STP-TrellisNets+：用于多步地铁站客流量预测的时空并行TrellisNets

IEEE TRANSACTIONSONKNOWLEDGEANDDATAENGINEERING——2023

汇报人: 韦浩文

时间：2025-8-10

目录

1

背景

背景

为什么要研究地铁站客流量预测问题

在地铁乘客数量急剧增加的背景下，地铁站客（MSP）流量预测有助于应对地铁系统的拥挤问题，实现地铁系统高效管理。

为什么要研究多步MSP流量预测

1.现有研究主要集中在单步预测上，多步预测研究较少

2.多步预测比单步预测更有实用价值

- 延长预测时长
- 揭示更细粒度的客流量变化

背景

MSP流量预测具有较强的空间-时间相关性，即一个站点的MSP流量既和过去的流量在时间上相关，也和其他站点的MSP流量在空间上相关

STP-TrellisNets+

首次将时间卷积框架TrellisNet应用于多步MSP流量预测

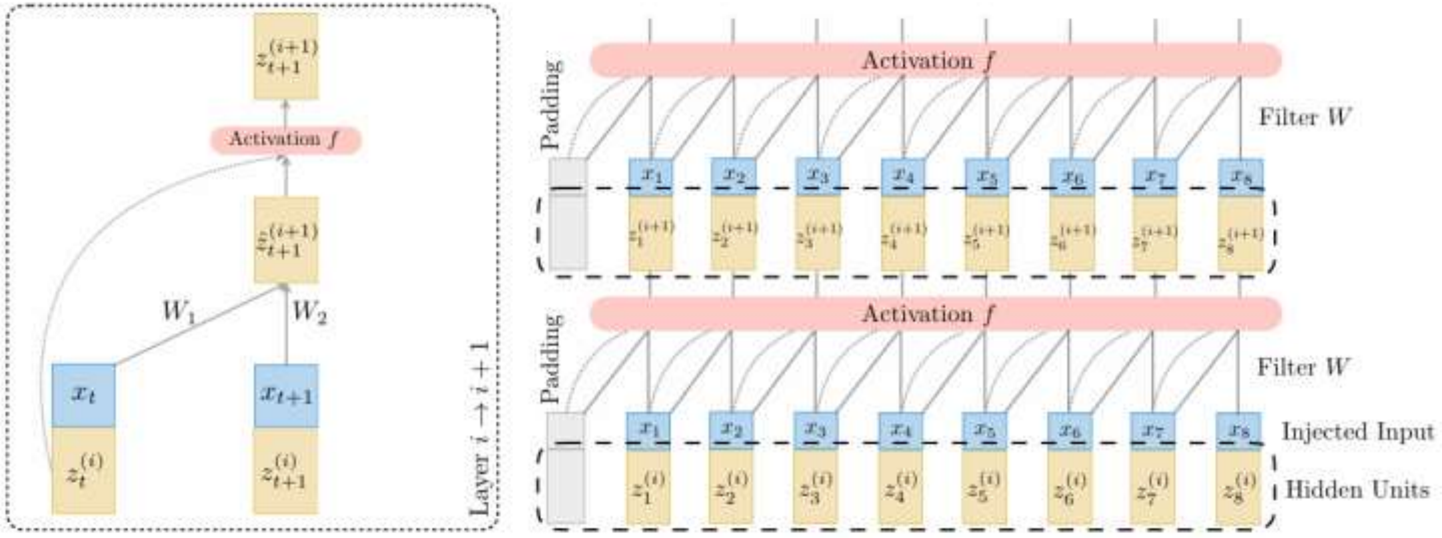
- ■ 时间模块（CP-TrellisNetsED）——时间相关性
  - ✓ C-TrellisNet——短期时间相关性
  - ✓ P-TrellisNetsED——长期时间相关性
- ■ 空间模块（GC-TrellisNetsED）——空间相关性
  - ✓ DGCN
  - ✓ D-TrellisNetsED

背景

TrellisNet

《TRELLIS NETWORKS FOR SEQUENCE MODELING 》ICLR 2019

兼具时间卷积网络和循环网络的特点



(a) TrellisNet at an atomic level

(b) TrellisNet on a sequence of units

1.特殊的时间卷积网络

- ➢ 跨层权重共享
- ➢ 直接输入注入

2.泛化截断循环网络

截断循环网络等价于权重矩阵有特殊稀疏结构的 TrellisNet

2 方法

STP-TrellisNets+

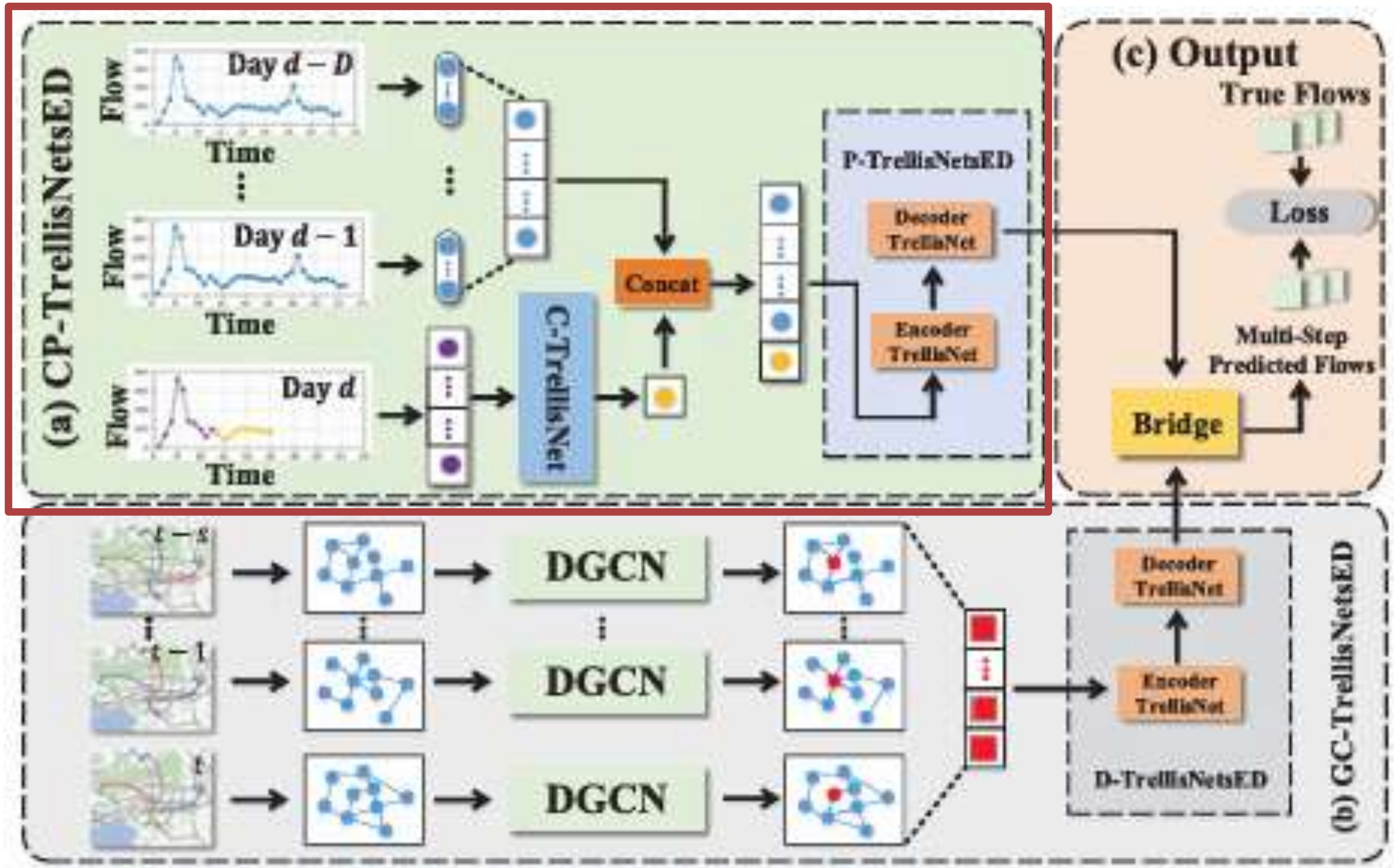


Fig. 3. The overall architecture of STP-TrellisNets+, where d represents the current day and FC denotes the Fully Connected layer. The MSP inflows and MSP outflows history are taken as input for MSP flow prediction.

CP-TrellisNetsED——C-TrellisNet

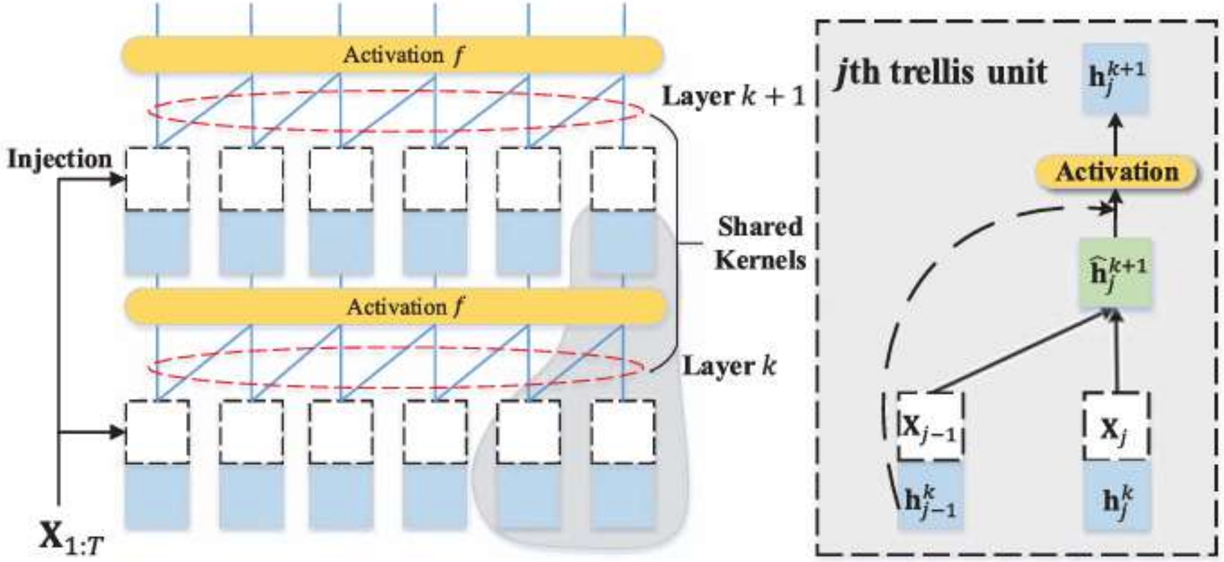


Fig. 4. An example of the TrellisNet architecture for MSP flow prediction.

$$\begin{cases} \widehat{\mathbf{h}}_j^{k+1} = W_1\left[\mathbf{X}_{j-1} \mid\mid \mathbf{h}_{j-1}^k\right] + W_2\left[\mathbf{X}_j \mid\mid \mathbf{h}_j^k\right] \\ \mathbf{h}_j^{k+1} = f\left(\widehat{\mathbf{h}}_j^{k+1}, \mathbf{h}_{j-1}^k\right) \end{cases}, \tag{1}$$

$$\mathbf{h}_{1:T}^{k+1} = f\left(\left(\mathbf{h}_{1:T}^k \mid\mid \mathbf{X}_{1:T}\right) * \mathbf{W},\ \mathbf{h}_{1:T-1}^k\right), \tag{2}$$

$$\Gamma^{(c)}(\mathbf{X}_{1:T}) = \mathbf{h}_T^{c+1}, \tag{3}$$

CP-TrellisNetsED——P-TrellisNetsED

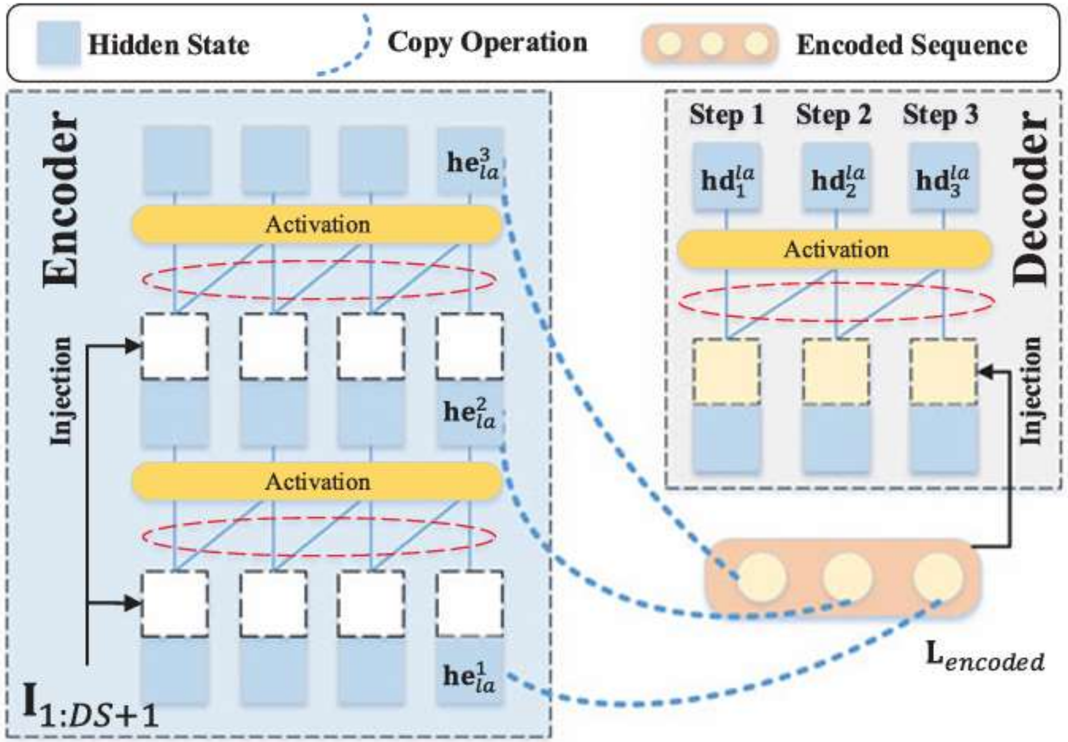


Fig. 5. An example of the TrellisNetsED architecture for multi-step MSP flow prediction.

$$\mathbf{P}_{1:DS} = \left(\underbrace{\overbrace{\mathbf{x}_{t+1-Dm}, \cdots, \mathbf{x}_{t+S-Dm}}^{S}, \cdots, \overbrace{\mathbf{x}_{t+1-m} \cdots, \mathbf{x}_{t+S-m}}^{S}}_{DS} \right)$$

$$\mathbf{I}_{1:DS+1} = \mathbf{P}_{1:DS} \mid\mid \mathbf{h}_{T}^{c+1}$$

$$\mathbf{he}_{1:DS+1}^{S} = f\left(\left(\mathbf{he}_{1:DS+1}^{S-1} \mid\mid \mathbf{I}_{1:DS+1}\right) * \mathbf{W}_1,\ \mathbf{he}_{1:DS}^{S-1}\right),$$

$$\mathbf{L}_{encoded} = \mathrm{Encoder}(\mathbf{I}_{1:DS+1}) = \{\underbrace{\mathbf{he}_{la}^{S}, \mathbf{he}_{la}^{S-1}, \ldots, \mathbf{he}_{la}^{1}}_{S}\},$$

$$\begin{aligned}\mathrm{Decoder}(\mathbf{L}_{encoded}) &= \mathbf{hd}_{1:S}^{la} \\ &= f\left(\left(\mathbf{hd}_{1:S}^{la-1} \mid\mid \mathbf{L}_{encoded}\right) * \mathbf{W}_2, \boxed{\mathbf{hd}_{1:S}^{la-1}}\right),\end{aligned}$$

$$\widehat{\mathbf{X}}_{1:S} = \Upsilon^{(S)}\left(\mathbf{P}_{1:DS} \mid\mid \Gamma^{(c)}(\mathbf{X}_{1:T})\right), \quad (4)$$

GC-TrellisNetsED

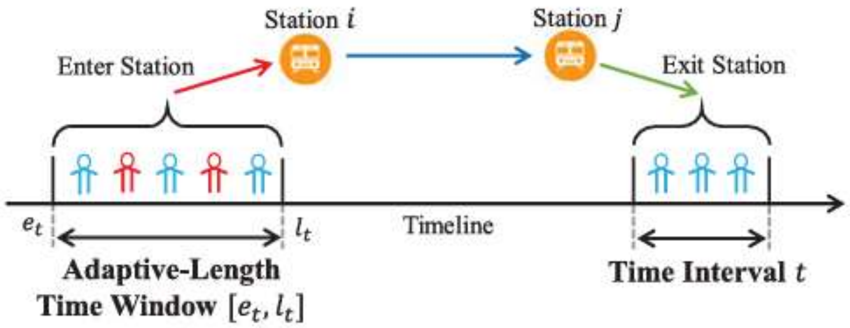


Fig. 6. Illustration example of transfer flow-based metric, where a blue person represents a passenger exiting station j during time interval t who takes the metro from station i, and a red person represents a passenger who enters station i during time window $[e_t, l_t]$, but does not exit at station j.

$$\xi_{ij}^{t} = \frac{A_{ij}^{t}}{O_{i}^{t}}, \tag{5}$$

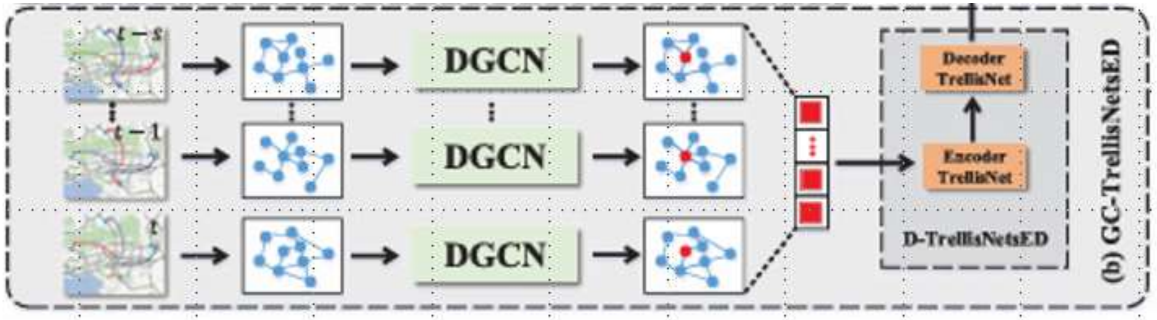
$$\mathcal{G}_t = (\mathcal{V}, \mathcal{E}_t, \mathbf{M}_t)$$

$$\mathbf{M}_t = \{a_{ij}^{t}\}$$

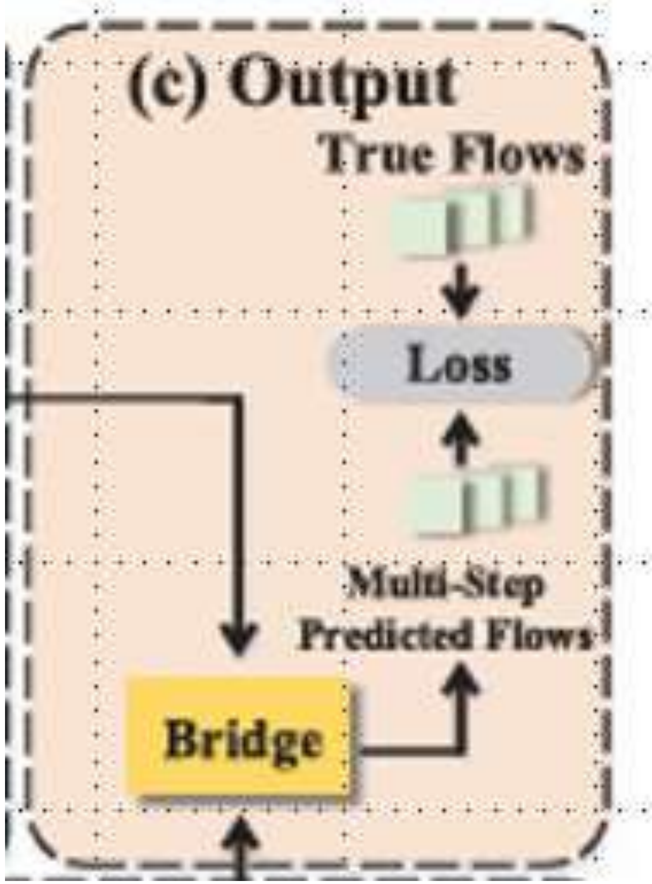
$$a_{ij}^{t} = \begin{cases} \xi_{ij}^{t}, & \text{if} \quad i \neq j \\ 0, & \text{if} \quad i = j \end{cases}. \tag{6}$$

$$\mathbf{Y}_{1:n,t} = \left(f_{1,t}^{\text{in}}, f_{2,t}^{\text{in}}, \ldots, f_{n,t}^{\text{in}}\right)^{\text{tr}},$$

$$\widehat{\mathbf{Y}}_{1:S} = \Upsilon^{(S)}\left(\Psi(\mathbf{M}_{t-s}, \mathbf{Y}_{1:n,t}), \ldots, \Psi(\mathbf{M}_t, \mathbf{Y}_{1:n,t})\right), \tag{8}$$



Output



$$\widehat{\mathbf{Z}}_{1:S} = \tanh\left(\mathbf{W}_x \circ \widehat{\mathbf{X}}_{1:S} + \mathbf{W}_y \circ \widehat{\mathbf{Y}}_{1:S}\right),$$

$$Loss(\boldsymbol{\Theta}) = \frac{1}{bn}\sum_{a=1}^{b}\sum_{i=1}^{n}\sum_{\tau=1}^{S}\left(Z_{i,t+\tau}^{a} - \widehat{Z}_{i,t+\tau}^{a}\right)^2,$$

3

实验

实验-数据集、评估指标

数据集

TABLE 2
Detailed Information of the Evaluated Datasets

Properties	Datasets: Shenzhen	Datasets: Hangzhou
# of Stations	165	81
# of Records	1.5 billion	70 million
Time span	6/1/2017 - 6/30/2017	1/1/2019 - 1/25/2019

评估指标

- 1.平均绝对误差

$$\mathrm{MAE} = \frac{1}{bn}\sum_{j=1}^{b}\sum_{i=1}^{n}\left|Z_{i,t+\tau}^{j} - \widehat{Z}_{i,t+\tau}^{j}\right|,$$

- 2.均方根误差

$$\mathrm{RMSE} = \sqrt{\frac{1}{bn}\sum_{j=1}^{b}\sum_{i=1}^{n}\left(Z_{i,t+\tau}^{j} - \widehat{Z}_{i,t+\tau}^{j}\right)^{2}},$$

- 3.平均绝对百分比误差

$$\mathrm{MAPE} = \frac{1}{bn}\sum_{j=1}^{b}\sum_{i=1}^{n}\left|\frac{Z_{i,t+\tau}^{j} - \widehat{Z}_{i,t+\tau}^{j}}{Z_{i,t+1}^{j}}\right|$$

实验-与真实值的比较

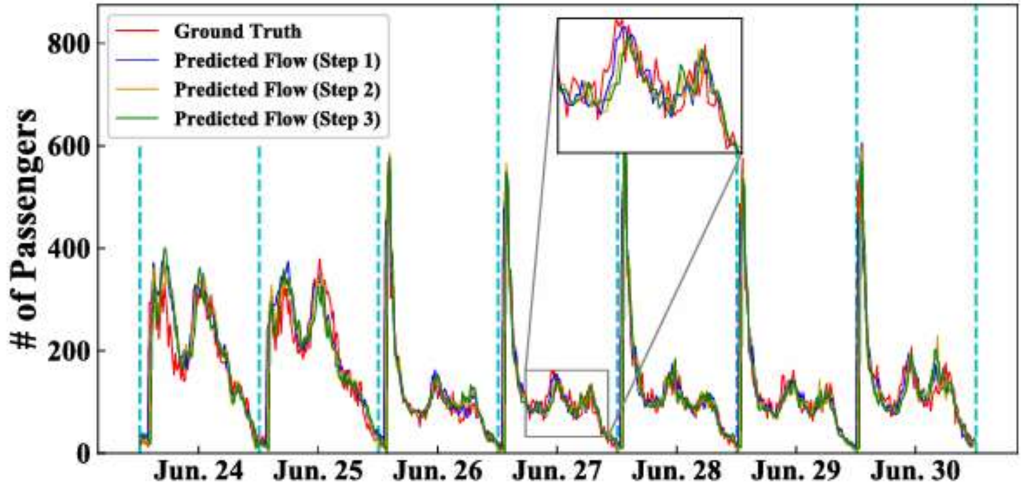


Fig. 8. Comparison of the multi-step MSP outflows predicted by STP-TrellisNets+ and the ground truth values of station Shaoniangong in the Shenzhen dataset.

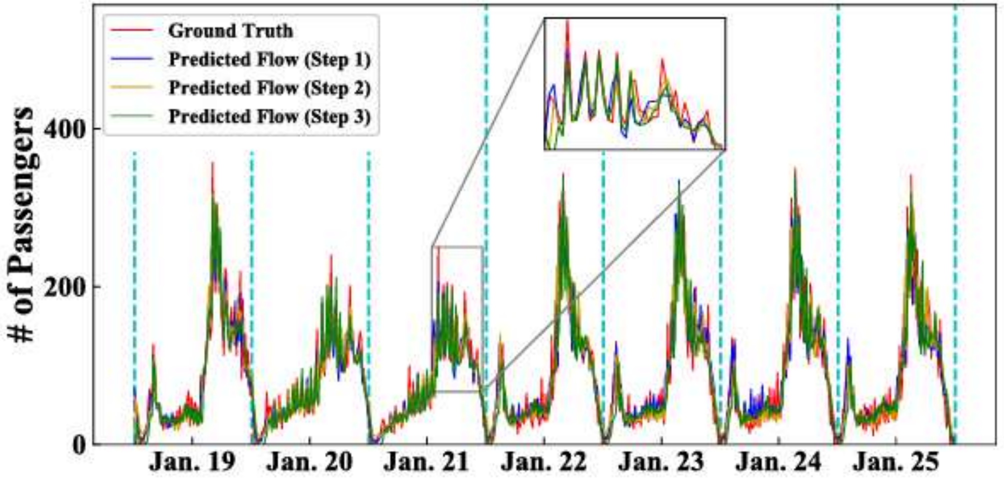


Fig. 9. Comparison of the multi-step MSP outflows predicted by STP-TrellisNets+ and the ground truth values of station Xiasha Jiangbin in the Hangzhou dataset.

实验-与基线方法的比较

TABLE 4
Multi-Step Prediction Performance Under 10-Minute Time Interval Setting

Datasets	Models	MAE			RMSE			MAPE (%)		
		Step 1	Step 2	Step 3	Step 1	Step 2	Step 3	Step 1	Step 2	Step 3
Shenzhen	Last-Value	22.62	36.78	47.63	42.80	50.08	61.23	30.79	31.62	32.33
	HA	35.66	35.89	36.07	76.82	76.91	76.99	52.28	52.23	52.21
	VAR	22.57	28.76	35.60	43.67	48.90	57.37	30.05	31.89	32.14
	MLP	19.70	25.32	32.43	31.42	39.43	46.76	26.57	27.80	28.36
	LSTM	17.86	25.44	31.28	30.17	38.41	45.88	25.77	26.43	27.81
	ConvLSTM	17.73	25.23	32.89	29.43	37.61	44.45	23.67	24.28	25.90
	DCRNN	16.75	24.87	31.08	27.03	36.86	43.91	21.57	22.49	23.67
	STGCN	15.57	16.14	17.15	26.79	27.46	28.57	20.71	21.75	23.94
	Graph WaveNet	15.57	15.77	16.39	26.04	26.35	27.29	20.89	21.33	22.53
	STP-TrellisNets+	**14.60**	**14.86**	**14.95**	**24.54**	**25.10**	**25.31**	**14.72**	**15.15**	**15.24**
Hangzhou	Last-Value	29.78	35.34	43.22	51.27	58.35	64.88	24.59	25.66	26.69
	HA	32.06	32.15	32.21	70.04	70.05	70.06	26.38	26.44	26.50
	VAR	34.63	40.57	46.28	55.77	60.39	69.70	24.56	25.77	26.29
	MLP	32.64	38.63	44.87	48.29	57.86	63.90	25.73	26.80	27.35
	LSTM	31.65	37.89	44.10	46.38	55.32	60.78	24.76	25.66	26.30
	ConvLSTM	30.62	37.88	43.75	40.15	48.28	55.73	22.33	23.67	25.19
	DCRNN	28.72	36.59	42.93	36.88	43.99	52.00	21.34	22.89	23.52
	STGCN	22.40	23.22	24.13	36.25	37.72	39.01	18.25	19.16	20.65
	Graph WaveNet	22.04	22.44	22.93	35.35	36.29	37.14	17.94	18.21	19.19
	STP-TrellisNets+	**19.40**	**20.09**	**21.53**	**30.72**	**32.02**	**34.03**	**14.55**	**15.36**	**16.04**

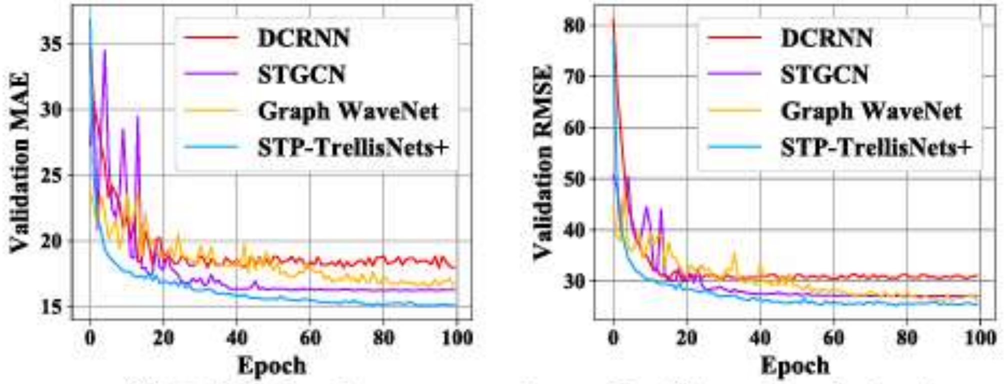
TABLE 5
Multi-Step Prediction Performance Under 30-Minute Time Interval Setting

Datasets	Models	MAE			RMSE			MAPE (%)		
		Step 1	Step 2	Step 3	Step 1	Step 2	Step 3	Step 1	Step 2	Step 3
Shenzhen	Last-Value	64.55	72.32	79.03	162.89	170.14	179.22	32.84	33.90	34.28
	HA	101.34	102.99	102.75	222.46	223.29	223.05	74.88	75.35	75.54
	VAR	60.43	66.34	72.24	150.02	158.43	163.04	26.33	27.96	28.61
	MLP	46.38	53.98	61.23	80.32	89.07	96.13	19.25	20.38	21.86
	LSTM	44.35	52.80	59.51	75.90	82.13	90.46	18.36	19.24	20.75
	ConvLSTM	40.92	47.20	56.42	66.38	74.56	81.81	15.78	16.31	17.64
	DCRNN	34.64	44.17	52.78	58.03	63.07	72.71	14.38	15.39	16.78
	STGCN	27.51	32.26	38.33	44.26	53.52	65.43	14.50	16.11	19.06
	Graph WaveNet	27.80	30.88	34.00	42.71	47.76	52.90	14.41	15.83	17.32
	STP-TrellisNets+	**26.64**	**29.03**	**31.23**	**40.44**	**42.68**	**48.96**	**11.03**	**11.80**	**12.01**
Hangzhou	Last-Value	84.61	90.34	97.30	176.33	181.18	188.64	24.56	25.07	26.45
	HA	85.60	86.34	85.99	197.10	197.25	197.07	21.69	21.51	21.15
	VAR	83.89	89.07	96.32	200.87	208.23	216.78	26.43	27.46	28.98
	MLP	68.37	78.30	88.48	120.20	127.28	138.35	24.12	25.69	26.05
	LSTM	64.26	71.08	77.09	110.30	115.62	121.47	22.75	23.90	24.93
	ConvLSTM	62.19	69.33	75.24	105.56	112.73	119.39	22.67	23.04	24.98
	DCRNN	58.28	67.90	74.51	67.32	78.93	84.13	17.13	18.86	20.09
	STGCN	36.94	42.19	50.52	56.92	67.55	81.94	9.90	11.30	17.10
	Graph WaveNet	36.18	39.65	42.40	54.83	61.21	66.33	9.82	10.77	11.83
	STP-TrellisNets+	**34.45**	**37.79**	**40.34**	**50.89**	**55.17**	**59.64**	**9.71**	**10.34**	**11.76**

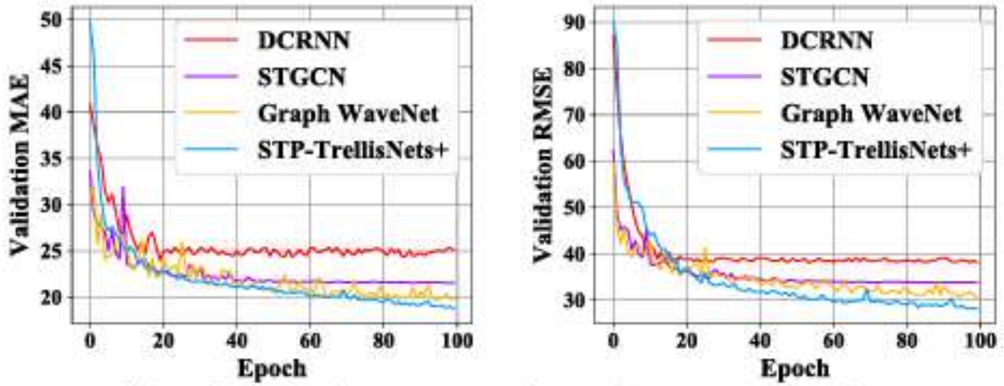
➢ 基线方法

Last-Value、HA、VAR、MLP、LSTM、ConvLSTM、DCRNN、STGCN、Graph WaveNet

实验-与其他时空方法的损失收敛比较



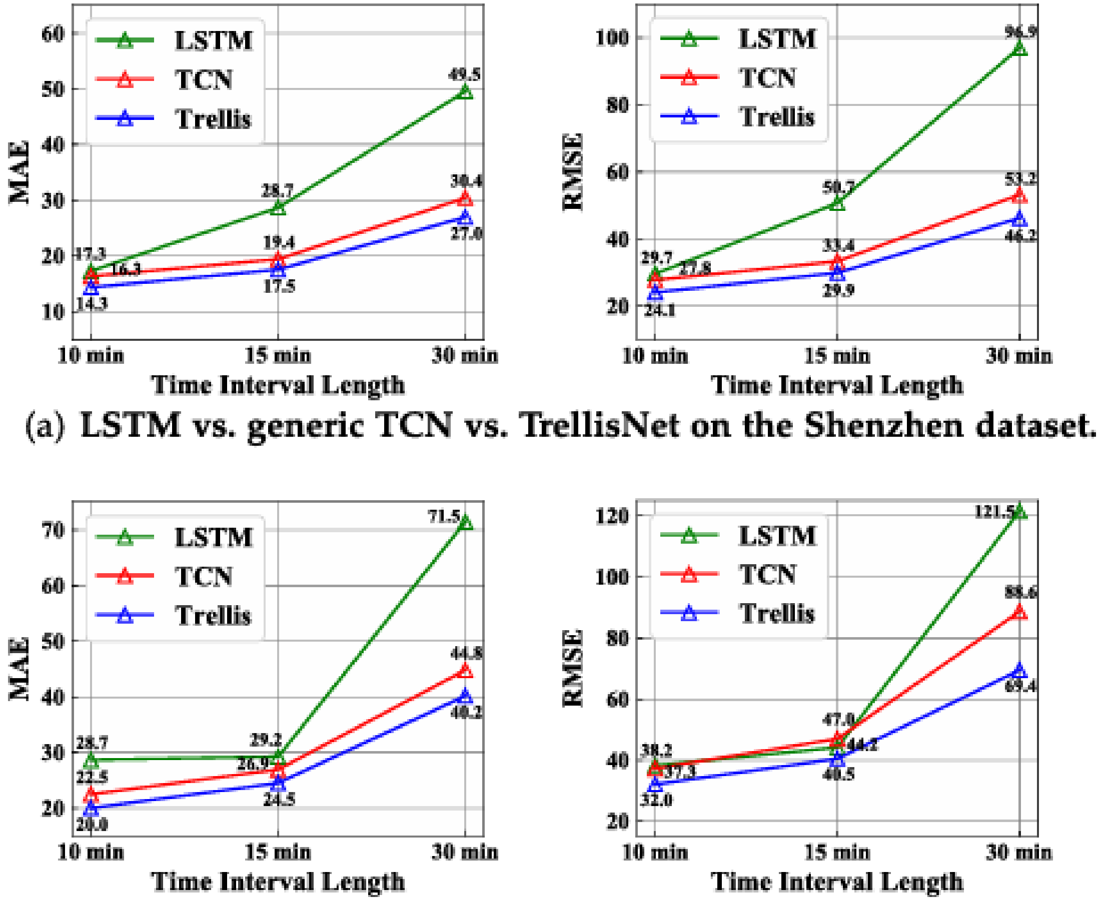
(a) Validation loss v.s epoch on the Shenzhen dataset.



(b) Validation loss v.s epoch on the Hangzhou dataset.

Fig. 10. The validation loss versus training epochs comparison among DCRNN, STGCN, Graph WaveNet and STP-TrellisNet+.

实验-处理时间相关性的模型比较

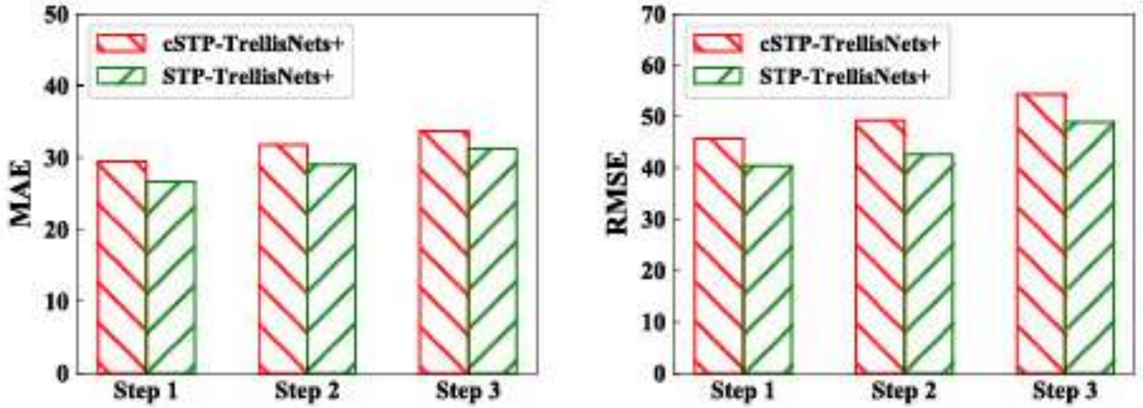


(a) LSTM vs. generic TCN vs. TrellisNet on the Shenzhen dataset.

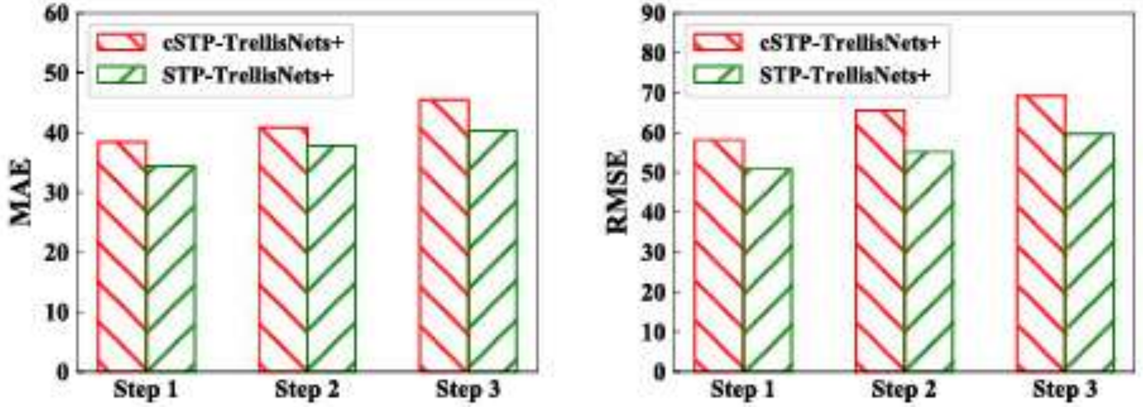
(b) LSTM vs. generic TCN vs. TrellisNet on the Hangzhou dataset.

Fig. 11. The MSP flow prediction performance comparison among LSTM, generic TCN, and C-TrellisNet.

实验-处理MSP流量不连续性的必要



(a) **STP-TrellisNets+ vs. cSTP-TrellisNets+ on the Shenzhen dataset.**



(b) **STP-TrellisNets+ vs. cSTP-TrellisNets+ on the Hangzhou dataset.**

Fig. 12. The multi-step prediction performance comparison between the STP-TrellisNets+ and cSTP-TrellisNets+.

实验-与模型变体的比较

TABLE 6
Comparison With Architectural Variants Under 10-Minute and 30-Minute Time Interval Settings

Models	10-minute	30-minute
	Step 1/2/3	Step 1/2/3
P-TrellisNetED	16.22/17.31/18.03	31.26/34.16/36.18
CP-TrellisNetsED	16.01/16.98/17.40	30.69/ 33.72/ 35.44
Single DGCN	15.08/15.90/16.12	28.34/ 31.59/ 32.92
STP-TrellisNets+	**14.60/14.86/14.95**	**26.64/29.03/31.23**

实验-多步预测策略分析

TABLE 7
Comparison With Variants of Multi-Step Prediction Strategies Under 10-Minute and 30-Minute Time Interval Settings

Models	10-minute	30-minute
	Step 1/2/3	Step 1/2/3
Step-by-Step	14.75/18.36/21.02	31.05/36.36/40.64
Plus FC Layer	15.20/16.22/17.07	51.92/38.18/53.61
Multi-TrellisNets	15.16/15.43/15.78	29.60/34.90/40.34
Multi-STP-TrellisNets	15.01/15.39/15.53	28.45/33.11/40.23
STP-TrellisNets+	**14.60/14.86/14.95**	**26.64/29.03/31.23**

实验- CP-TrellisNetsED输入长度与GC-TrellisNetsED中DGCN数量的联合影响

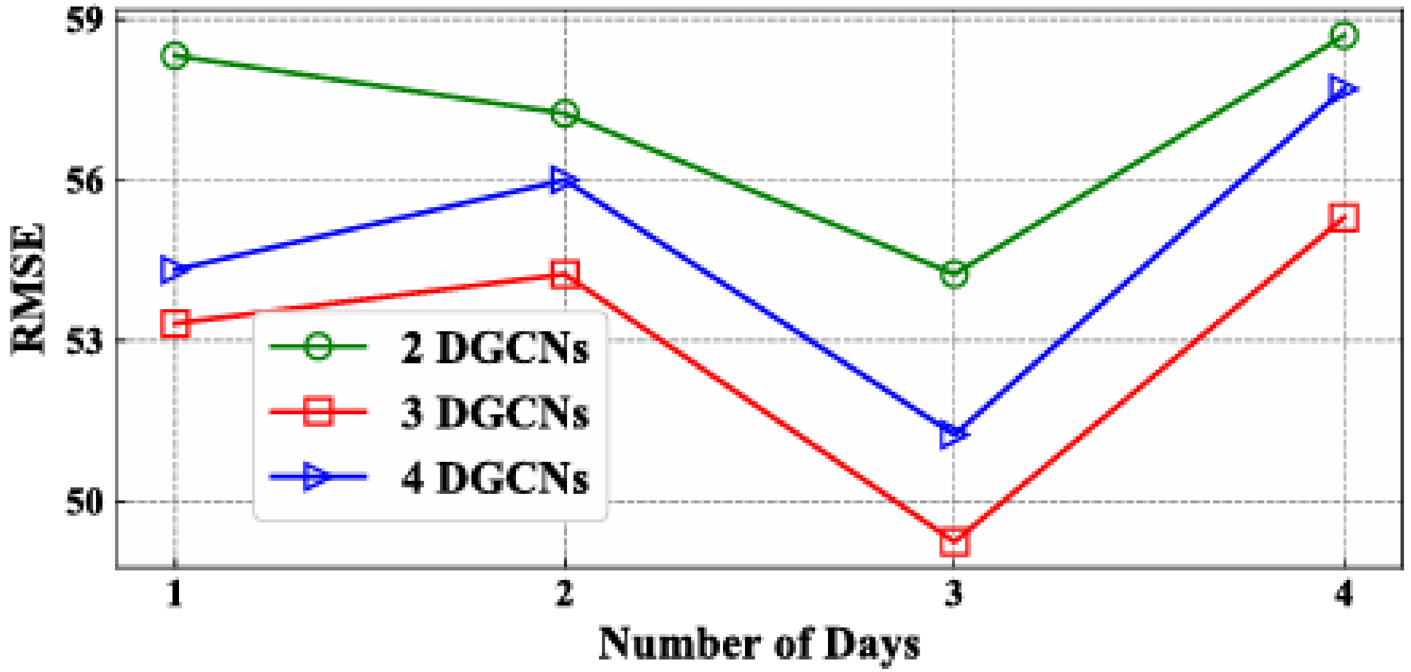


Fig. 13. Joint influence of CP-TrellisNetsED' input length and the number of DGCNs.

4 总结

总结

- **总结:** 本文提出了一种用于多步MSP流量预测的新框架——STP-TrellisNets+。 STP-TrellisNets+包含两个并行的时间模块（CP-TrellisNetsED）和空间模块(GC-TrellisNetsED)， CP-TrellisNetsED通过串联TrellisNet和一个基于TrellisNet的编码器编码器捕捉短期和长期时间相关性， GC-TrellisNetsED将多个DGCN的输出输入一个基于TrellisNet的编码器编码器捕捉空间相关性的动态变化。实验证明了STP-TrellisNets+的有效性。

谢谢大家!

汇报人: 韦浩文

时间: 2025-8-10