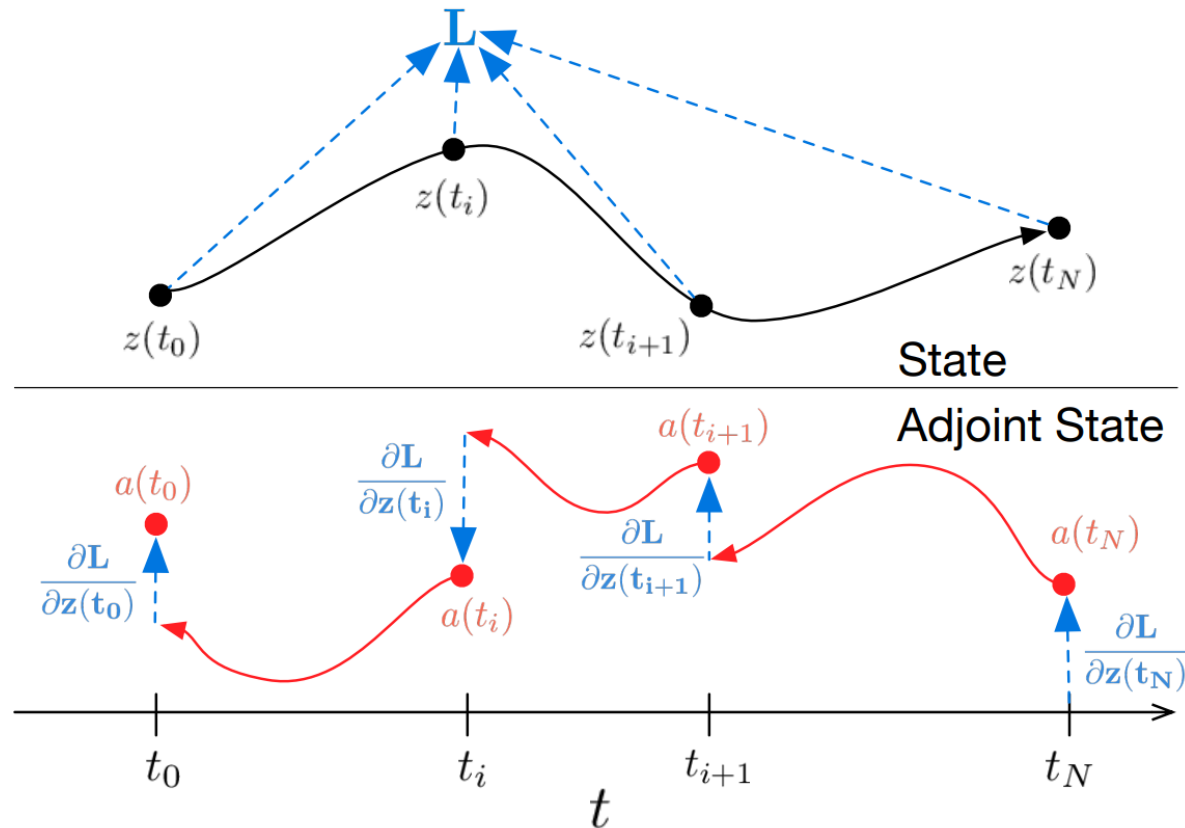
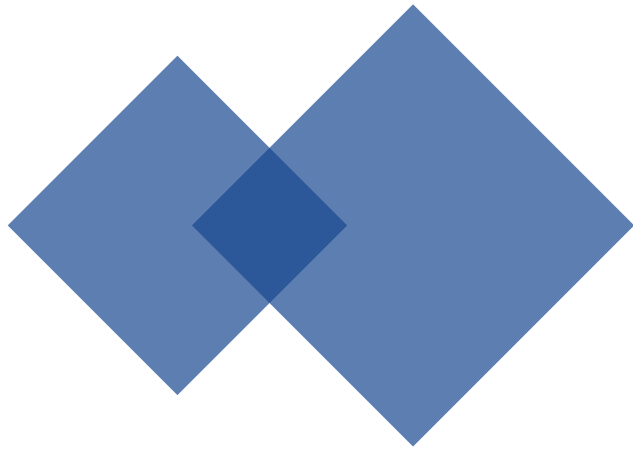


STGODE

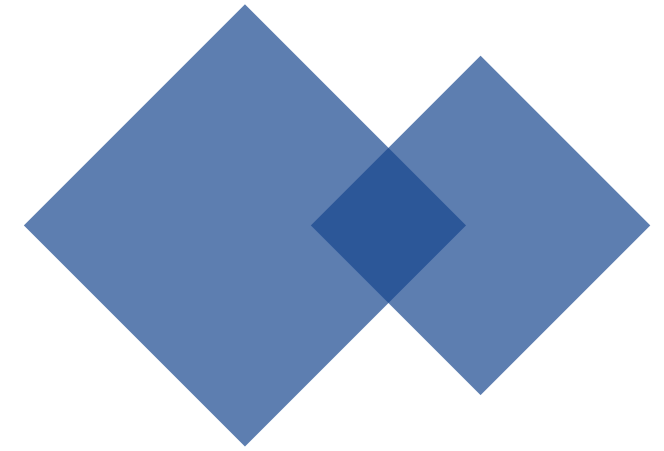
Spatial-Temporal **Graph ODE**
Networks for Traffic Flow Forecasting





STGODE

Spatial-**T**emporal **G**raph **ODE**
Networks for Traffic Flow
Forecasting



24.4.16

Presented by Yyyq



常微分方程 → 带初值的常微分方程 → 常微分方程的解 → 残差网络和神经常微分方程

常微分方程

$$\frac{dx}{dt} + x = 1 \Rightarrow \frac{dx}{dt} = 1 - x$$

$$\Rightarrow f(x) = x - \frac{1}{2}x^2 + C$$

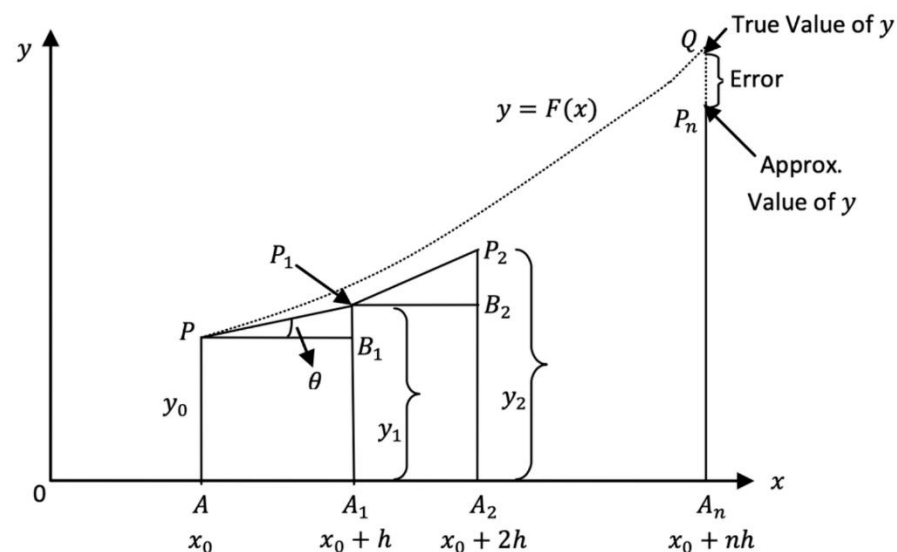
$$F(x, y(x), y'(x), \dots, y^{(n)}(x)) = 0$$

初始条件：初值 $(x_0, f(x_0)) = (0, 1) \Rightarrow C = 1$

ODE Solver

$$\frac{dy}{dx} = f(x, y), \quad y_0 = y(x_0)$$

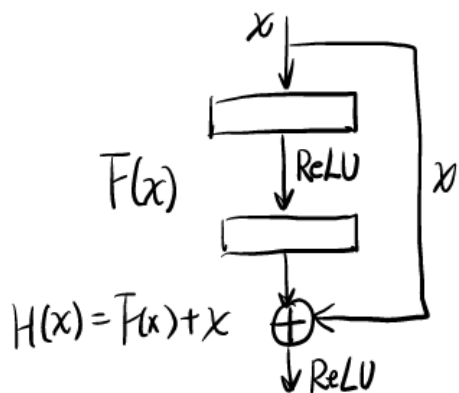
经典的数值求解方法：微分思想，直线代替曲线





常微分方程 \rightarrow 带初值的常微分方程 \rightarrow 常微分方程的解 \rightarrow 残差网络和神经常微分方程

Neural ODE: (ResNet)



$$H(x) = F(x) + x$$

$$h_{t+1} = h_t + f(h_t, \theta)$$

$$\Rightarrow \frac{h_{t+1} - h_t}{\Delta t} = f(h_t, \theta), \Delta t = 1$$

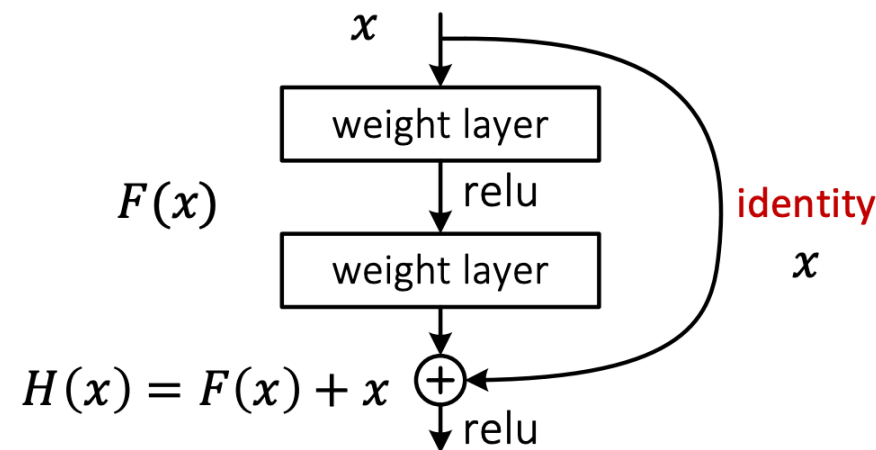
$$\Rightarrow \text{对 } \Delta t \text{ 取极限: } \lim_{\Delta t \rightarrow 0} \frac{h_{t+1} - h_t}{\Delta t} = f(h_t, \theta)$$

(即对 h 求导)

$$\Rightarrow \frac{dh(t)}{dt} = f(h(t), t, \theta)$$

$$\Rightarrow \text{用 NN 代替 } f, \text{ 即 } \frac{dh(t)}{dt} = \text{NN}(h(t), t, \theta)$$

• Residual net





常微分方程 → 带初值的常微分方程 → 常微分方程的解 → 残差网络和神经常微分方程

$$\frac{dh(t)}{dt} = NN(h(t), t, \theta) \quad \begin{cases} t=0: h(0) \text{ 即网络的输入} \\ t=T: h(T) \text{ 即网络的输出.} \end{cases}$$

前向计算: 输入初值 $h(0)$.

$$L(x(t)) = L(x(0)) + \int_0^t f(x(\tau), \tau) d\tau$$

反向传播: (利用 loss 计算梯度, 更新 θ)

GPT:

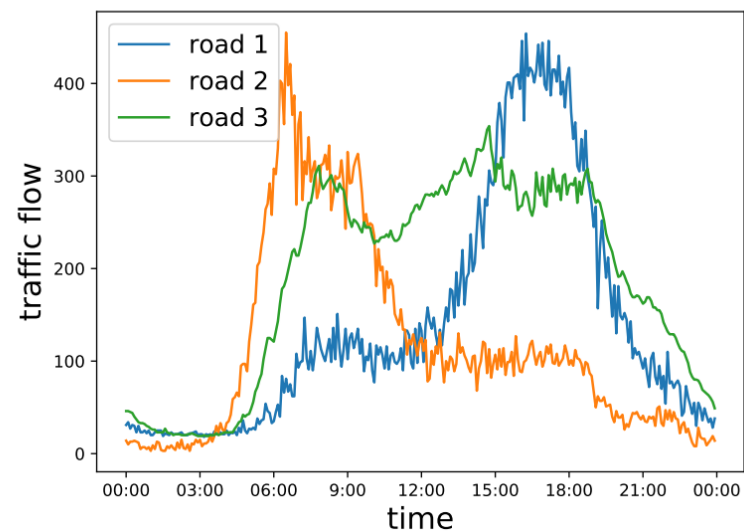
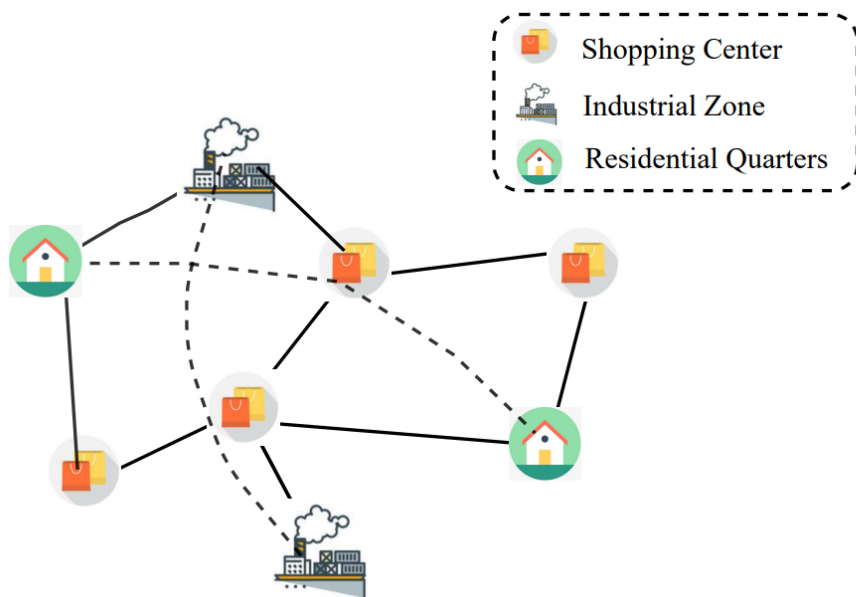
- 在神经ODE中, 模型的结构是动态的, 其行为随着时间的推移而改变。
- 神经ODE的核心思想是使用ODE来建模数据的动态变化过程, 而不是直接建模输入和输出之间的映射关系。
- 具体地, 给定一个输入数据, 神经ODE通过ODE的解来推断出数据的演化路径。
- 这使得神经ODE具有动态适应数据的能力, 能够更好地适应数据的复杂性和变化性。
- 灵活性强、模型参数少、能够处理动态数据等优点

01

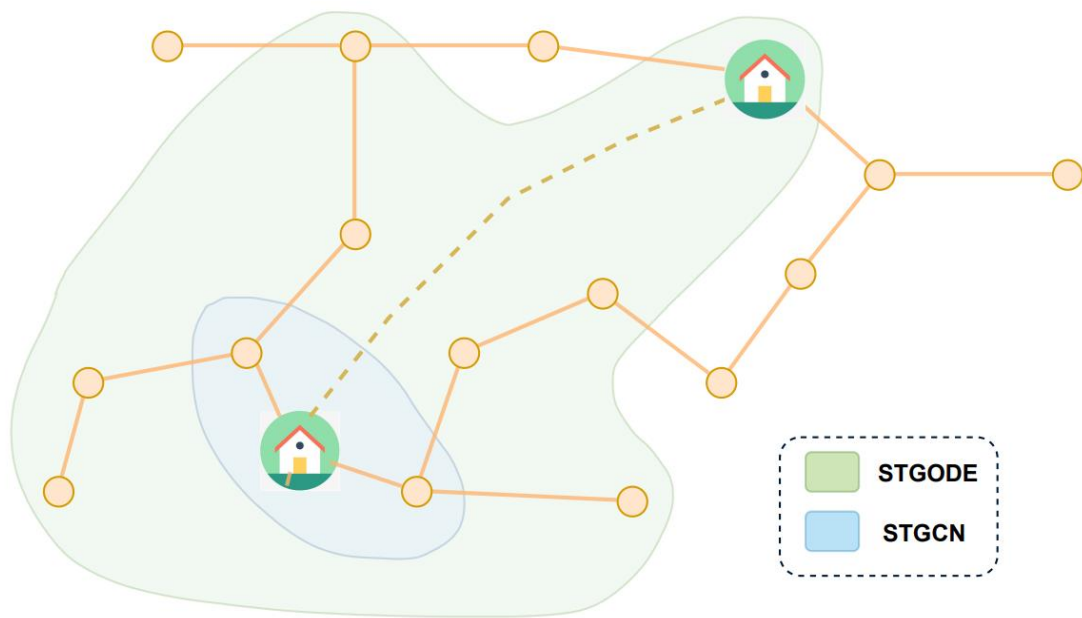


问题描述

- 近距离和远距离的空间依赖同时存在
- 多样化的时间模式



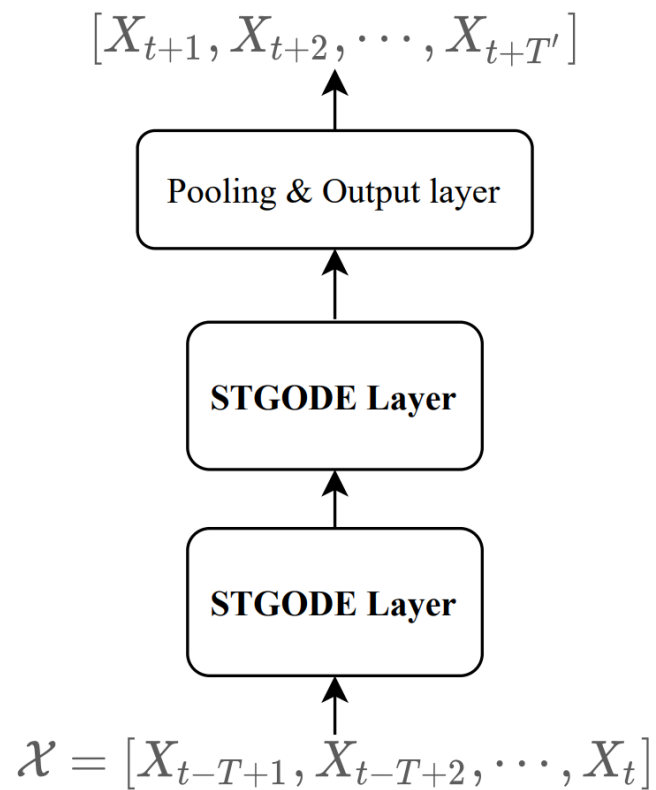
- 引出连续图神经网络 (CGNN)
 - 具有残差连接的连续图神经网络(Continuous GNNs)
 - 证明了具有残差连接的离散层可以看作是常微分方程 (ODE) 的离散化
 - 更深的网络，提高提取远距离时空相关性的能力
- 引入两类邻接矩阵：空间邻接矩阵和语义邻接矩阵



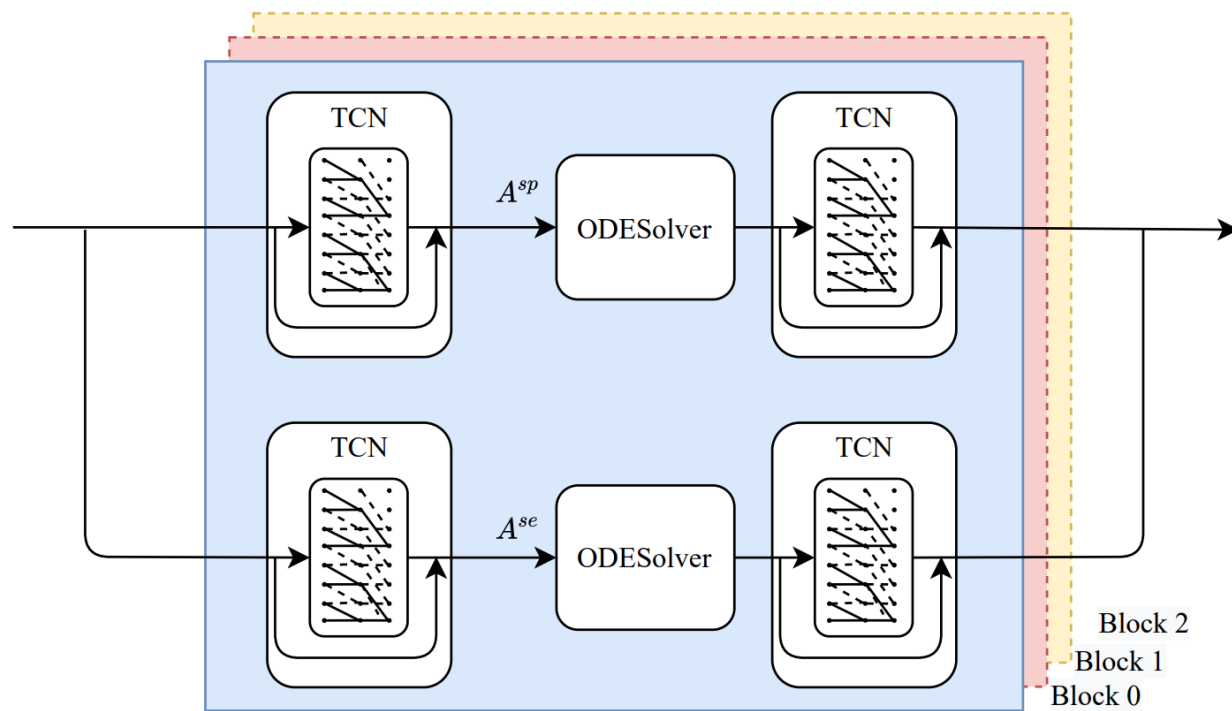
03



模型结构: Overview



(a) Framework



(b) STCODE Layer



- 空间邻接矩阵：基于距离

$$A_{ij}^{sp} = \begin{cases} \exp\left(-\frac{d_{ij}^2}{\sigma^2}\right) & , \text{ if } \exp\left(-\frac{d_{ij}^2}{\sigma^2}\right) \geq \epsilon \\ 0 & , \text{ otherwise} \end{cases}$$

- 语义邻接矩阵：DTW算法

$$D(i, j) = \text{dist}(x_i, y_j) + \min(D(i-1, j), D(i, j-1), D(i-1, j-1)),$$

$$A_{ij}^{SE} = \begin{cases} 1, & \text{DTW}(X^i, X^j) < \epsilon \\ 0, & \text{otherwise} \end{cases}$$

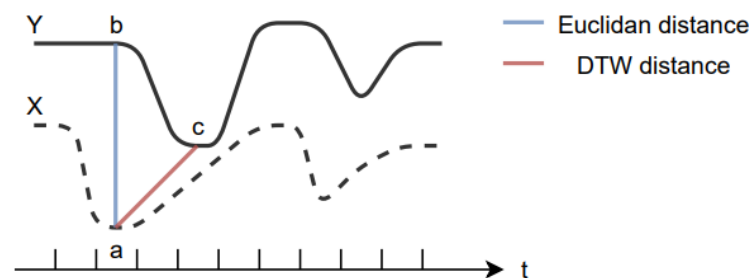


Figure 4: An example of the difference between the Euclidean distance and the DTW distance.

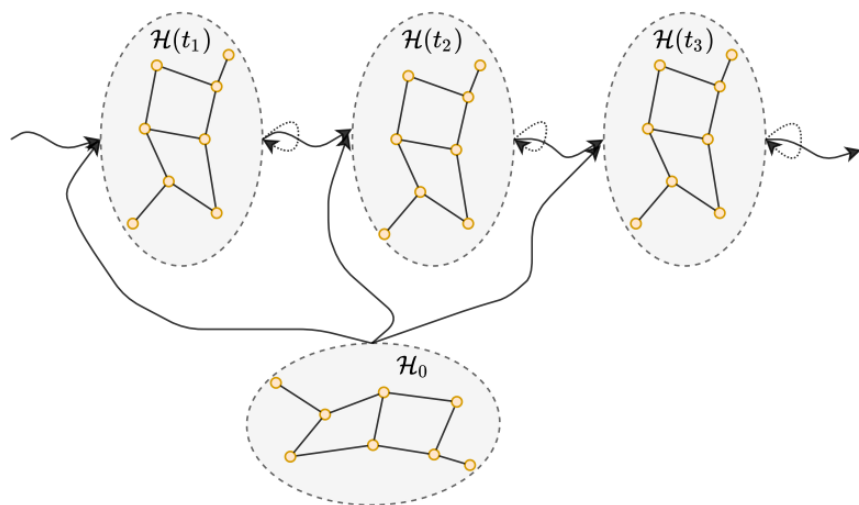


➤ 用ODE改进常规GCN会遇到的过平滑问题

$$H_{l+1} = \text{GCN}(H_l) = \sigma(\hat{A}H_lW),$$



$$\mathcal{H}(t) = \text{ODESolve}\left(\frac{d\mathcal{H}(t)}{dt}, \mathcal{H}_0, t\right),$$



证明了使用的初始连接 H_0 的有效性

证明了具有残差连接的离散层可以看作是常微分方程 (ODE) 的离散化

证明了式子有解析解

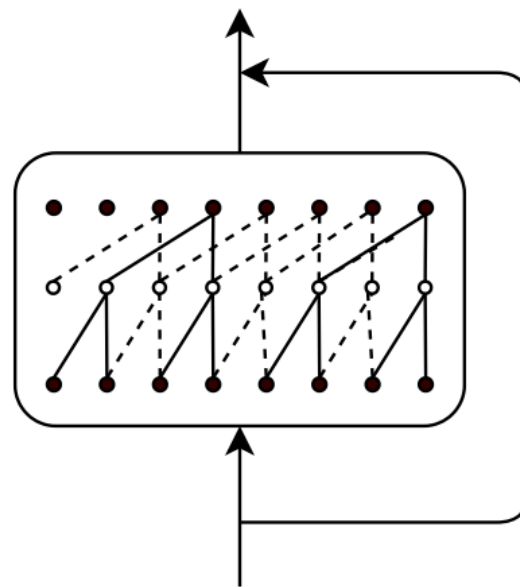
04



算法实现：时间卷积块TCN

- 沿时间轴的一维扩张时间卷积网络
- 扩大感受野: $d^l \rightarrow 2^{l-1}$

$$H_{tcn}^l = \begin{cases} X & , l = 0 \\ \sigma(W^l *_{d^l} H_{tcn}^{l-1}) & , l = 1, 2, \dots, L \end{cases}$$

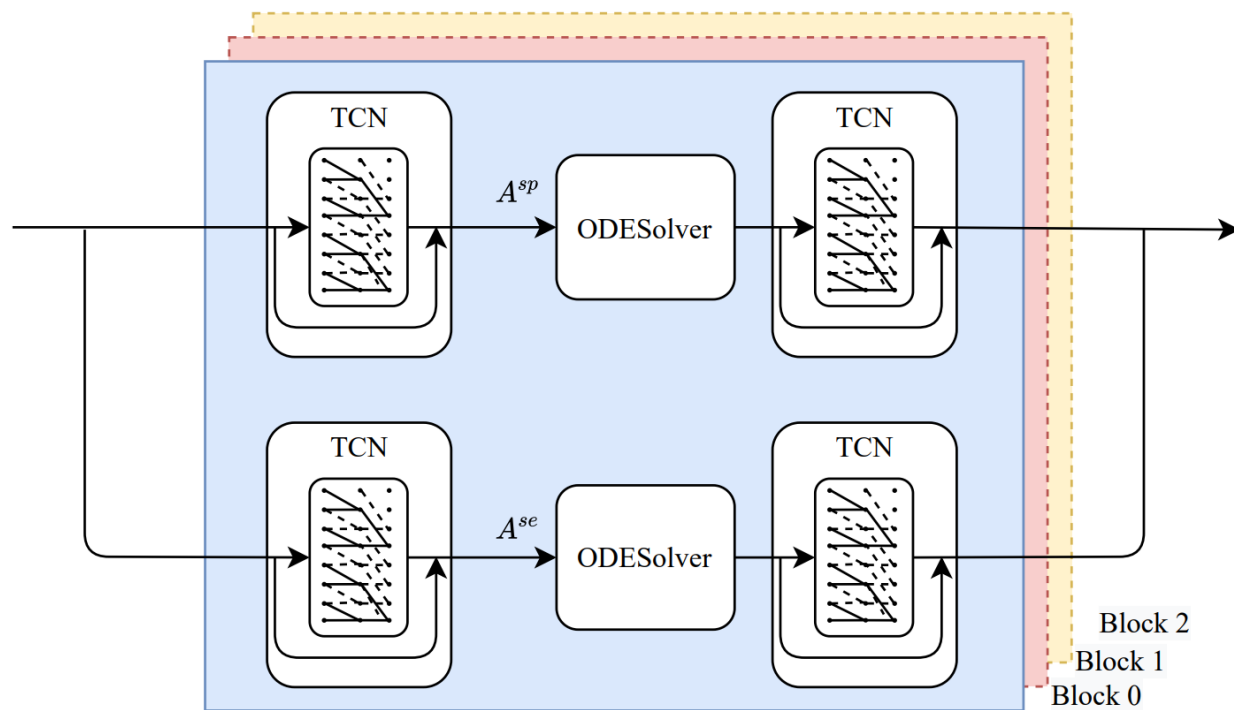


04



算法实现: STGODE Layer

- 三明治结构: 两个TCN夹着STGODE块
- 两种空间关系并行: 空间邻接矩阵和语义邻接矩阵
- 最大池化层: 选择性地聚合来自不同块的信息





Datasets	#Sensors	#Edges	Time Steps
PeMSD7(M)	228	1132	12672
PeMSD7(L)	1026	10150	12672
PeMS03	358	547	26208
PeMS04	307	340	16992
PeMS07	883	866	28224
PeMS08	170	295	17856

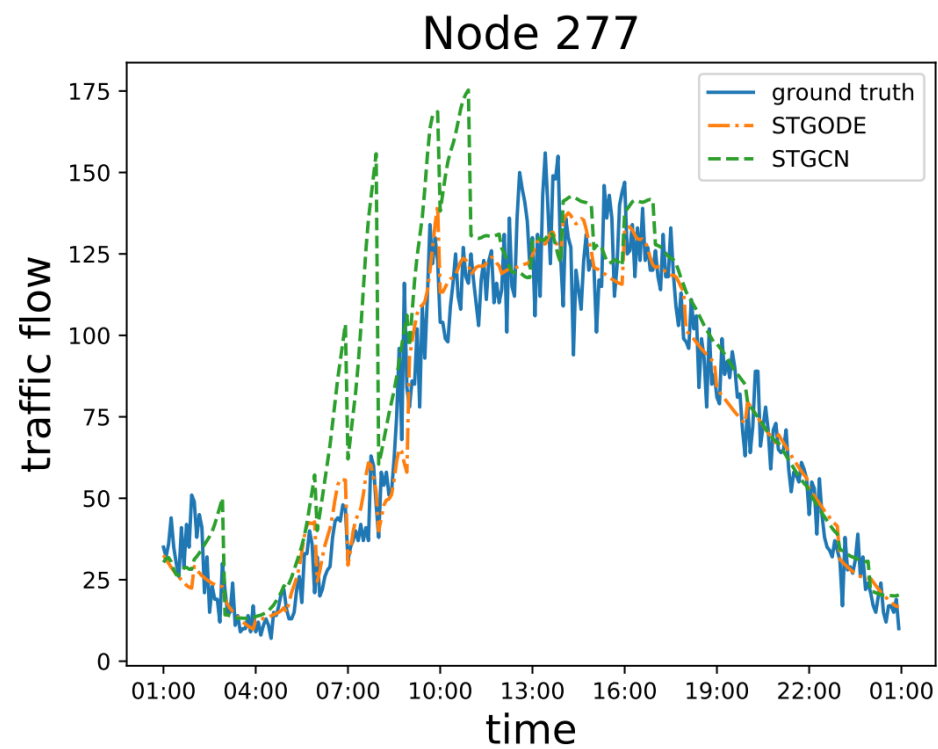
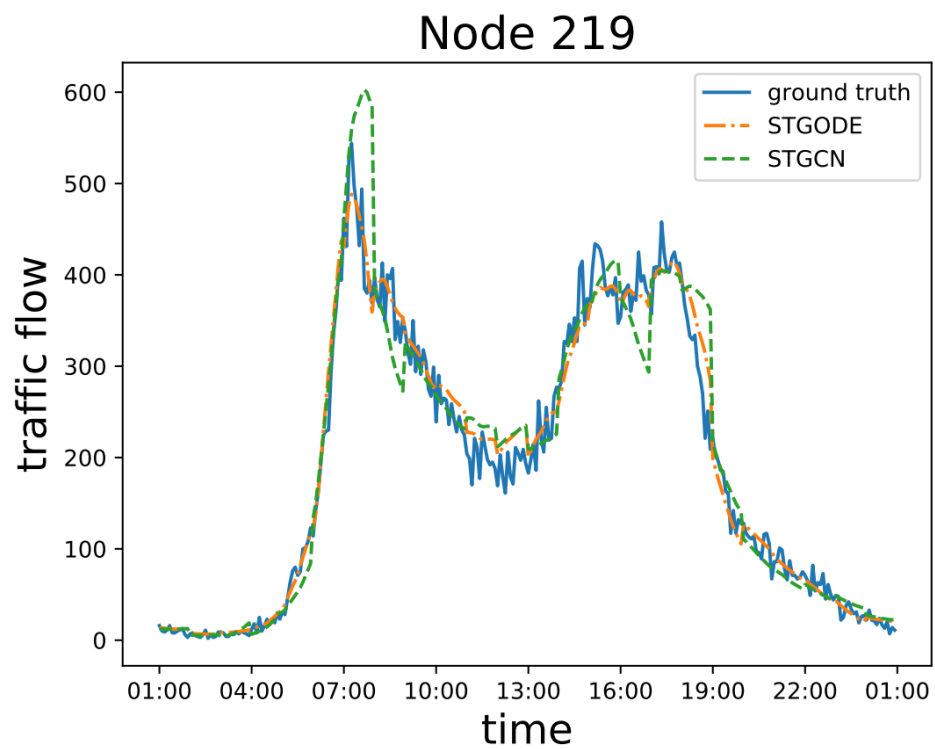
Table 1: Datasets description

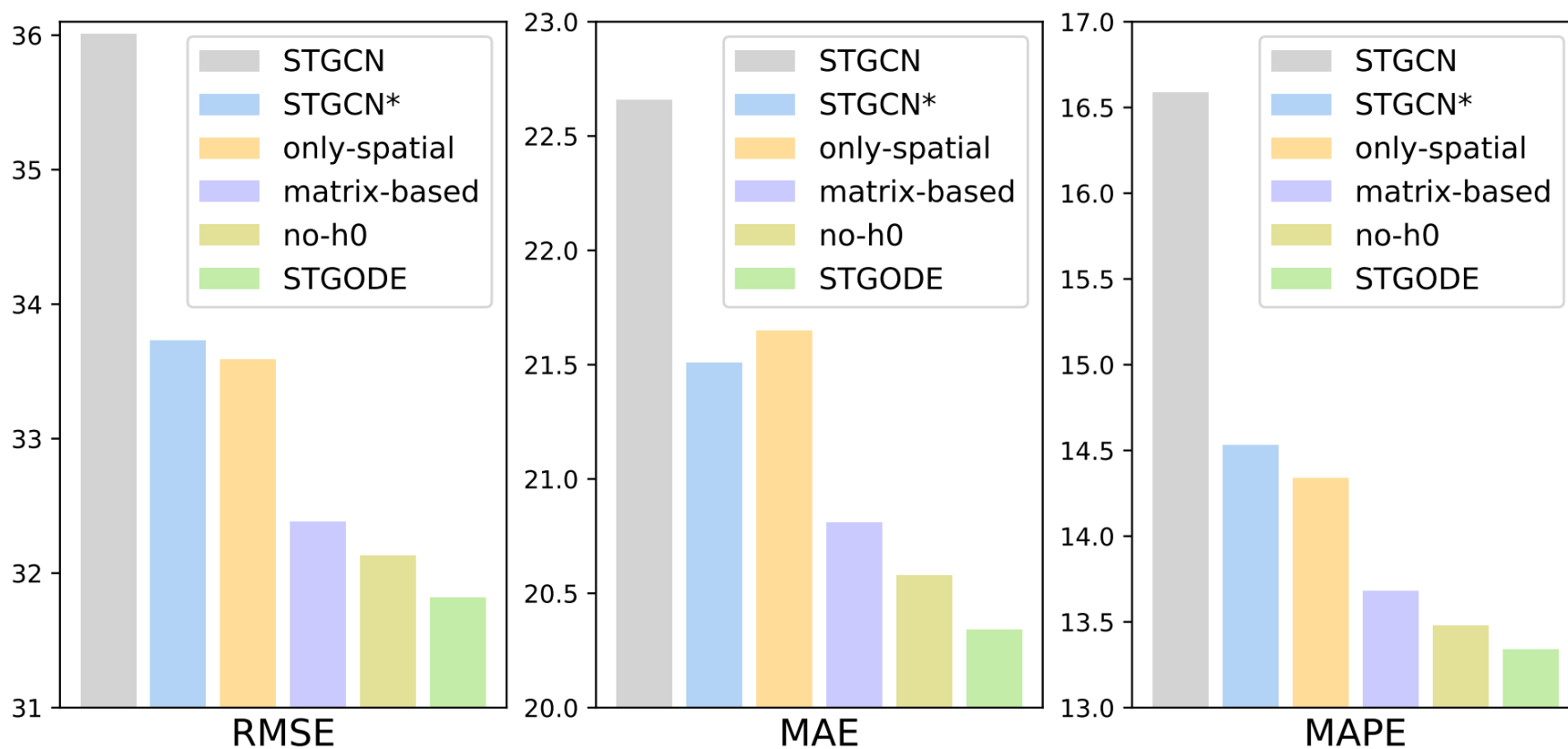
主要基线：GraphWaveNet、ASTGCN、STSGCN

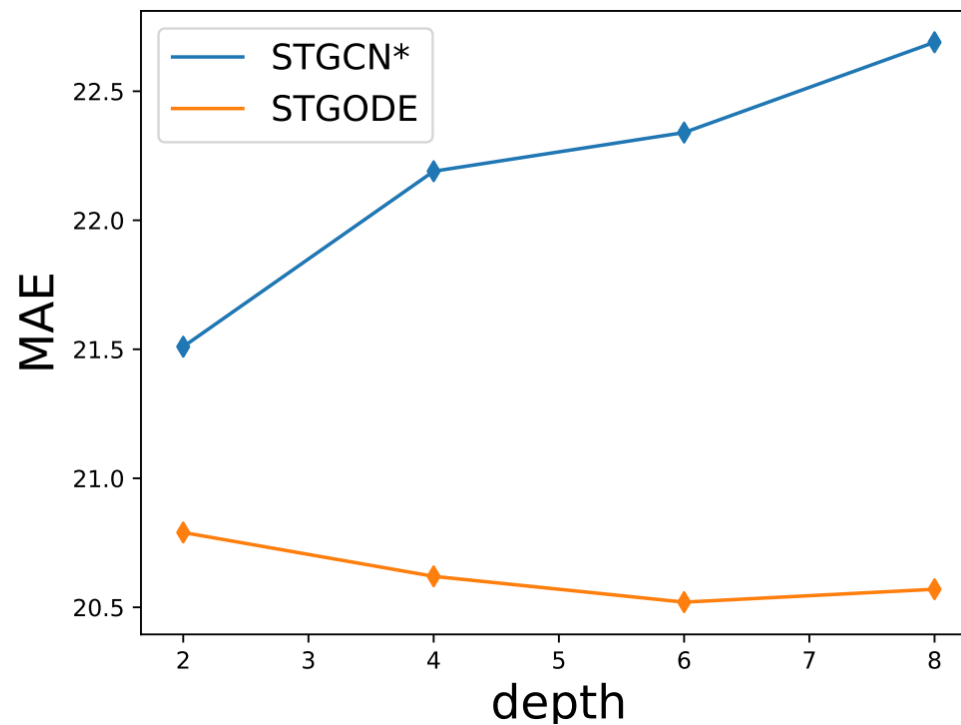
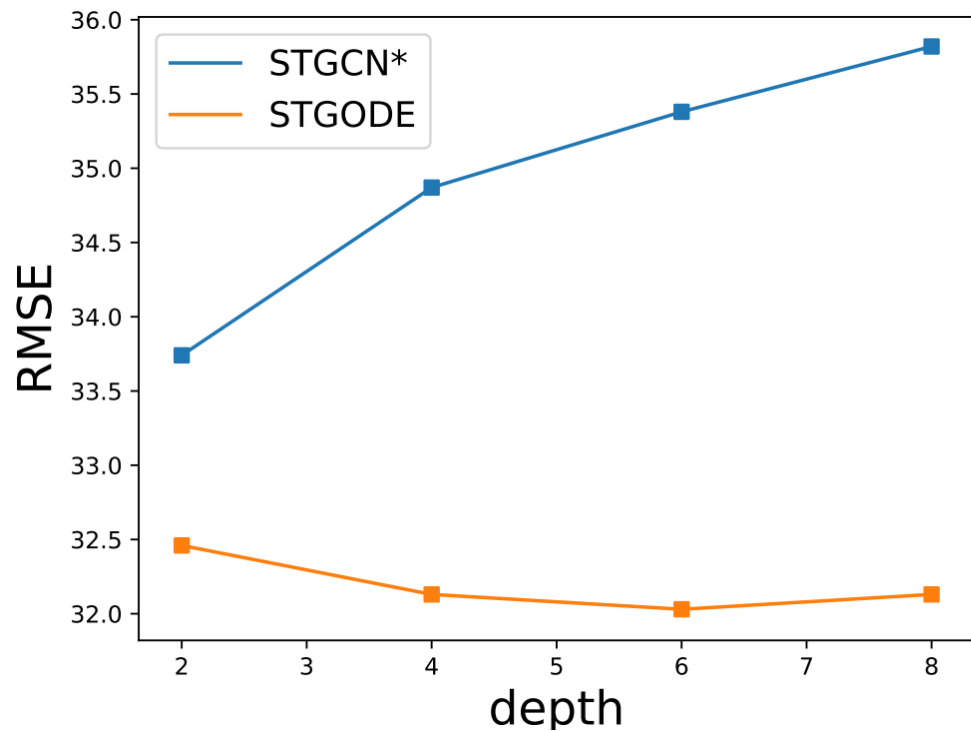


Dataset	Metric	ARIMA	STGCN	DCRNN	ASTGCN(r)	GraphWaveNet	STSGCN	STODE
PeMSD7(M)	RMSE	13.20	7.55	7.18	6.87	6.24	5.93	5.66
	MAE	7.27	4.01	3.83	3.61	3.19	3.01	2.97
	MAPE	10.38	9.67	9.81	8.84	8.02	7.55	7.36
PeMSD7(L)	RMSE	12.39	8.28	8.33	7.64	7.09	6.88	5.98
	MAE	7.51	4.84	4.33	4.09	3.75	3.61	3.22
	MAPE	15.83	11.76	11.41	10.25	9.41	9.13	7.94
PeMS03	RMSE	47.59	30.42	30.31	29.56	32.77	29.21	27.84
	MAE	35.41	17.55	17.99	17.34	19.12	17.48	16.50
	MAPE	33.78	17.43	18.34	17.21	18.89	16.78	16.69
PeMS04	RMSE	48.80	36.01	37.65	35.22	39.66	33.65	32.82
	MAE	33.73	22.66	24.63	22.94	24.89	21.19	20.84
	MAPE	24.18	14.34	17.01	16.43	17.29	13.90	13.77
PeMS07	RMSE	59.27	39.34	38.61	37.87	41.50	39.03	37.54
	MAE	38.17	25.33	25.22	24.01	26.39	24.26	22.99
	MAPE	19.46	11.21	11.82	10.73	11.97	10.21	10.14
PeMS08	RMSE	44.32	27.88	27.83	26.22	30.04	26.80	25.97
	MAE	31.09	18.11	17.46	16.64	18.28	17.13	16.81
	MAPE	22.73	11.34	11.39	10.6	12.15	10.96	10.62

Table 2: Performance comparison of baseline models and STGODE on PeMS datasets.









谢谢观看

MANY THANKS !

24.4.16

