



研究报告

图神经网络在交通流量
预测中的应用

23.1.10





交通流的预测任务可以看做是一个“seq2seq”的任务，
给定图上前 t_1 个的交通数据来预测接下来 t_2 个交通数据。

模型分类	具体算法	缺点
传统时间序列分析 (统计模型)	历史平均(HA)、 自回归综合移动平均(ARIMA) 向量自回归(VAR)	这些方法需要数据满足一定的假设。 难以处理不稳定、非线性数据
传统机器学习	KNN、SVM	虽然能够处理复杂数据，但是 (1) 难以同步考虑高维交通数据的时空相关性 (2) 依赖于特征工程，需要较多专家经验
传统深度学习	CNN, LSTM	虽然可以提取时空特征，但是只能处理结构化的 2D或者3D等规则数据
基于图的深度学习	GCN, ChebNet, GGCN	没有同时考虑动态时空相关性

交通流的（各个时间点间的）时间动态性和（交通网络图各个节点间的）空间依赖性。



- 图定义: $G = (V, E, A)$, 其中 V 为节点集, E 为边集, A 为邻接矩阵。
- 对于单个时间步 t , G 的节点特征矩阵 $X_t \in R^{N \times d}$, 其中 N 为节点数, d 为特征数。
- 基于图的交通预测定义: 找到一个函数 f , 生成 $y = f(X; G)$
 - 其中 y 为待预测的交通状态,
 - $\chi = \{\chi^1, \chi^2, \dots, \chi^T\}$ 为图 G 上定义的历史流量状态,
 - T 为历史窗口大小中的时间步长数。
- 考虑外部因素 ε : 例如天气和节假日, $y = f(X, \varepsilon; G)$



DC DCNN结合正反向的流量卷积获取空间依赖特征

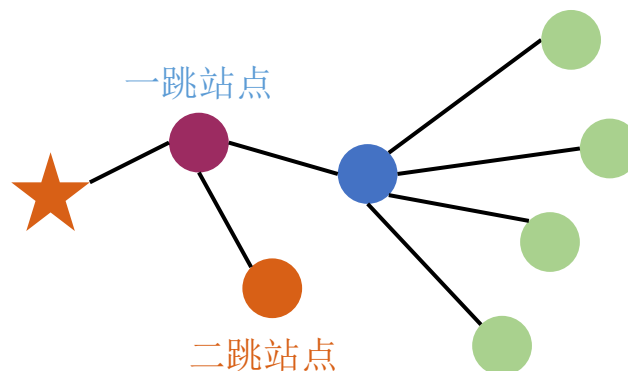
RNN DCGRU获取时间依赖特征

➤ 扩散 (随机游走)

$$\mathcal{P} = \sum_{k=0}^{\infty} \alpha(1 - \alpha)^k (D_O^{-1}W)^k$$

➤ 扩散卷积

$$X_{:,p} \star_{\mathcal{G}} f_{\theta} = \sum_{k=0}^{K-1} \left(\theta_{k,1} (D_O^{-1}W)^k + \theta_{k,2} (D_I^{-1}W^{\top})^k \right) X_{:,p} \quad \text{for } p \in \{1, \dots, P\}$$



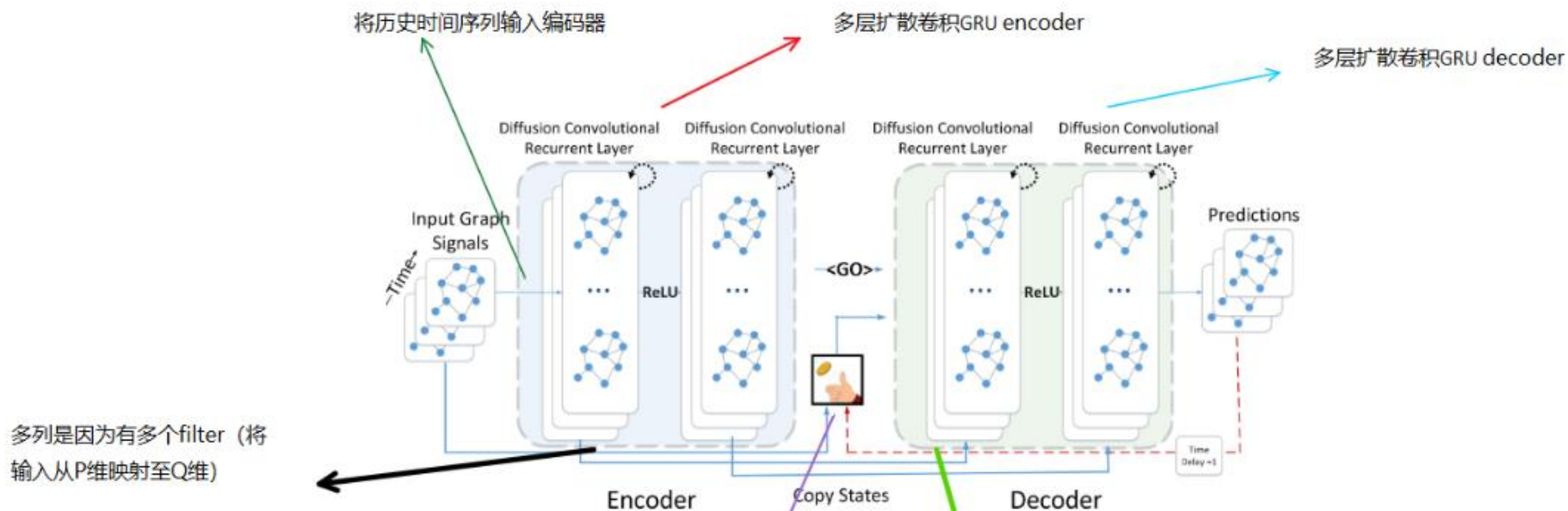


Figure 2: System architecture for the *Diffusion Convolutional Recurrent Neural Network* designed for spatiotemporal traffic forecasting. The historical time series are fed into an encoder whose final states are used to initialize the decoder. The decoder makes predictions based on either previous ground truth or the model output.

- ☐ 训练的时候, 一定概率给decoder的是真实输入值, 一定概率给decoder的是上一个时刻的输出结果 (也即这一时刻的预测值)
- ☐ 测试的时候, 给的就是上一时刻的输出结果

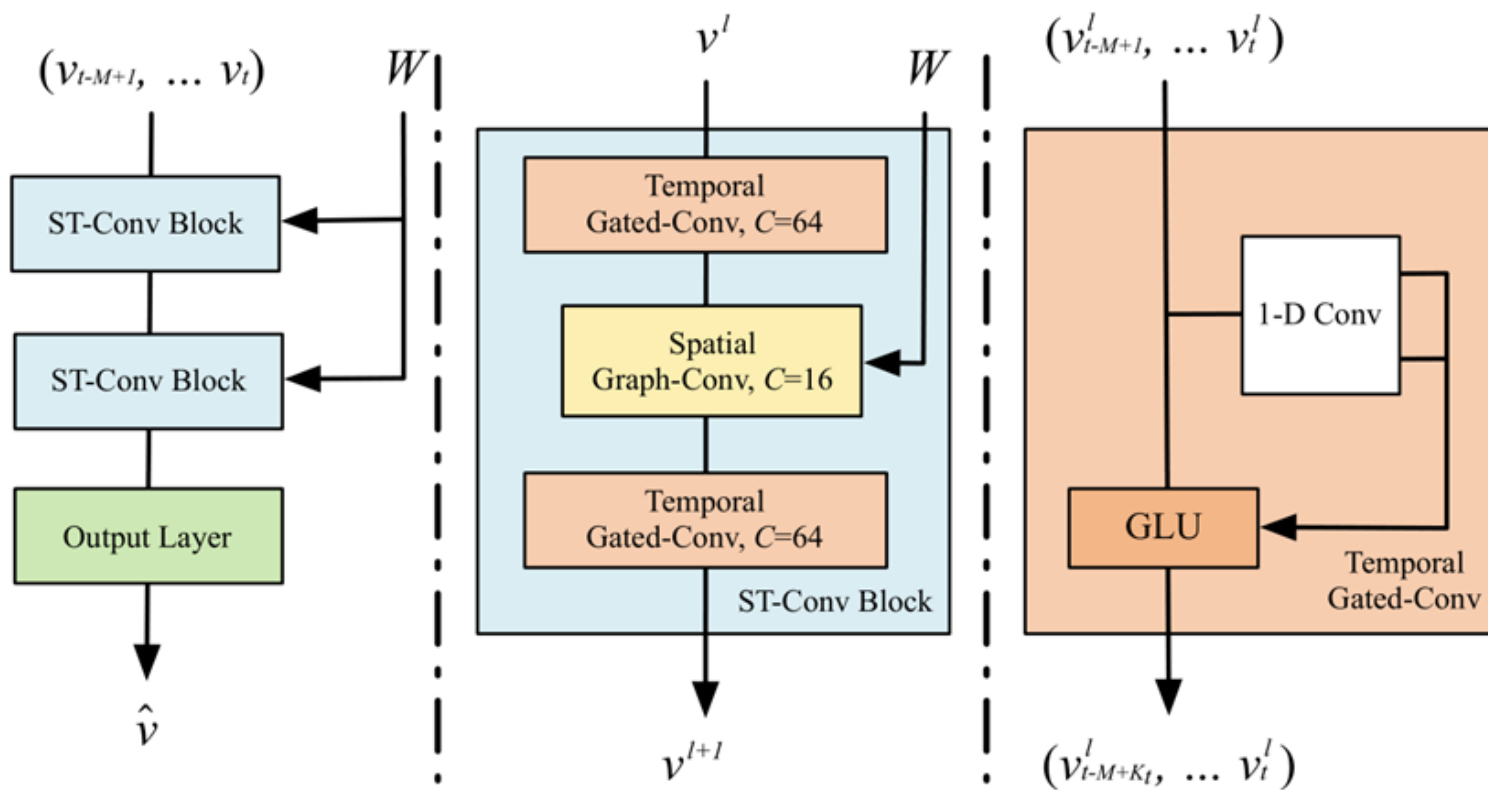
decoder使用相应的encoder的 H_{t-1}

03



经典模型: STGCN (IGCAI 2018)

在交通研究中首次应用纯卷积结构同时提取图结构时间序列的时空特征。

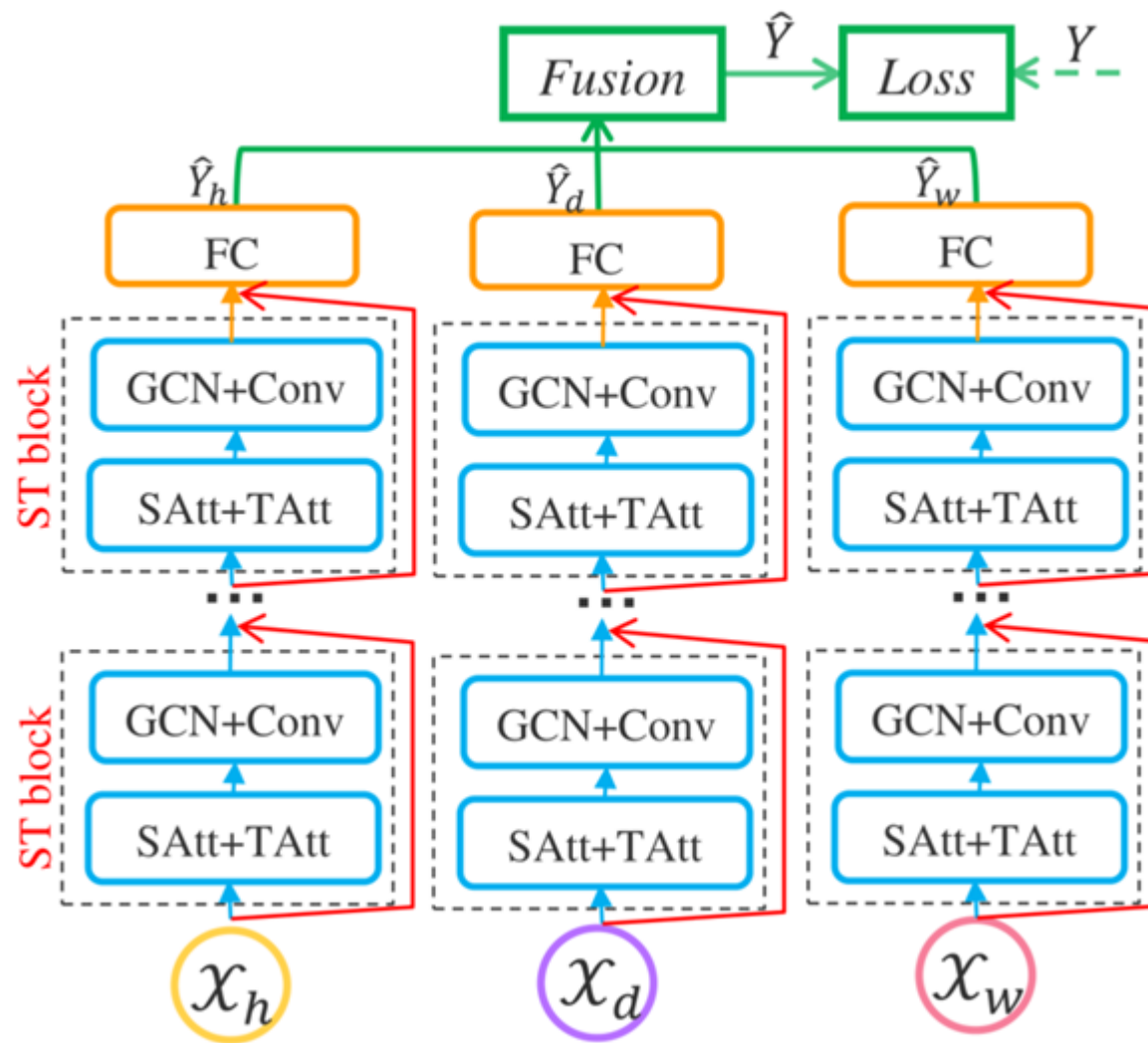


03

经典模型: ASTGCN (AAAI 2019)

图卷积 + 注意力机制

- 由三个独立的组件组成，
对应时间序列上三个不同长短的周期模式
(当前、日、周)
- $ST\ BLOCK = (Satt + Tatt) + (GCN + conv)$

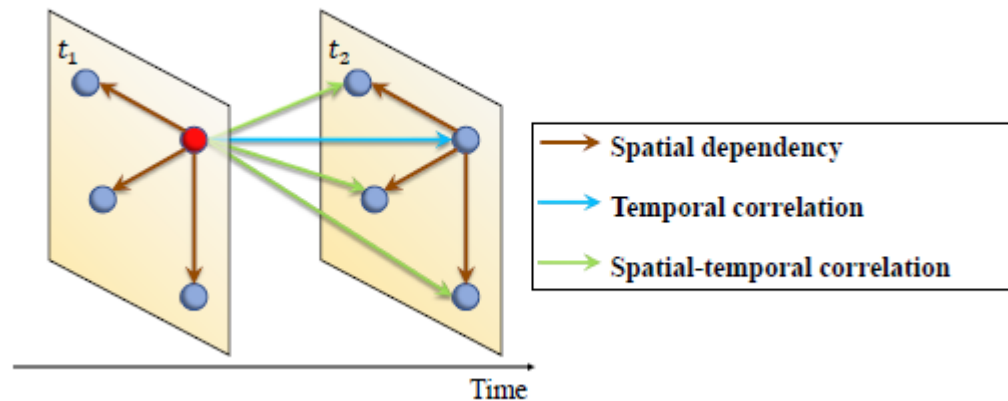


03



引申模型: STSGCN (AAAI 2020)

Synchronous: 同步的
Heterogeneity: 异质性



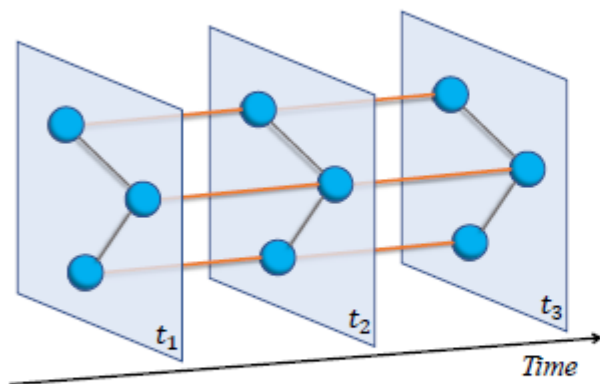
两个单独的组件，捕获空间依赖性和时间相关性，不能很好的体现时空相关性。



Synchronous: 同步的
Heterogeneity: 异质性

局部时空图

局部时空图



(a) Localized Spatial-Temporal Graph

局部时空图的邻接矩阵

$A^{(t_1)}$	$A^{t_1 \rightarrow t_2}$	
$A^{t_2 \rightarrow t_1}$	$A^{(t_2)}$	$A^{t_2 \rightarrow t_3}$
	$A^{t_3 \rightarrow t_2}$	$A^{(t_3)}$

在时间片i 和时间片j 节点与自身的连接关系

$$A'_{i,j} = \begin{cases} 1, & \text{if } v_i \text{ connects to } v_j \\ 0, & \text{otherwise} \end{cases}$$

在时间片i 的空间图的邻接矩阵

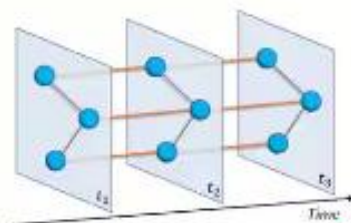
(b) Adjacency matrix of Localized Spatial-Temporal Graph

03

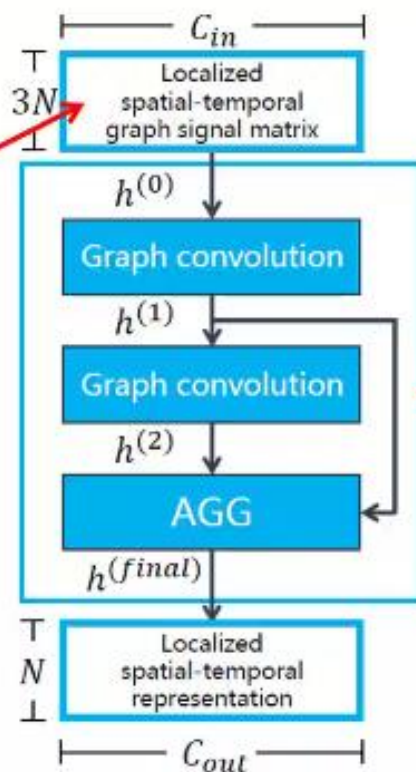


引申模型: STSGCN (AAAI 2020)

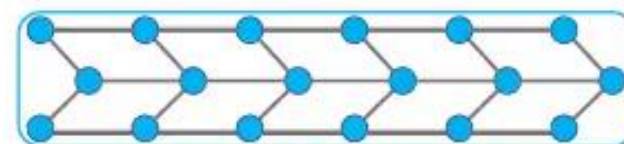
Synchronous: 同步的
Heterogeneity: 异质性



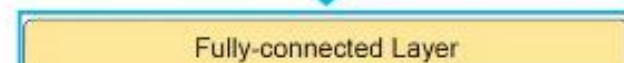
$A(t_1)$	$A^{t_1 \rightarrow t_2}$	
$A^{t_2 \rightarrow t_1}$	$A(t_2)$	$A^{t_2 \rightarrow t_3}$
	$A^{t_3 \rightarrow t_2}$	$A(t_3)$



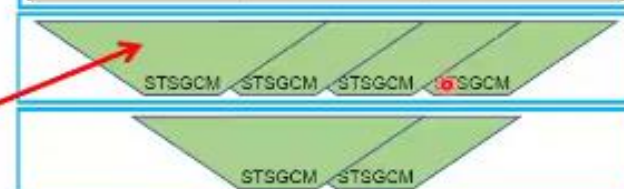
时空图卷积模块
(STSGCM)



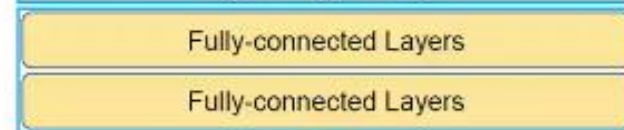
Input spatial-temporal network series



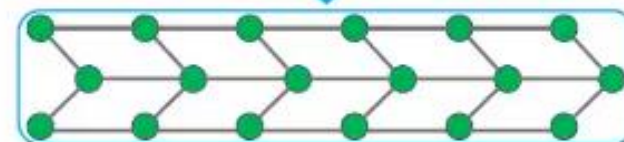
Input transform layer



时空图卷积层 (STSGCL)



Output transform layers



Target spatial-temporal network series

03

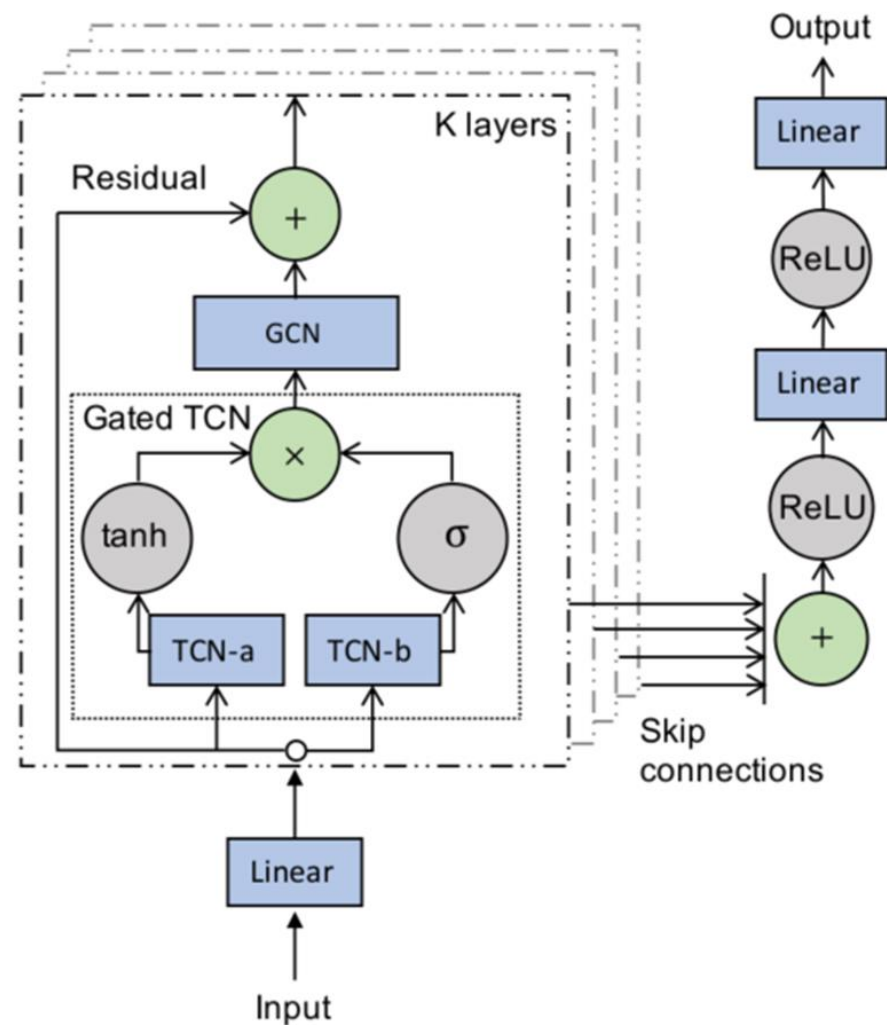


经典模型: Graph WaveNet (IJCAI 2019)

空间依赖: GCN + 自适应邻接矩阵 + 扩散卷积

时间依赖: 扩张因果卷积 (dilated casual convolution) + 门控机制

$$h = g(\Theta_1 \star \chi + b) \odot \sigma(\Theta_2 \star \chi + c),$$

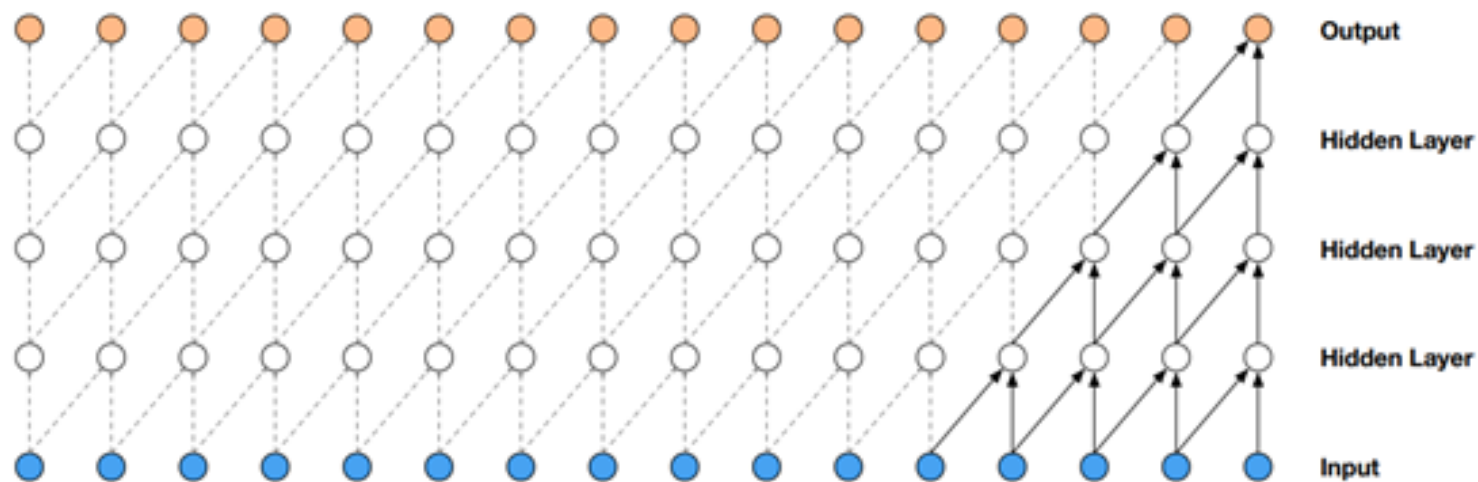


03



经典模型：Graph WaveNet (IJCAI 2019)

扩张因果卷积：替换传统叠加CNN+池化层



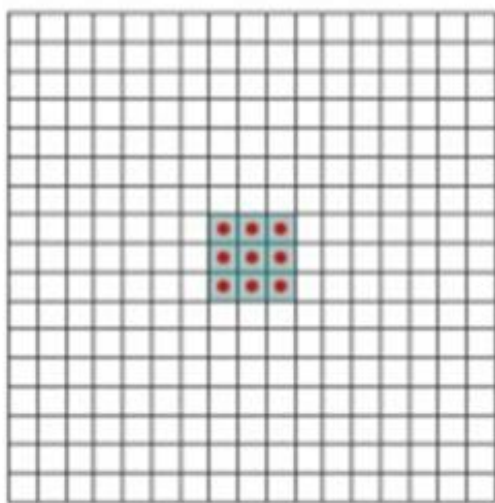
因果卷积

03

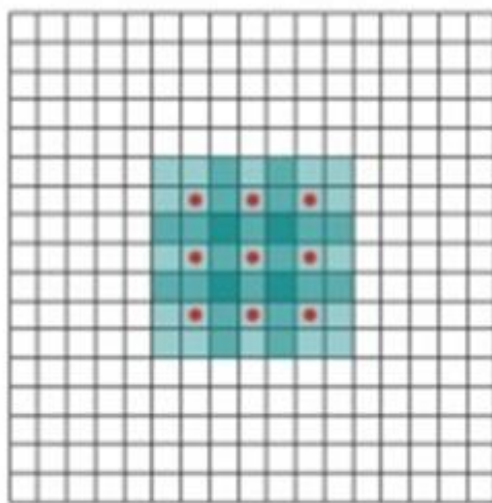


经典模型：Graph WaveNet (IJCAI 2019)

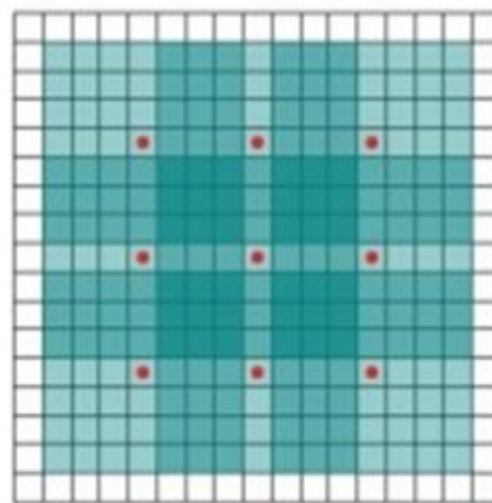
扩张因果卷积：替换传统叠加CNN+池化层



(a)



(b)



知乎 @阮糖糖



- **METR-LA**, 洛杉矶高速公路数据集, 207个传感器并收集了连续4个月的数据

备注: 数据应用在DCRNN文章中。

- **PeMS04**, 加利福尼亚高速公路数据集, shape为(307, 16992, 3), 307个传感器, 16992组数据 (5分钟采集一次, 共59天), 3维特征 (flow, occupy, speed)

- **PeMS08**, 加利福尼亚高速公路数据集, shape为(170, 17856, 3), 170个节点, 连续62天

备注: 数据应用在ASTGCN文章中。

- **TaxiBJ**, 北京出租车数据集, 原始数据shape=(5596, 2, 32, 32)

备注: 数据应用在ST-ResNet文章中。

- **NYC-Taxi**, 纽约出租车数据集, 原始数据shape=(1920, 10, 20, 2)

备注: 数据应用在STDN文章中。



- 均方根误差 (RMSE)
- 平均绝对误差 (MAE)
- 平均绝对百分比误差 (MAPE)



谢谢观看

MANY THANKS !

23.1.10