

《基于深度学习的交通流量预测与优化》 模型实验报告

2020 年 11 月 3 日

1 预备知识

1.1 时空数据

时空数据 (spatial temporal data), 即ST数据, 是具有空间坐标和时间戳的数据, 包含了空间维度和时间维度的信息。在现实中, 交通运输、公共卫生、环境科学、犯罪分析、电子商务、社交网络等许多领域的数据都属于时空数据, 如何从时空数据中发掘有价值的信息, 为实际应用提供指导和帮助具有重要研究意义, 也是当下数据挖掘前沿研究的热点之一。

与图像、时间序列等数据相比, ST数据最大的特点在于其同时具备空间和时间的属性[1]。

1. 空间属性

一方面, 地理层次影响数据的粒度, 地理层次更高的数据具有更粗的粒度级。另一方面, 数据之间的相关性和相似性受地点之间的距离的影响, 距离相近的数据之间的相似性更高。同时, 一个位置的数据会受到周围和特定地点数据的影响, 一般来说, 距离越近, 或地点之间的关联性越强, 对该地点的数据的影响越大。

2. 时间属性

每个位置的数据按照时间顺序排列, 形成时间属性。第一, ST数据具有邻近性, 相邻时间的数据的相似性比更遥远的数据要高。第二, ST数据具有周期性, 通常存在一定的周期模式, 例如交通网络每日的高峰期和低峰期。第三, ST数据具有趋势性, 即在更长的时间跨度上, 数据的变化规律存在一定的趋势。

基于以上属性, 时空数据具有高噪声和非线性等特性, 且本身具有高度自相关性, 这些特性使得极大地增加了时空数据挖掘任务的复杂程度。

1.2 交通流量预测问题

本文研究的交通预测问题为短期交通流量预测问题。对目标预测区域 t 时刻的交通流量定义无向图 $G_t = (V_t, \epsilon, W)$, 其中 V_t 为 n 个节点的集合, 表示区域内 n 个交通流的监测点, ϵ 为边的集合, 表示监测点间的连接, $W \in \mathbb{R}^{n \times n}$ 表示带权重的邻接矩阵。

预测任务为: 建立一个函数 $F(\cdot)$, 将历史 H 个时间步的交通流量图中 n 个节点的流量向量 v_t , 映射到未来 M 个时间步的流量向量, 即

$$\hat{v}_{H+1}, \dots, \hat{v}_M = F([v_1, \dots, v_H])$$

1.3 图卷积网络

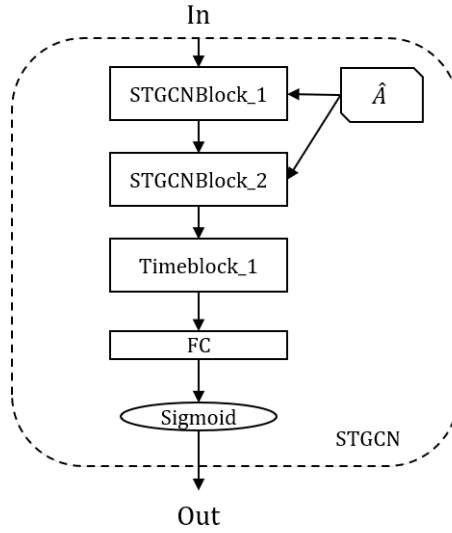
由于卷积神经网络 (CNN) 无法处理非欧式数据, 因此学者提出如何将卷积操作拓展到图数据上, 开发了图卷积网络 (Graph Convolutional Network)。其中一种方法基于图谱理论 (spectral graph theory), 利用图的傅里叶变换实现图卷积操作[2], 如下式:

$$\hat{y} = \sigma(U g_\theta(\Lambda) U^T x) \quad (1.1)$$

记图的拉普拉斯矩阵 (Laplacian matrix) 为 L , 其中 $U \in \mathbb{R}^{n \times n}$ 为 L 的特征向量, $\Lambda \in \mathbb{R}^{n \times n}$ 为 L 的特征值的对角矩阵, $g_\theta(\cdot)$ 表示要学习的参数的函数。

2 基本模型

基本模型参考了Yu, Yin和Zhu(2018)的研究[3]。模型的基本模块为STGCN模块, 该模块结合了捕捉数据空间相关性和时间相关性的两种模块, 整体结构如下图:



其中， \hat{A} 为图的邻接矩阵，将历史的图数据输入模型后，依次用两个STGCN模块充分提取数据的时间和空间特征，然后通过一个时间特征模块和全连接层，最终经过sigmoid激活函数后输出预测值。每个STGCN模块都由空间特征和时间特征组件构成。

2.1 空间特征组件

空间特征组件采用图卷积操作捕捉数据的空间相关性，为降低计算复杂度，提高效率，采用Chebyshev多项式[4]作为卷积核：

$$g_{\theta}(\Lambda) = \sum_{k=0}^{K-1} \theta_k T_k(\tilde{\Lambda})$$

其中 $T_k(\cdot)$ 是k阶的Chebyshev多项式， θ_k 是要训练的参数。 $\tilde{\Lambda}$ 为归一化后的拉普拉斯矩阵特征值对角矩阵，

$$\tilde{\Lambda} = 2\Lambda/\lambda_{max} - I$$

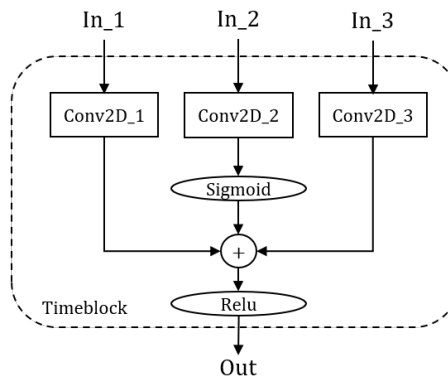
将该卷积核代入式(1.1)，则输出为

$$y_{output} = \sigma\left(\sum_{k=0}^{K-1} \beta_k T_k(\tilde{L})x\right) \quad (2.1)$$

其中 $\tilde{L} = 2L/\lambda_{max} - I$ 。

2.2 时间特征组件

时间特征组件结构如下图所示：



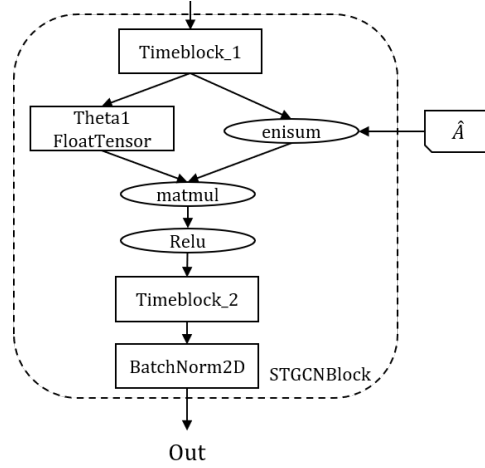
将历史三个时刻的流量图 G_{t-2}, G_{t-1}, G_t 分别输入一个一维卷积层来提取时间相关性特征，其中由中间的一维卷积层输出的特征经过sigmoid函数处理，然后将三个特征整合，用ReLU激活函数后输出时间特征 T_{t+1} 。全过程如下式：

$$T_{t+1} = \text{ReLU}(\Gamma *_{\mathcal{T}} G_{t-2} + \sigma(\Gamma *_{\mathcal{T}} G_{t-1}) + \Gamma *_{\mathcal{T}} G_t)$$

Γ 为一维卷积核。

2.3 STGCN模块

STGCN模块的具体结构如下图所示：



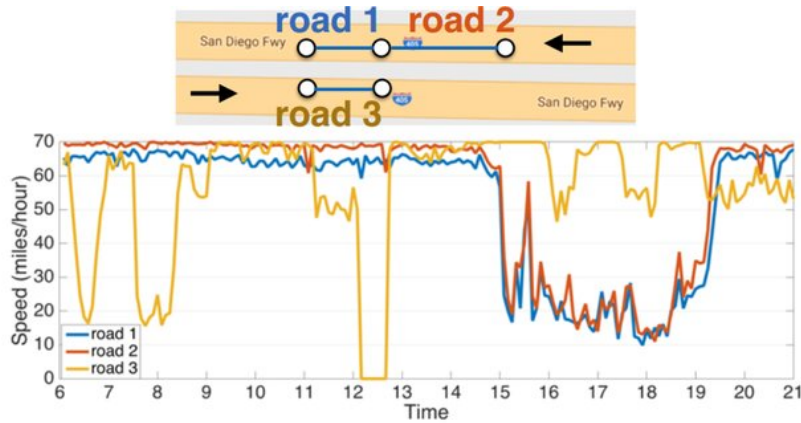
主要由两个时间特征组件和一个空间特征组件组成。输入的历史流量图数据依次经过交错排列的时间和空间特征组件，得到时空相关性特征，对提取的特征进行批处理操作后输出。

3 实验

3.1 实验设置

3.1.1 数据集

我们使用了METR-LA数据集进行试验。该交通数据集包含从洛杉矶县高速公路上的环路检测器收集的交通信息。选择了207个传感器，收集了从2012年3月1日到2012年6月30日4个月的数据。



空间相关性主要由路网结构决定。对于监测的三条道路，其中，道路1与道路2位于同一条公路上，交通速度相似；道路1和道路3在公路的相反方向。虽然欧几里得距离很近,但交通速度差异较大。

3.1.2 数据处理

1. 线性插值填充缺失值数据集中往往有大量的缺失值，会对训练造成影响，使用线性插值对缺失值进行填补。

设缺失值前后两个时间片的观测值为 y_1, y_2 ，则使用线性插值将其补为：

$$\frac{y_1 + y_2}{2}$$

2. Z-score 归一化

Z-Score归一化通过 $(x - \mu)/\sigma$ 将两组或多组数据转化为无单位的Z-Score分值，使得数据标准统一化，提高了数据可比性。

3. 基于距离计算邻接矩阵邻接矩阵A代表了图的拓扑结构,而道路的长度决定了节点间相互影响的程度.将归一化后的距离作为邻接矩阵的值.

$$w_{ij} = \begin{cases} \exp\left(-\frac{d_{ij}^2}{\sigma^2}\right), & i \neq j \text{ and } \exp\left(-\frac{d_{ij}^2}{\sigma^2}\right) \geq \varepsilon \\ 0, & \text{otherwise} \end{cases}$$

3.1.3 超参数设置

学习率为0.01，时间片间隔为5分钟，预测接下来5分钟的交通情况。

3.1.4 评价指标

我们使用在交通预测中常用的三个指标对模型的性能进行评价。

- Mean Absolute Error(MAE)，表示成对观测值之间误差的量度。

$$\text{MAE} = \frac{\sum_{i=1}^n |y_i - x_i|}{n} = \frac{\sum_{i=1}^n |e_i|}{n}$$

- Mean Absolute Percentage Error(MAPE)，为每个绝对误差的和除以实际值。

$$\text{M} = \frac{1}{n} \sum_{t=1}^n \left| \frac{A_t - F_t}{A_t} \right|$$

- Root Mean Squared Error (RMSE)，为误差平方平均值的平方根。

3.2 单步实验结果

指标	实验值
MAE	3.58
MAPE	11.23
RMSE	7.76

模型在各项指标上都取得了较为满意的结果，MAE和RMSE与同类模型相比较都有明显优势，体现出了图卷积神经网络的强大。但MAPE明显偏高，说明对于异常值的处理不够准确，无法捕捉因天气等产生的异常情况。

4 未来展望

目前模型在交通流量预测中，外部因素对其影响很大，也往往是产生异常值的原因，在接下来研究中如何引入外部数据到模型中是研究的重点。对异常情况的处理也是值得考虑的。同时对于基本模型，可以考虑使用更能捕捉时间依赖性的LSTM或GRU对时间相关性组件进行改进，提高预测精确度。

参考文献

- [1] Zhang, J. , Zheng, Y. , Qi, D. , Li, R. , & Yi, X. . (2016). DNN-based prediction model for spatio-temporal data. the 24th ACM SIGSPATIAL International Conference. ACM.
- [2] Bruna, J. , Zaremba, W. , Szlam, A. , & Lecun, Y. . (2013). Spectral networks and locally connected networks on graphs. Computer ence.
- [3] Yu, B. , Yin, H. , & Zhu, Z. . (2018). Spatio-Temporal Graph Convolutional Networks: A Deep Learning Framework for Traffic Forecasting. Twenty-Seventh International Joint Conference on Artificial Intelligence IJCAI-18.
- [4] Hammond, D. K. , Vandergheynst, P. , & Rémi Gribonval. (2011). Wavelets on graphs via spectral graph theory. Applied & Computational Harmonic Analysis, 30(2), 129-150.