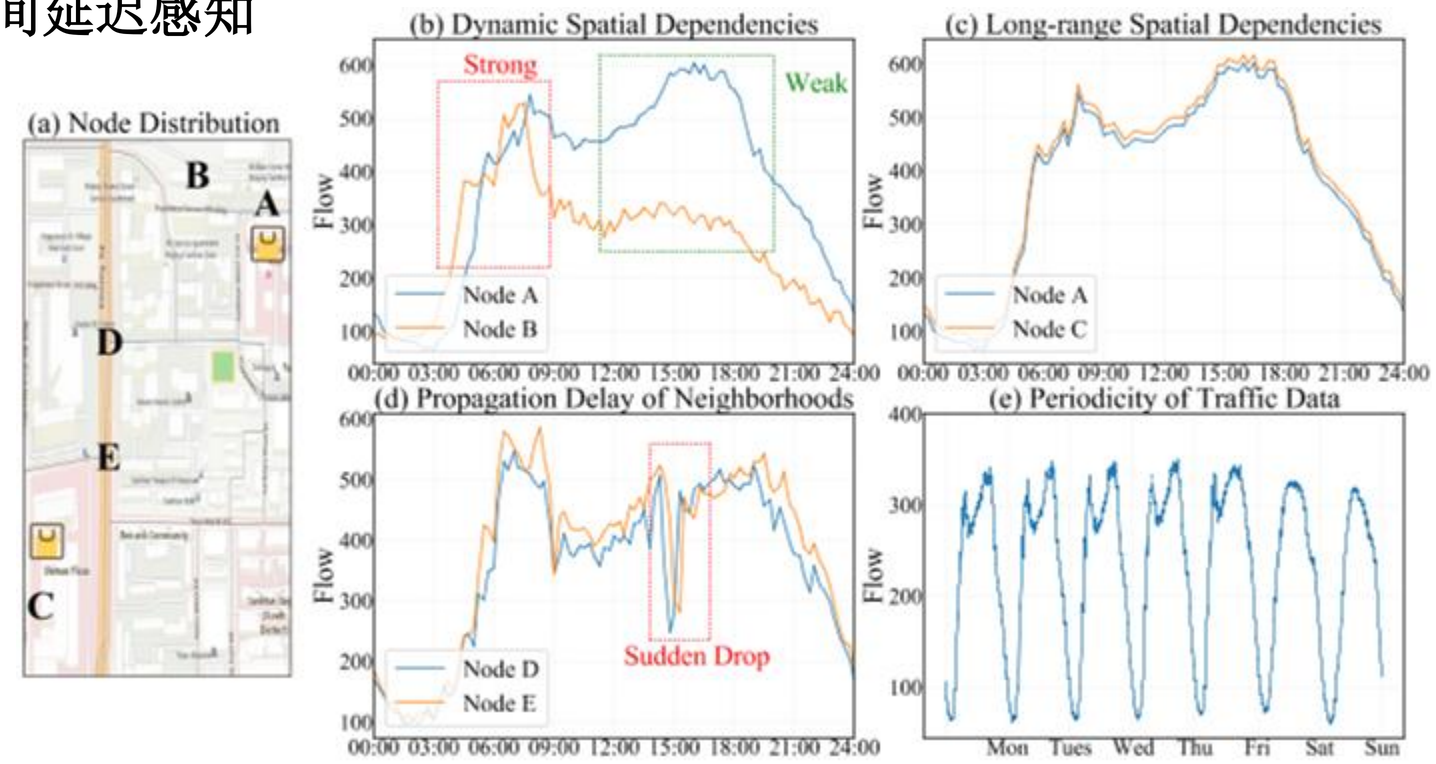
00

问题描述：PDFormer

1、Dynamic动态

2、Long-range长距离

3、Propagation Delay-aware传播时间延迟感知

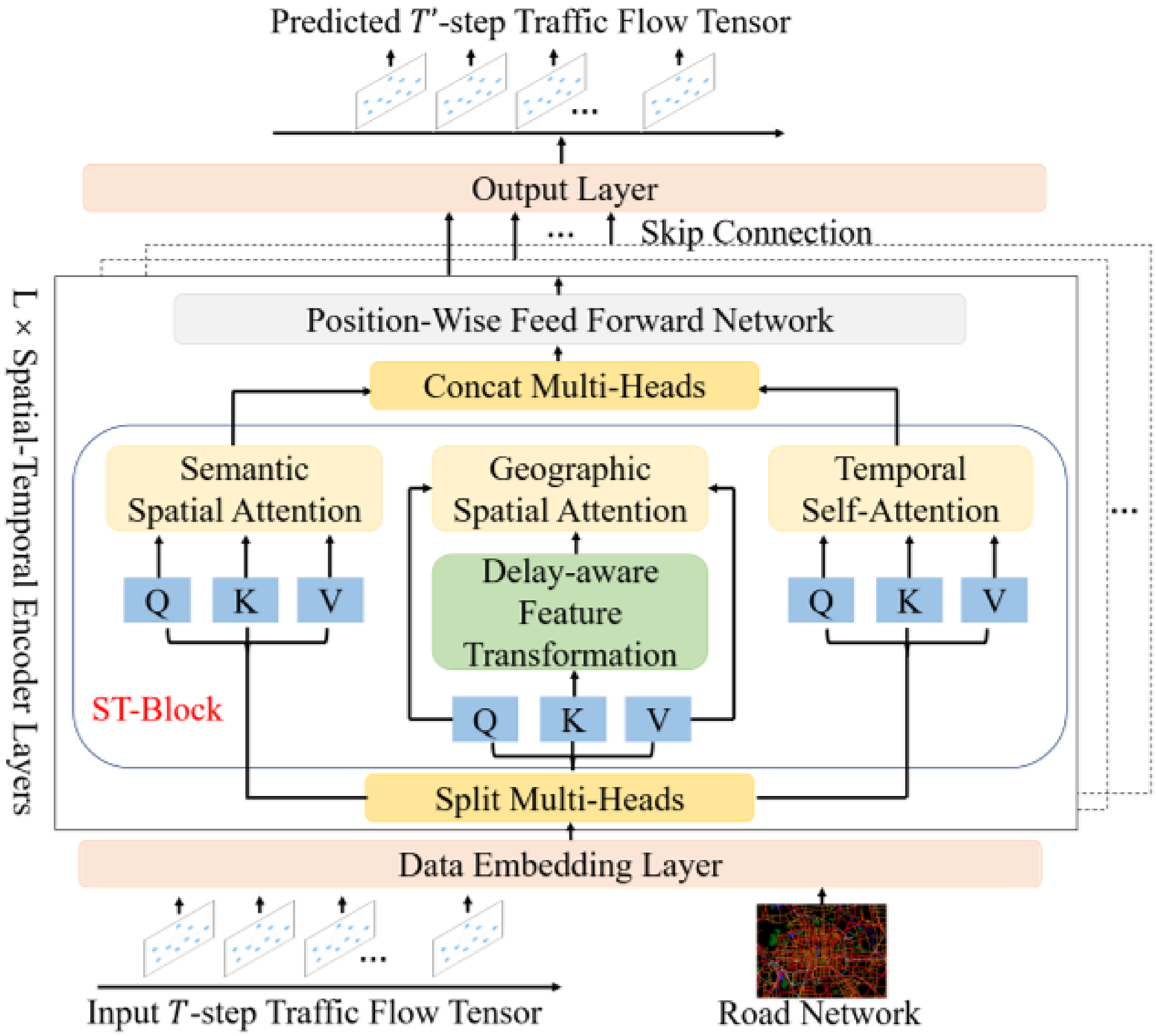


01 创新点

- **Spatial Self-attention：**捕获空间依赖关系
  - local geographic neighborhood
  - global semantic neighborhood
- **Delay-aware Feature Transformation：**集成历史交通模式
- **Temporal Self-attention：**识别时间动态性

02

模型结构



Predicted T'-step Traffic Flow Tensor
Output Layer
Skip Connection
Position-Wise Feed Forward Network
Concat Multi-Heads
Semantic Spatial Attention
Geographic Spatial Attention
Temporal Self-Attention
Q
K
V
Delay-aware Feature Transformation
ST-Block
Split Multi-Heads
L × Spatial-Temporal Encoder Layers
Data Embedding Layer
Input T-step Traffic Flow Tensor
Road Network

03 算法描述：Data Embedding Layer

➢ Spatial Graph Laplacian Embedding

对无向图G进行编码， 使得G蕴含的信息可计算。

Graph Laplacian 的最小非零特征值
对应的特征向量是图G最优的一维编码

$$G = (V, A) \Rightarrow \triangle = I - D^{-\frac{1}{2}} A D^{-\frac{1}{2}} \Rightarrow \triangle = U^T \wedge U \Rightarrow X_{spe} \in R^{N \times d}$$

➢ Temporal Periodic Embedding

$$X_w, X_d \in R^{T \times d}$$

$$\mathcal{X}_{emb} = \mathcal{X}_{data} + X_{spe} + X_w + X_d + X_{tpe}$$

*Graph Laplacian：https://zhuanlan.zhihu.com/p/56568843

03 算法描述：Spatial-Temporal Encoder Layer

➢ 空间自注意力(Spatial Self-Attention, SSA)模块

$$\begin{cases} Q_t^{(S)} = X_{t::}W_Q^S, K_t^{(S)} = X_{t::}W_K^S, V_t^{(S)} = X_{t::}W_V^S \\ A_t^{(S)} = \dfrac{(Q_t^{(S)})(K_t^{(S)})^\top}{\sqrt{d'}} \in \mathbb{R}^{N\times N} \\ SSA(Q_t^{(S)}, K_t^{(S)}, V_t^{(S)}) = softmax(A_t^{(S)})V_t^{(S)} \end{cases}$$

$$GeoSSA(Q_t^{(S)}, K_t^{(S)}, V_t^{(S)}) = softmax(A_t^{(S)} \odot M_{geo})V_t^{(S)}$$

$$SemSSA(Q_t^{(S)}, K_t^{(S)}, V_t^{(S)}) = softmax(A_t^{(S)} \odot M_{sem})V_t^{(S)}$$

- **geographic masking matrix** M_{geo} ：依据距离的远近
- **semantic masking matrix** M_{sem} ：DTW算法，计算节点之间历史交通流量的相似性

03

算法描述：Spatial-Temporal Encoder Layer

- **semantic masking matrix** M_{sem} ：DTW算法，计算节点之间历史交通流量的相似性

例子：计算序列A（1-1-3-3-2-4）和序列B（1-3-2-2-4-4）两个序列的相似性

	A(1)=1	A(2) =1	A(3) =3	A(4) =3	A(5) =2	A(6) =4
B(1) =1	0	0	2	2	1	3
B(2) =3	2	2	0	0	1	1
B(3) =2	1	1	1	1	0	2
B(4) =2	1	1	1	1	0	2
B(5) =4	3	3	1	1	2	0
B(6) =4	3	3	1	1	2	0

03

算法描述：Spatial-Temporal Encoder Layer

➢ 空间自注意力(Spatial Self-Attention, SSA)模块

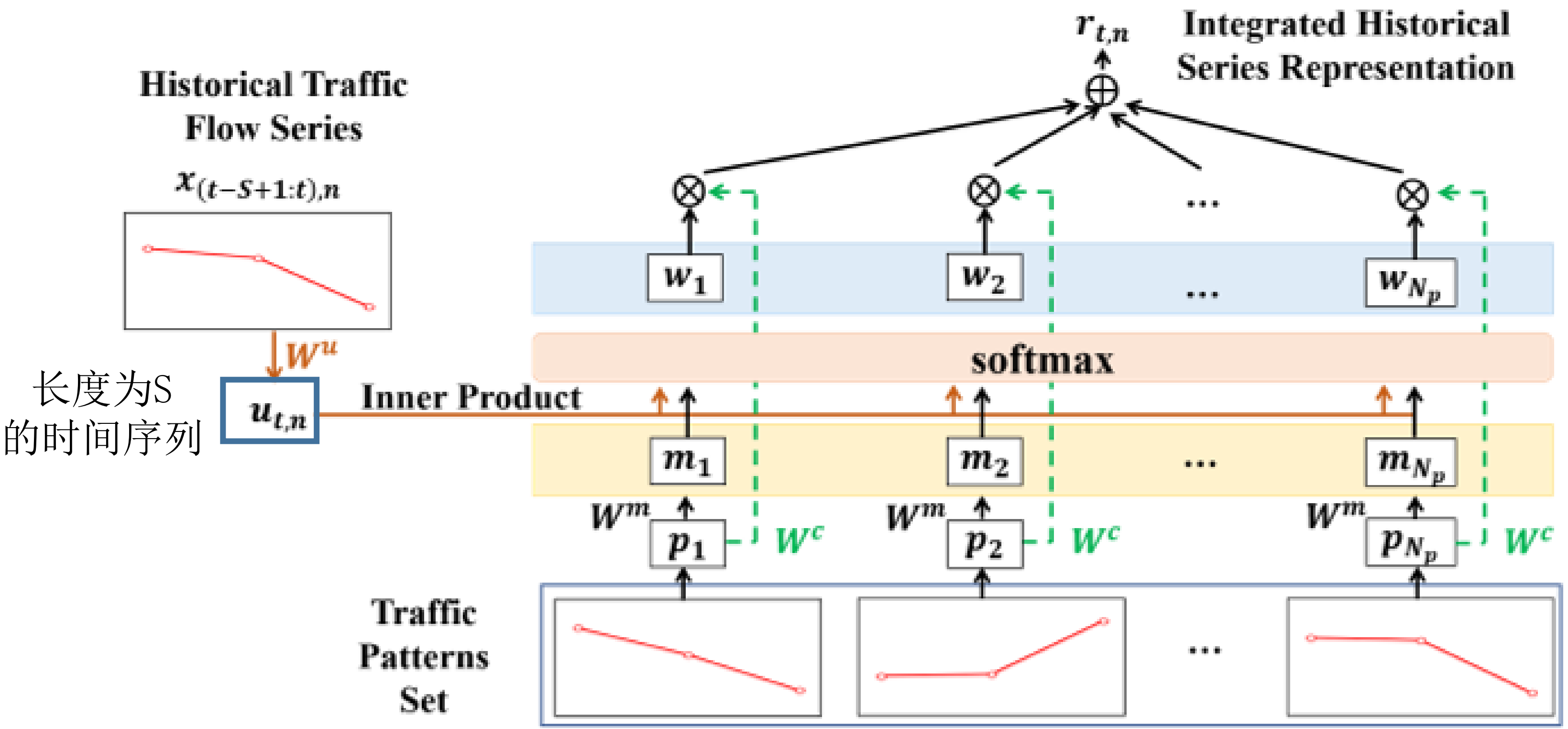
$$
\begin{cases}
Q_t^{(S)} = X_{t::}W_Q^S, K_t^{(S)} = X_{t::}W_K^S, V_t^{(S)} = X_{t::}W_V^S \\
A_t^{(S)} = \dfrac{(Q_t^{(S)})(K_t^{(S)})^\top}{\sqrt{d'}} \in \mathbb{R}^{N\times N} \\
\begin{cases}
GeoSSA(Q_t^{(S)}, K_t^{(S)}, V_t^{(S)}) = softmax(A_t^{(S)} \odot M_{geo})V_t^{(S)} \\
SemSSA(Q_t^{(S)}, K_t^{(S)}, V_t^{(S)}) = softmax(A_t^{(S)} \odot M_{sem})V_t^{(S)}
\end{cases}
\end{cases}
$$

- **geographic masking matrix** M_{geo} ：依据距离的远近
- **semantic masking matrix** M_{sem} ：DTW算法，计算节点之间历史交通流量的相似性

03 算法描述：Spatial-Temporal Encoder Layer

➢ 延迟感知特征转换(Delay-aware Feature Tansformation, DFT)模块

K-shape聚类算法



$$r_{t,n} = \sum_{i=1}^{N_p} w_i (p_i W^c)$$

$$w_i = softmax(u_{t,n}^\top m_i)$$

$$m_i = p_i W^m$$

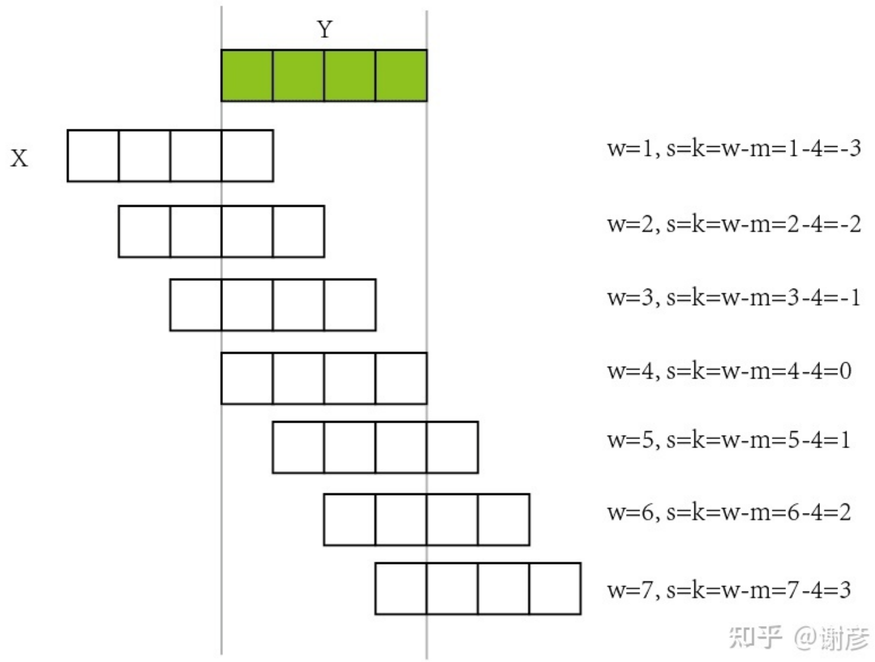
$$\mathcal{P} = \{p_i | i \in [1, \ldots, N_p]\}$$

03 算法描述：Spatial-Temporal Encoder Layer

➢ **延迟感知特征转换(Delay-aware Feature Transformation, DFT)模块**

K-shape聚类算法

- **距离计算：**互相关算法，计算两个等长序列的相似性



块重叠越多，形状越像，
互相关系数CC越大，SBD形态距离越小

$$SBD(\vec{x},\vec{y}) = 1 - \max_{w}\left(\frac{CC_w(\vec{x},\vec{y})}{\sqrt{R_0(\vec{x},\vec{x})\cdot R_0(\vec{y},\vec{y})}}\right)$$

03 算法描述：Spatial-Temporal Encoder Layer

➢ **延迟感知特征转换(Delay-aware Feature Transformation, DFT)模块**

K-shape聚类算法

- **距离计算：**互相关算法，计算两个等长序列的相似性

$$SBD(\vec{x},\vec{y}) = 1 - \max_{w}\left(\frac{CC_w(\vec{x},\vec{y})}{\sqrt{R_0(\vec{x},\vec{x})\cdot R_0(\vec{y},\vec{y})}}\right)$$

- **质心计算：**使质心与其他时间序列的相似度NCC最小，与簇中各序列的NCC最大。

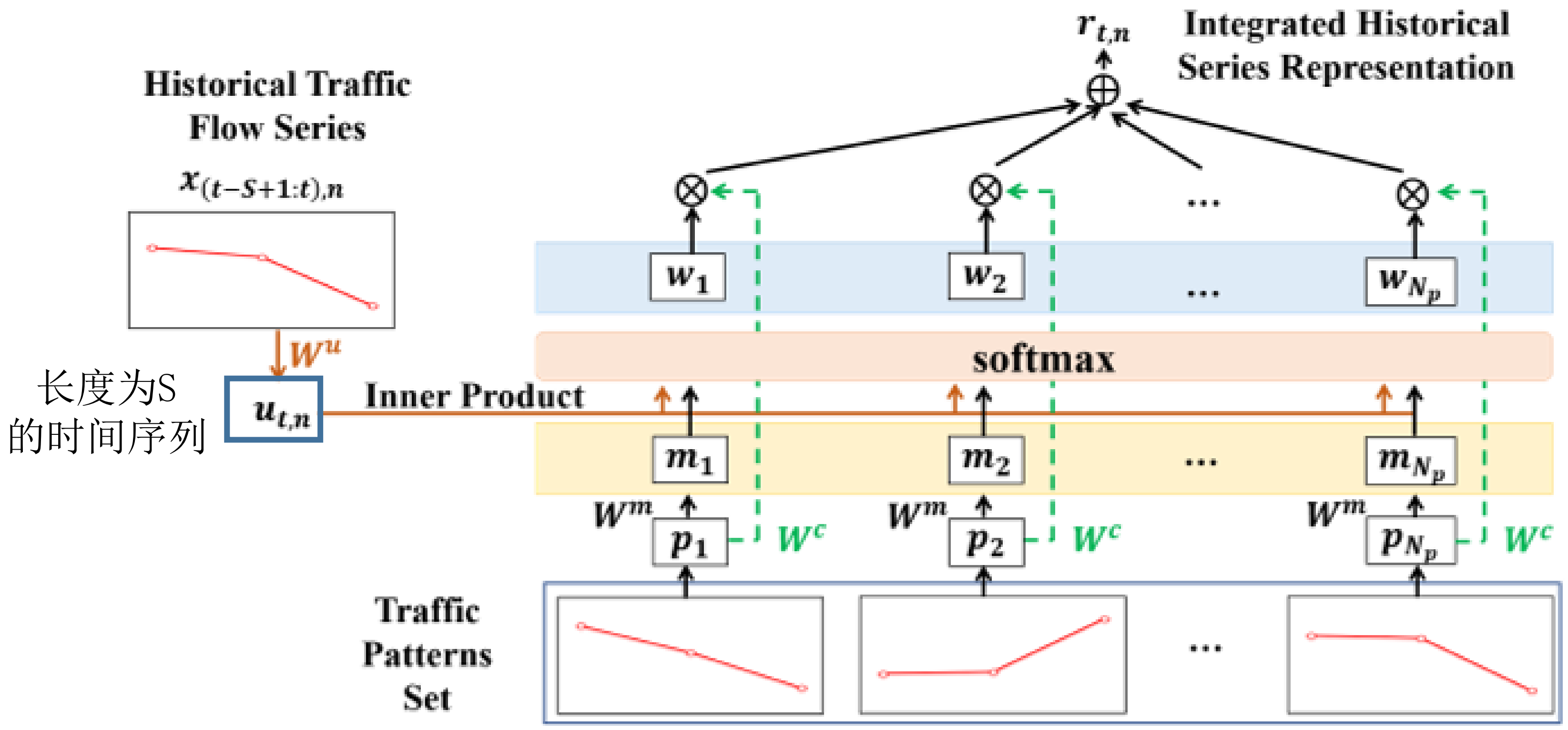
$$C_k = argmin \sum_{(\vec{x}_i \epsilon p_k)} NCC_c(\vec{x}_i, \vec{c})^2 = argmax \sum_{(\vec{x}_i \epsilon p_k)} \left(\frac{max_w CC_w(\vec{x}_i, \vec{c})}{\sqrt{R_0(\vec{x}_i, \vec{x}_i) * R_0(\vec{c}, \vec{c})}}\right)^2$$

- **迭代更新**

03 算法描述：Spatial-Temporal Encoder Layer

➢ 延迟感知特征转换(Delay-aware Feature Transformation, DFT)模块

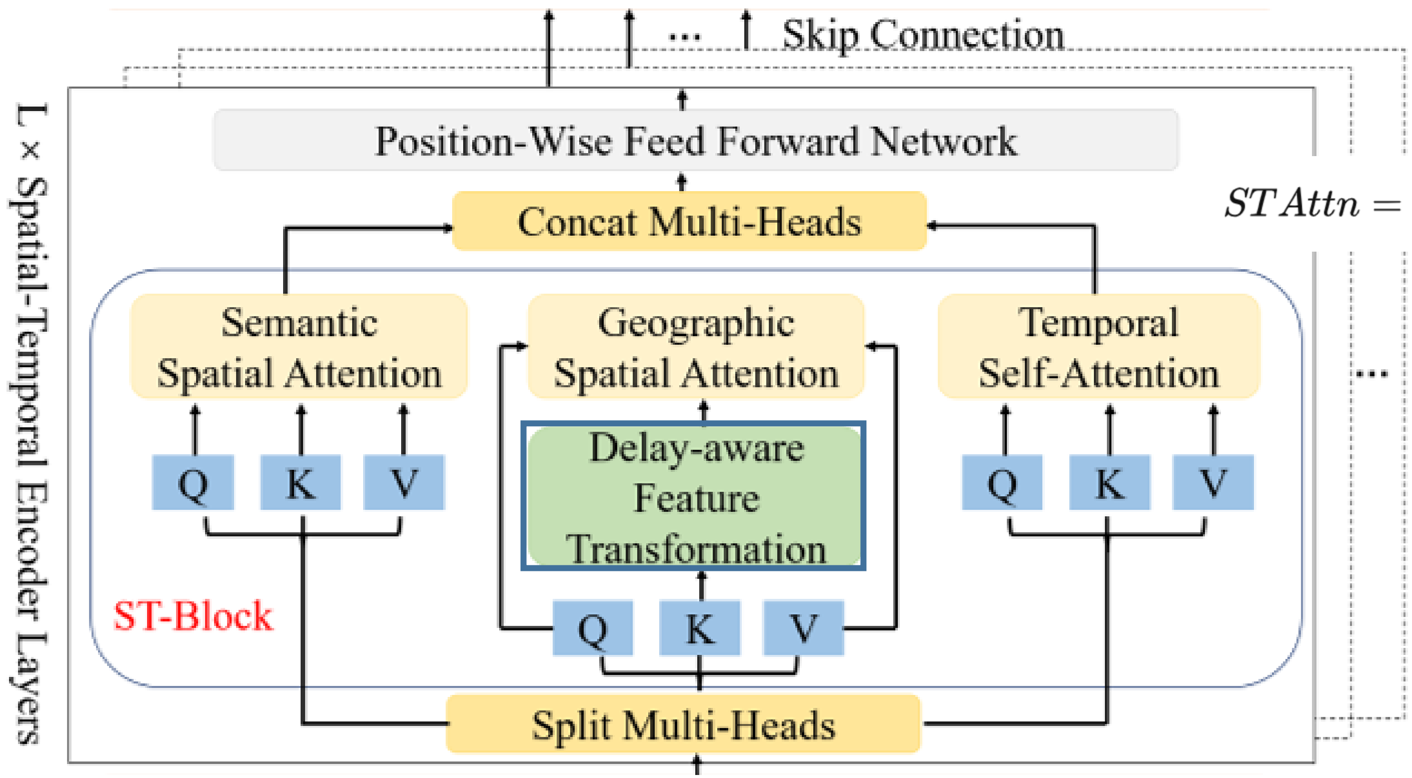
K-shape聚类算法



$$r_{t,n} = \sum_{i=1}^{N_p} w_i (p_i W^c)$$

$$\tilde{K}_t^{(S)} = K_t^{(S)} + R_t$$

03 算法描述：Spatial-Temporal Encoder Layer



$$STAttn = \oplus(Z^{geo}_{1...h_{geo}}, Z^{sem}_{1,...,h_{sem}}, Z^{t}_{1,...,h_t})W^O$$

03 算法描述：Output Layer

$$\hat{\mathcal{X}} = Conv_2(Conv_1(\mathcal{X}_{hid}) \in \mathbb{R}^{T' \times N \times C}$$

04 数据集

Datasets	#Nodes	#Edges	#Timesteps	#Time Interval	Time range
PeMS04	307	340	16992	5min	01/01/2018-02/28/2018
PeMS07	883	866	28224	5min	05/01/2017-08/31/2017
PeMS08	170	295	17856	5min	07/01/2016-08/31/2016
NYCTaxi	75 (15x5)	484	17520	30min	01/01/2014-12/31/2014
CHIBike	270 (15x18)	1966	4416	30min	07/01/2020-09/30/2020
T-Drive	1024 (32x32)	7812	3600	60min	02/01/2015-06/30/2015

05

实验结果1

Model	PeMS04			PeMS07			PeMS08		
	MAE	MAPE(%)	RMSE	MAE	MAPE(%)	RMSE	MAE	MAPE(%)	RMSE
VAR	23.750	18.090	36.660	101.200	39.690	155.140	22.320	14.470	33.830
SVR	28.660	19.150	44.590	32.970	15.430	50.150	23.250	14.710	36.150
DCRNN	22.737	14.751	36.575	23.634	12.281	36.514	18.185	11.235	28.176
STGCN	21.758	13.874	34.769	22.898	11.983	35.440	17.838	11.211	27.122
GWNET	19.358	13.301	31.719	21.221	9.075	34.117	15.063	9.514	24.855
MTGNN	19.076	12.961	31.564	20.824	9.032	34.087	15.396	10.170	24.934
STSGCN	21.185	13.882	33.649	24.264	10.204	39.034	17.133	10.961	26.785
STFGNN	19.830	13.021	31.870	22.072	9.212	35.805	16.636	10.547	26.206
STGODE	20.849	13.781	32.825	22.976	10.142	36.190	16.819	10.623	26.240
STGNCDE	19.211	12.772	31.088	20.620	8.864	34.036	15.455	9.921	24.813
STTN	19.478	13.631	31.910	21.344	9.932	34.588	15.482	10.341	24.965
GMAN	19.139	13.192	31.601	20.967	9.052	34.097	15.307	10.134	24.915
TFormer	18.916	12.711	31.349	20.754	8.972	34.062	15.192	9.925	24.883
ASTGNN	18.601	12.630	31.028	20.616	8.861	34.017	14.974	9.489	24.710
PDFormer	**18.321**	**12.103**	**29.965**	**19.832**	**8.529**	**32.870**	**13.583**	**9.046**	**23.505**

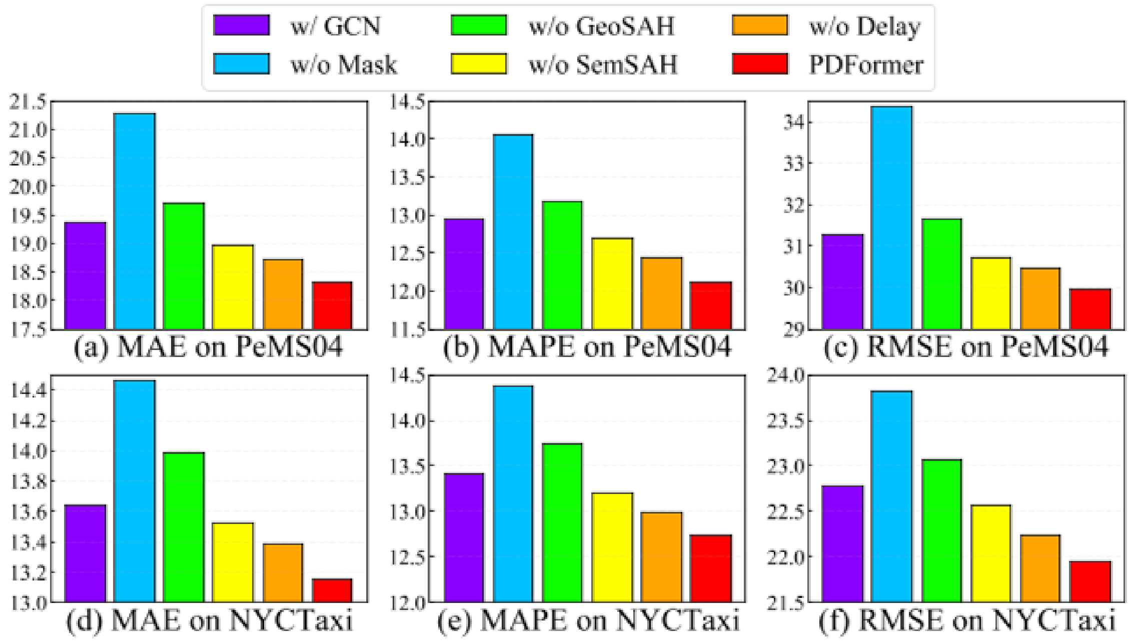
05

实验结果2

Datasets	NYCTaxi						T-Drive						CHIBike					
Metrics	inflow			outflow			inflow			outflow			inflow			outflow		
Models	MAE	MAPE(%)	RMSE	MAE	MAPE(%)	RMSE	MAE	MAPE(%)	RMSE	MAE	MAPE(%)	RMSE	MAE	MAPE(%)	RMSE	MAE	MAPE(%)	RMSE
STResNet	14.492	14.543	24.050	12.798	14.368	20.633	19.636	17.831	34.890	19.616	18.502	34.597	4.767	31.382	6.703	4.627	30.571	6.559
DMVSTNet	14.377	14.314	23.734	12.566	14.318	20.409	19.599	17.683	34.478	19.531	17.621	34.303	4.687	32.113	6.635	4.594	31.313	6.455
DSAN	14.287	14.208	23.585	12.462	14.272	20.294	19.384	17.465	34.314	19.290	17.379	34.267	4.612	31.621	6.695	4.495	31.256	6.367
DCRNN	14.421	14.353	23.876	12.828	14.344	20.067	22.121	17.750	38.654	21.755	17.382	38.168	4.236	31.264	5.992	4.211	30.822	5.824
STGCN	14.377	14.217	23.860	12.547	14.095	19.962	21.373	17.539	38.052	20.913	16.984	37.619	4.212	31.224	5.954	4.148	30.782	5.779
GWNET	14.310	14.198	23.799	12.282	13.685	19.616	19.556	17.187	36.159	19.550	15.933	36.198	4.151	31.153	5.917	4.101	30.690	5.694
MTGNN	14.194	13.984	23.663	12.272	13.652	19.563	18.982	17.056	35.386	18.929	15.762	35.992	4.112	31.148	5.807	4.086	30.561	5.669
STSGCN	15.604	15.203	26.191	13.233	14.698	21.653	23.825	18.547	41.188	24.287	19.041	42.255	4.256	32.991	5.941	4.265	32.612	5.879
STFGNN	15.336	14.869	26.112	13.178	14.584	21.627	22.144	18.094	40.071	22.876	18.987	41.037	4.234	32.222	5.933	4.264	32.321	5.875
STGODE	14.621	14.793	25.444	12.834	14.398	20.205	21.515	17.579	38.215	22.703	18.509	40.282	4.169	31.165	5.921	4.125	30.726	5.698
STGNCDE	14.281	14.171	23.742	12.276	13.681	19.608	19.347	17.134	36.093	19.230	15.873	36.143	4.123	31.151	5.913	4.094	30.595	5.678
STTN	14.359	14.206	23.841	12.373	13.762	19.827	20.583	17.327	37.220	20.443	15.992	37.067	4.160	31.208	5.932	4.118	30.704	5.723
GMAN	14.267	14.114	23.728	12.273	13.672	19.594	19.244	17.110	35.986	18.964	15.788	36.120	4.115	31.150	5.910	4.090	30.662	5.675
TFormer	13.995	13.912	23.487	12.211	13.611	19.522	18.823	16.910	34.470	18.883	15.674	35.219	4.071	31.141	5.878	4.037	30.647	5.638
ASTGNN	13.844	13.692	23.177	12.112	13.602	19.201	18.798	16.101	33.870	18.790	15.584	33.998	4.068	31.131	5.818	3.981	30.617	5.609
PDFormer	**13.152**	**12.743**	**21.957**	**11.575**	**12.820**	**18.394**	**17.832**	**14.711**	**31.606**	**17.743**	**14.649**	**31.501**	**3.950**	**30.214**	**5.559**	**3.837**	**29.914**	**5.402**

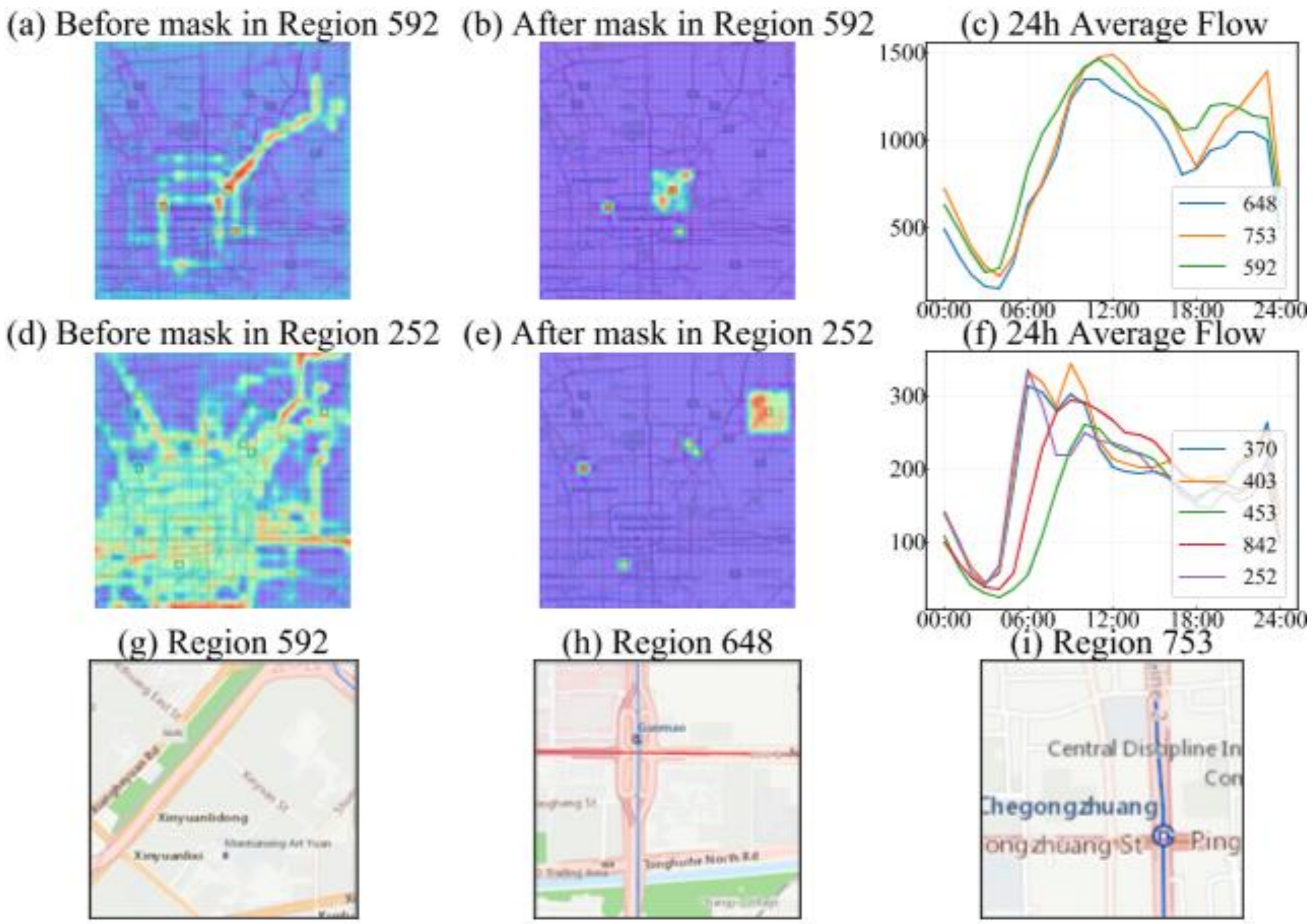
05

消融实验



Dataset	PeMS04		NYCTaxi	
Model	Training	Inference	Training	Inference
GMAN	501.578	38.844	130.672	4.256
ASTGNN	208.724	52.016	119.092	4.601
PDFormer	133.871	8.120	85.305	2.734
STTN	100.398	12.596	68.036	2.650
TFormer	71.099	7.156	76.169	2.575

06 案例分析


(a) Before mask in Region 592
(b) After mask in Region 592
(c) 24h Average Flow
1500
1000
500
648
753
592
00:00
06:00
12:00
18:00
24:00
(d) Before mask in Region 252
(e) After mask in Region 252
(f) 24h Average Flow
300
200
100
370
403
453
842
252
00:00
06:00
12:00
18:00
24:00
(g) Region 592
(h) Region 648
(i) Region 753
Xinyuanlidong
Xinyuanxili
Central Discipline In
Chegongzhuang
ongzhuang St
Ping

谢谢观看

MANY THANKS !

23.3.2