

基于 DCGRU-RF 模型的路网短时交通流预测



熊 亭¹ 戚 湧¹ 张伟斌²

1 南京理工大学计算机科学与工程学院 南京 210094

2 南京理工大学电子工程与光电技术学院 南京 210094

(609688837@qq.com)

摘 要 随着城市化进程的加快,我国城市机动车数量快速增加,使得现有路网容量难以满足交通运输需求,交通拥堵、环境污染、交通事故等问题与日俱增。准确高效的交通流预测作为智能交通系统的核心,能够有效解决交通出行和管理方面的问题。现有的短时交通流预测研究往往基于浅层的模型方法,不能充分反映交通流特性。文中针对复杂的交通网络结构,提出了一种基于 DCGRU-RF(Diffusion Convolutional Gated Recurrent Unit-Random Forest)模型的短时交通流预测方法。首先,使用 DCGRU(Diffusion Convolutional Gated Recurrent Unit)网络刻画交通流时间序列数据中的时空相关性特征;在获取数据中的依赖关系和潜在特征后,选择 RF(Random Forest)模型作为预测器,以抽取的特征为基础构建非线性预测模型,得出最终的预测结果。实验以两条城市道路中的 38 个检测器为实验对象,选取了 5 周工作日的交通流数据,并将所提方法与其他常见交通流量预测模型进行比较。结果表明,DCGRU-RF 模型能够进一步提高预测精度,准确度可达 95%。

关键词: 交通流量预测;组合预测模型;深度学习;路网

中图法分类号 TP181

Short-term Traffic Flow Prediction Based on DCGRU-RF Model for Road Network

XIONG Ting¹, QI Yong¹ and ZHANG Wei-bin²

1 School of Computer Science and Engineering, Nanjing University of Science and Technology, Nanjing 210094, China

2 School of Electronic and Optical Engineering, Nanjing University of Science and Technology, Nanjing 210094, China

Abstract With the acceleration of urbanization, the number of motor vehicles in cities in China is increasing rapidly, which makes the existing road network capacity difficult to meet the transportation needs, traffic congestion, environmental pollution and traffic accidents are increasing day by day. Accurate and efficient traffic flow prediction, as the core of ITS, can effectively solve the problems of traffic travel and management. The existing short-term traffic flow prediction researches mainly use the shallow model method, so they cannot fully reflect the traffic flow characteristics. Therefore, this paper proposed a short-term traffic flow prediction method based on DCGRU-RF model for complex traffic network structure. The DCGRU network is used to characterize the spatio-temporal correlation features in the traffic flow time series data. After obtaining the dependencies and potential features in the data, the RF model is selected as the predictor, and the nonlinear prediction model is constructed based on the extracted features, and finally getting the prediction result. In this experiment, 38 detectors in two urban roads were selected as experimental objects, traffic flow data of five working days were selected, and the proposed model was compared with other common traffic flow prediction models. The experimental results show that DCGRU-RF model can further improve the prediction accuracy, the accuracy can reach 95%.

Keywords Traffic flow forecast, Combined forecasting model, Deep learning, Road network

1 引言

20 世纪末期,智能交通系统(Intelligent Traffic System, ITS)的概念被提出,将各学科先进的技术方法融合运用于交

通运输、服务控制等领域^[1]。准确高效的交通流预测作为 ITS 的核心之一,是正确引导车辆通行的关键,预测的结果会直接影响交管部门的控制方案、人们外出的路线选择,进而影响路网交通的使用效率和系统的先进性,因此研究道路网的

到稿日期:2019-01-25 返修日期:2019-05-20 本文已加入开放科学计划(OSID),请扫描上方二维码获取补充信息。

基金项目:国家重点研发计划政府间国际科技创新合作重点专项(2016YFE0108000);江苏省重点研发计划(产业前瞻与共性关键技术)项目(BE2017163);江苏省交通运输科技与成果转化项目(2018Y51)

This work was supported by the Key Program for International S&T Cooperation Projects of China(2016YFE0108000), Key R & D plan of Jiangsu Province(BE2017163), Jiangsu Transportation Technology and Achievement Transformation Project(2018Y51).

通信作者:戚湧(790815561@qq.com)

短时交通流预测的理论和方法具有深刻的现实意义和实际应用价值。

目前,研究人员已经开发了一些短时交通流量预测模型,来协助交管部门引导出行路线和提高车辆流动的效率。已提出的模型大致分为 3 类^[2]:基于线性系统理论的模型、基于知识发现的智能预测模型、组合预测模型。

第一类模型的可解释性强,其理论便于理解,是最早应用于交通流预测领域的方法,包括时间序列模型、卡尔曼滤波模型等。Levin 等^[3]使用 Box-Jenkins 时间序列分析方法来预测高速公路的交通流量,并通过多次实验证明 ARIMA(0,1,1) 具有最显著的统计意义。Vasanth 等^[4]为了弥补输入数据量过大导致数据难以获取和利用的缺点,提出 SARIMA 模型并使用少量的训练集对交通流进行短期预测。Guo 等^[5]提出了自适应卡尔曼滤波方法,其包含随机结构的定义、状态空间的表示和自适应卡尔曼递归。实验证明,该方法具有优越的预测性能。

第二类模型不需要过多的先验知识,侧重于对真实数据的拟合效果,能较好地刻画交通流数据中的随机性和非线性特点。Wu 等^[6]使用 SVR 方法预测旅行时间,实验结果表明 SVR 预测器可以显著降低相对平均误差和预测通行时间的均方根误差。Hou 等^[7]讨论了工作区的流量预测问题,并根据工作区的特征(如工作类型、封闭车道数量、车道宽度等)研究了包括随机森林模型在内的 4 种预测模型。Ma 等^[8]将路网交通流量表示为图像,并采用卷积神经网络(Convolutional Neural Network, CNN)的深度学习架构来提取图像所包含的时空流量特征,实现了对路网交通速度的预测。

第三类组合预测模型是由 Bates 等于 1969 年提出的一种预测策略,旨在弥补单一基础模型无法完全刻画交通流特性的缺陷,其通过组合不同模型独有的优势来获得更可靠、更有效的预测结果。Hu 等^[9]为了实现更高的精度和更短的学习时间,提出了一种 PSO-SVR 预测方法,该方法使用粒子群优化(PSO)搜索最优 SVR 参数。为了找到更适合搜索 SVR 参数的 PSO 方法, Hu 等分析了 3 种处理粒子偏离搜索空间的策略,并选择其中的最佳策略。由于路网中多个路段之间存在时空相关性,Chen 等^[10]利用该特性提出了一种组合模型。首先利用改进后的遗传算法挑选出重要的特征集合,并优化 LSSVR 模型参数;然后由训练好的 LSSVR 模型得出路网流量预测结果。Wu 等^[11]提出了一种结合 CNN 和 LSTM 的新型深度架构来预测未来的交通流量。其利用一维 CNN 来捕获交通流的空间特征,并利用两个 LSTM 来挖掘交通流的短期变化和周期性。

近年来,随着道路基础设施的建设和交通数据采集技术的提高,交通流预测不再局限于道路中的单一检测点。基于路网结构的多检测点预测由于更能满足道路通行的需要而受到越来越多的关注。深度学习作为一种基于无监督特征学习和特征层次结构的学习方法,使用多层体系结构来提取数据中的固有特征。作为一种新兴的机器学习方法,深度学习已引起学术界和工业界的广泛兴趣^[12]。此外,随机森林(Random Forest, RF)模型^[13]作为一种集成学习方法,将大量弱模

型组合构建成强模型,可以有效地分析交通大数据中的非线性特点。因此,本文提出基于 DCGRU-RF 模型的路网短时交通流预测方法,利用扩散卷积门限单元(Diffusion Convolutional Gated Recurrent Unit)循环神经网络^[14],对流量数据中的时空相关性进行建模,提取数据中的潜在特征,并且选择 RF 模型作为预测器,得出未来时间间隔内的交通流预测结果,充分揭示了交通流数据中的时空相关性和非线性特点。

此外,本文只对基于路网结构的交通流预测模型进行研究和验证,对于如何提高大型路网中多检测点的预测效率本文不做过多的描述。

2 DCGRU 网络和 RF 模型

2.1 GRU 网络

为简化运算,提高计算效率,Cho 等^[15]于 2014 年提出了 GRU 网络,该网络是以 LSTM 网络为基础的改进型 RNN 网络,即将输入门与遗忘门合并为一个重置门。GRU 包含两个门:重置门和更新门。GRU 的结构如图 1 所示。

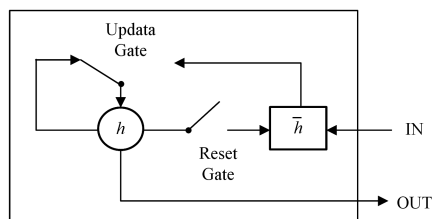


图 1 GRU 单元的内部结构

Fig. 1 Internal structure of GRU unit

(1)重置门。用于控制前一时间刻状态信息的省略程度,重置门的值越大代表忽略的程度越低,当前输入值的影响就越小。定义 r_t 为 t 时刻的重置门,其计算式如式(1)所示:

$$r_t = \sigma(W_r \cdot [h_{t-1}, x_t]) + b_r \quad (1)$$

其中, $W_r = [W_{rx}, W_{rh}]$ 表示输入数据到重置门、上一时刻激活状态到重置门的权重矩阵; x_t 为 t 时刻的输入值; h_{t-1} 是上一时刻的激活值; σ 表示激活函数,通常为 sigmoid 函数。

(2)更新门。用于决定前一时间刻激活值 h_{t-1} 对当前时刻激活值的影响程度,且前一时间刻的状态信息值与更新门的值呈正相关。定义 z_t 为 t 时刻的更新门,其计算式如式(2)所示:

$$z_t = \sigma(W_z \cdot [h_{t-1}, x_t]) + b_z \quad (2)$$

其中, $W_z = [W_{zx}, W_{zh}]$ 表示当前时刻输入数据到更新门、上一时刻激活状态到更新门的权重矩阵。

与 LSTM 类似,GRU 神经网络结构中也包含一个候选激活值 \tilde{h} 和当前时刻激活值,两者数值的大小与重置门和更新门有关,其计算式分别如式(3)和式(4)所示:

$$\tilde{h}_t = \tanh(W_{\tilde{h}} \cdot [r_t * h_{t-1}, x_t]) + b_{\tilde{h}} \quad (3)$$

$$h_t = (1 - z_t) * h_{t-1} + z_t * \tilde{h}_t \quad (4)$$

其中, $W_{\tilde{h}} = [W_{\tilde{h}x}, W_{\tilde{h}h}]$ 。

由上述公式可知,更新门和重置门的作用是控制信息依赖的时间尺度,当 r_t 值较大时,过往信息影响较小,GRU 只能支持短时依赖;当 z_t 值较大时,过往信息影响较大,GRU

可以支持长时信息依赖。依据前向传播算法和误差反向传递机制,可以得到网络参数的求导公式。

2.2 DCGRU 网络

复杂网络不仅研究网络的静态结构,还关注网络的动力学特性,随机游走属于复杂网络中的一种动态过程,是一系列由连续随机步组成的轨迹的数学表示^[16]。给定网络 $G=(V, E)$ 和起始节点 $v \in V$, 随机选择一个邻居节点, 将其移动到该邻居; 然后, 再次选择这个新节点的邻居并移动它, 以此类推。以这种形式选择的节点随机序列称为网络上的随机游走。这一随机过程未来的状态在条件上独立于过去的状态, 只与当前状态有关, 即某一节点移动到邻居节点的概率是独立于过去的轨迹的。

由此, 将道路网络表示成有向加权图 $G=(V, E, W)$, 其中 V 表示道路网中 N 个道路检测点的集合; E 表示道路网中的链路集合; $W \in \mathbb{R}^{N \times N}$ 表示节点接近度 (如其道路网络距离的函数) 的加权邻接矩阵。将在图 G 上观察到的交通流表示为图信号 $X \in \mathbb{R}^{N \times P}$, 其中 P 是每个节点的特征数量 (如速度、流量特征)。设 $X^{(t)}$ 表示在时间戳 t 处观察到的图形信号。该扩散过程可表征为在图 G 上的随机游走, 其中重启概率 $\alpha \in [0, 1]$, 状态转移矩阵为 $D_0^{-1}W$ 。矩阵 $D_0 = \text{diag}(W1)$ 为出度对角矩阵, $1 \in \mathbb{R}^N$ 表示值为 1 的 N 维向量。根据随机游走理论, 在游走 n 步之后, 这种马尔可夫过程收敛到一个稳定分布的状态 $P \in \mathbb{R}^{N \times N}$, 其中第 i 行 $P_{i,:} \in \mathbb{R}^N$ 表示节点 $v_i \in V$ 扩散至其他节点的可能性。引理 1 给出了稳定分布的闭式^[17]。

引理 1 扩散过程的稳定分布可以表示为无向图上无限步随机游走的加权组合, 形式如下:

$$P = \sum_{k=0}^{\infty} \alpha (1-\alpha)^k (D_0^{-1}W)^k \quad (5)$$

其中, k 代表扩散步数。在实际中, 通常使用一个明确的值来截断扩散过程, 并针对每步截断分配一个可训练的权重。本文所指的扩散过程是双向的, 即包括反向的一个扩散过程。增加这一环节主要是因为包含反向传递的双向扩散能为模型提供更大的灵活性, 能够捕获来自上下游流量的影响。

在引理 1 的基础上, 利用图信号 $X \in \mathbb{R}^{N \times P}$ 和滤波器 f_θ 定义一种扩散卷积算子, 有向图结构中有限 K 步双向随机游走的加权组合, 如式(6)所示:

$$X_{:,p} *_{\mathcal{G}} f_\theta = \sum_{k=0}^{K-1} (\theta_{k,1} (D_0^{-1}W)^k + \theta_{k,2} (D_1^{-1}W^{-1})^k), \quad p \in \{1, \dots, P\} \quad (6)$$

其中, $\theta \in \mathbb{R}^{K \times 2}$ 是滤波器中的参数, $D_0^{-1}W$ 和 $D_1^{-1}W^{-1}$ 分别表示正反向扩散过程的状态转移矩阵。虽然卷积计算需要耗费大量计算资源, 但是如果传感器网络所构成的图 G 是一个稀疏图, 那么就可使用复杂度为 $O(K)$ 的递归稀疏-密集矩阵乘法来完成式(6)的计算, 使得总时间复杂度为 $O(K|\mathcal{E}|)$, 其远远小于 $O(N^2)$ 。

利用式(7)中定义的卷积运算构建一个扩散卷积层, 将 P 维特征映射到 Q 维输出。定义参数张量为: $\Theta \in \mathbb{R}^{Q \times P \times K \times 2} = [\Theta]_{q,p}$, 其中 $\Theta_{q,p,:} \in \mathbb{R}^{K \times 2}$ 表示对应第 p 维输入和第 q 维输出的卷积过滤器的参数。因此, 扩散卷积层可以表示为:

$$H_{:,q} = \alpha \left(\sum_{p=1}^P X_{:,p} *_{\mathcal{G}} f_{\Theta_{q,p,:}} \right), q \in \{1, 2, \dots, Q\} \quad (7)$$

其中, $X \in \mathbb{R}^{N \times P}$ 是输入数据; $H \in \mathbb{R}^{N \times Q}$ 是维度降低后的输出数据; α 表示激活函数, 例如 ReLU, Sigmoid 函数。扩散卷积层能够学习图结构化的数据表示, 可以采用基于随机梯度的方法对其进行训练。

为进一步提取数据中的固有特性, 表征数据中的时间依赖关系, 本文使用循环神经网络 RNN 的变体 GRU 来模拟交通数据中的时间相关性, 利用上文提到的扩散卷积算子代替 GRU 网络中的矩阵乘法, 由此引出 DCGRU 网络, 如式(8)~式(11)所示:

$$r_t = \sigma(\Theta_r *_{\mathcal{G}} [x_t, h_{t-1}] + b_r) \quad (8)$$

$$z_t = \sigma(\Theta_u *_{\mathcal{G}} [x_t, h_{t-1}] + b_z) \quad (9)$$

$$\tilde{h}_t = \tanh(\Theta_c *_{\mathcal{G}} [x_t, (r_t) \odot h_{t-1}] + b_h) \quad (10)$$

$$h_t = z_t \odot h_{t-1} + (1 - z_t) \odot \tilde{h}_t \quad (11)$$

其中, x_t 和 h_t 表示在时刻 t 交通流量的输入数据和隐藏层的输出数据, r_t 和 z_t 表示在时刻 t 的重置门和更新门, $*_{\mathcal{G}}$ 对应式(7)定义的扩散卷积算子, Θ_r, Θ_u 和 Θ_c 表示重置门、更新门和激活状态过滤器的参数。DCGRU 网络使用 BPTT 算法训练参数。与传统的 GRU 模型相比, DCGRU 的区别在于: 在更新门 $u^{(t)}$ 、重置门 $r^{(t)}$ 和记忆单元时使用了扩散卷积算子, 使得空间依赖性能被直观地解释和有效地计算。

2.3 RF 模型

RF 模型作为数据挖掘和机器学习的强大工具, 是一种将大量的回归树结合继而得出预测结果的集成学习方法。RF 的预测过程可以通过评估预测因子的重要程度来直观地解释, 该算法对于数据中的噪声和异常值具有鲁棒性, 可以有效地运行在交通大数据上, 对高维数据有很好的适应性^[18]。为了得到测试数据的预测值, 所有的输入特征都会被代入 R 棵回归树中, 每棵树都会产生一个预测值。最后, 对所有回归树的预测结果进行平均, 以获得模型输出值。

3 基于 DCGRU-RF 的路网交通流预测模型

本文利用 DCGRU 网络能充分挖掘交通流时空相关性的优势以及 RF 模型对非线性数据的良好拟合能力, 依据组合模型的思想, 提出了一种 DCGRU-RF 组合预测模型, 旨在利用时空相关性的约束条件, 选用具有代表性的特征来提高模型的可解释性以及预测精度。该模型的结构如图 2 所示。

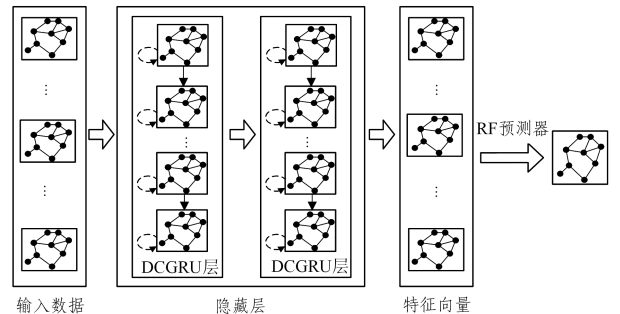


图 2 DCGRU-RF 模型结构

Fig. 2 DCGRU-RF model structure

因为本文只探讨预测模型的有效性, 所以对路网结构的简化工作不做过多阐述。假设选择路网中的 r 个检测点

开展预测工作,且为每个检测器构造 p 个特征,则基于路网结构观测到的交通流数据可以表示为 $\mathbf{X} \in \mathbb{R}^{r \times p}$ 。将 DCGRU 网络的一条输入数据表示为 $[\mathbf{X}^{t-m\Delta t}, \mathbf{X}^{t-(m-1)\Delta t}, \dots, \mathbf{X}^t]$ 的形式,其中 Δt 表示时间间隔, m 表示历史时间窗口大小,则有如下表示关系:

$$\mathbf{X}^t = [\mathbf{X}_1^t, \mathbf{X}_2^t, \mathbf{X}_i^t, \dots, \mathbf{X}_r^t]^T \quad (12)$$

$$\mathbf{X}_i^t = [z_{i,1}, z_{i,2}, \dots, z_{i,p}]^T \quad (13)$$

其中, \mathbf{X}_i^t 代表 t 时刻检测器 i 的特征向量。特征向量集由检测器 i 及其相邻路段当前时刻以及前 t' 个时刻的交通流量和速度构成。对应的标签数据为 $[\mathbf{X}^{t-(m-1)\Delta t}, \mathbf{X}^{t-(m-2)\Delta t}, \dots, \mathbf{X}^{t+\Delta t}]$ 。在训练过程中使用 MAPE 作为损失函数,并通过批次随机梯度下降对模型进行学习。设 DCGRU 网络学习完毕后得出的输出向量为 \mathbf{H} ,则第 i 个输出向量 $\mathbf{H}_i \in \mathbb{R}^{m \times r \times q}$ 如式(14)所示:

$$\mathbf{H}_i = \Phi(\mathbf{E}_i) \quad (14)$$

其中, Φ 表示 DCGRU 网络模型, \mathbf{E}_i 表示第 i 条输入数据。

学习到交通流数据中的有用特征后,将输出向量 \mathbf{H} 作为 RF 模型的输入数据集,对应的标签数据为样本下一时刻的实际交通流量值。待模型训练完毕,通过预测器 RF 得出路网中 r 个检测器在下一时刻的流量数据,如式(15)所示:

$$\hat{\mathbf{y}}_i = f(\mathbf{H}_i) \quad (15)$$

其中, f 代表预测模型——RF 模型, $\hat{\mathbf{y}}_i \in \mathbb{R}^{r \times 1}$ 代表第 i 个样本下一时刻的交通流量预测值。

本文将基于 DCGRU-RF 模型的路网短时交通流预测分为 4 个步骤:

- (1) 依据式(12)、式(13)构造路网数据特征,并组成模型的训练集。
- (2) 采用 DCGRU 模型对交通流量数据的时空相关性进行建模,通过批次随机梯度下降对模型进行学习。
- (3) 利用学习好的 DCGRU 模型提取具有鉴别性的特征,得到式(14)所示的输出特征向量 \mathbf{H} 。
- (4) 将提取到的特征代入 RF 模型,并利用训练好的 RF 模型输出路网中 r 个检测点的流量预测值 $\hat{\mathbf{Y}}$ 。

4 实验结果及分析

4.1 数据来源与预处理

本文所采用的实验数据来源于某城市的两条高速公路,设两条道路的编号为 A 和 B。图 3 为道路 A 和 B 的十字路口构成的高速路网络结构分布图,其中高速路 A 为东西走向,高速路 B 为南北走向。

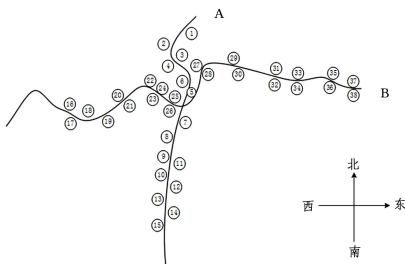


图 3 高速路网络结构

Fig. 3 Highway network structure

本文选取路网中 38 个检测器的数据用于特征学习和模型训练。为方便表示,用数字表示检测器编号,并将其标识在图 3 中的对应位置,其中选取高速路 A 中的 23 个检测器,高速路 A 南侧的检测器采集从西向东的交通流量数据,北侧的检测器则相反;选取高速路 B 上的 15 个检测器,高速路 B 西侧检测站采集从北向南的交通流量数据,东侧的检测器则相反。

本文采用的实验数据来自上述 38 个检测器,实验数据的时间跨度为 2015 年 7 月 27 日至 2015 年 8 月 31 日,数据的采样间隔为 5min。由于交通流数据在周末与工作日会呈现不同的规律,因此剔除周末的数据,只对周一至周五的流量进行分析研究。使用 70% 的数据用于模型的训练,10% 的数据用于验证,20% 的数据用于测试。

本文使用均值填充法对缺失值进行填补,其计算式如式(16)所示:

$$f_t = \frac{f_{t+1} - f_{t-1}}{2} \quad (16)$$

其中, f_t 表示 t 时刻的交通流缺失值, f_{t+1} 表示下一时刻的交通流数据, f_{t-1} 表示上一时刻的交通流数据。

使用 Z-score 正则化方法对数据进行归一化处理,处理后的交通流数据均值 $\mu = 0$,标准差 $\sigma = 1$ 。其计算式如式(17)所示:

$$f_t' = \frac{f_t - \bar{f}}{\sigma_f} \quad (17)$$

其中, \bar{f} 为原始交通流数据的平均值, f_t 为 t 时间间隔的交通流数据, σ_f 为原始交通流数据的标准差。

4.2 模型评价指标

为了公正评价预测模型的实际效果,需要选择评价指标对模型进行评估。给定一个交通流预测场景,假设 t 时刻某检测点的交通流量实际值为 $\hat{\mathbf{y}}_t$,预测模型学习训练后得到该检测点在 t 时刻的预测数值为 \mathbf{y}_t , N 表示测试样本的个数。本文选择以下常用的误差计算函数。

(1) 均方根误差^[19]。该误差表示所有测试样本的预测值与实际值差值的平方和与测试样本数相除后的平方根。

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (\mathbf{y}_t - \hat{\mathbf{y}}_t)^2} \quad (18)$$

(2) 绝对平均百分比误差^[19]。该数值是绝对平均误差的百分比数值,代表预测值和观测值之间的相对偏离程度。

$$\text{MAPE} = \frac{1}{N} \sum_{i=1}^N \left| \frac{\mathbf{y}_t - \hat{\mathbf{y}}_t}{\mathbf{y}_t} \right| \quad (19)$$

4.3 模型参数设定

在交通流量预测研究中,预测的精度和性能不仅受模型选择和构建的影响,还与模型参数的选择紧密相关。若模型的参数设置不当,则会影响预测模型和方法的有效性,使预测精度低于预期。然而研究工作中并没有明确的理论对参数设定提供指向性建议,只能通过实验对比来选择较优的参数值。

本文主要分析的模型参数有:特征构造中的相邻检测器点的数目 d , DCGRU 网络中的隐藏层节点数 N 、历史时间间隔长度 m 。

(1) 相邻检测器数目 d 对路网短时交通流预测结果的影响在复杂的道路网络中,影响预测精度的节点不再局限于

上下游位置,还包括其他路段中的相邻点,因此合理地选择相邻检测器的数目 d ,将对预测效果产生较大的影响。若 d 值过大,则会造成维数过高、运行时间过长、浪费计算资源的问题;若 d 值过小,则会出现预测效果不佳的现象。图 4 给出了两者之间的变化关系。

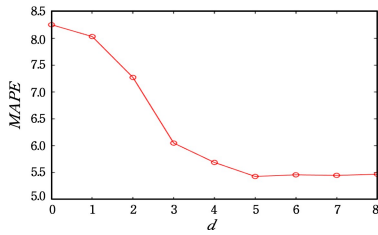


图 4 相邻检测器数 d 与 MAPE 的关系

Fig. 4 Relationship between number of adjacent detectors and MAPE

由图 4 可知,当 d 值过小时,检测器之间的空间相关性未被充分考虑,而当 d 值过大时,一些关联性较小的检测器也参与到特征构建中,导致冗余信息增多,预测效果并没得到显著改善,却造成了相当程度的计算资源浪费。因此,应从预测效率与系统资源消耗两方面折中选取相邻检测器的数目 d 。

(2) 隐藏层节点数 N 对路网短时交通流预测结果的影响

本文中的隐藏层是指 DCGRU 网络,隐藏层的构造会直接影响路网特征的学习程度,进而影响预测的效果。由于本文的数据量以及维度并不是很大,因此将隐藏层的层数固定为 2,现只需要根据实验数据尝试不同的隐藏层节点数 N ,以获取使模型效果最优的参数值。隐藏层节点数 N 与 MAPE 的关系如图 5 所示。由图 5 可知,当 N 过小时,由于欠拟合导致模型预测效果不佳,但当 N 过大时,又会出现过拟合的现象。在多次实验后发现, $N=30$ 时,预测误差最小。

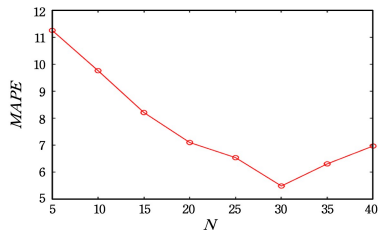


图 5 隐藏层节点数 N 与 MAPE 的关系

Fig. 5 Relationship between number of hidden layer nodes and MAPE

(3) 历史时间窗口大小 m 对路网短时交通流预测结果的影响

本文设置隐藏层的节点数为 30,相邻检测器节点数 d 为 4,在此基础上分析 m 对实验结果的影响,结果如图 6 所示。

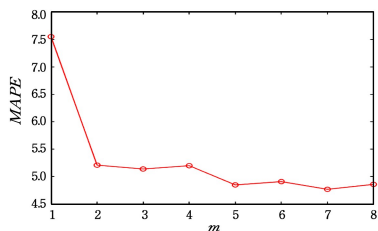


图 6 历史时间窗口大小 m 与 MAPE 的关系

Fig. 6 Relationship between historical time window size and MAPE

由图 6 可知,随着 m 的增大,MAPE 值逐步减小,当 m 增大到一定程度时,MAPE 值趋于稳定,因此可以将 m 设为 5,用于特征学习和模型训练。

4.4 预测结果及对比分析

本文依据 4.3 节的实验结果,选定最优参数值来完成 DCGRU-RF 模型的训练,并将最终预测结果以图表的形式展示,同时结合其他常见预测模型分析流量预测值与实际流量值之间的误差。此外,本文主要完成工作日的短时交通流预测工作。

本文设置用于特征构造的相邻检测器的数目 d 为 4,历史时间段长度 m 为 5,两层 DCGRU 网络中的隐藏层节点数 N 都为 30,训练迭代次数为 1500,批大小为 64;预测器 RF 的弱学习器的最大迭代次数为 320,决策树的最大深度为 20,最小样本数为 9,叶子节点最小样本数为 10,最大特征数为 16。本节选择 18 号检测器,以 2015 年 8 月 26 日的流量预测值和实际流量值为分析对象,对比结果如图 7 所示。

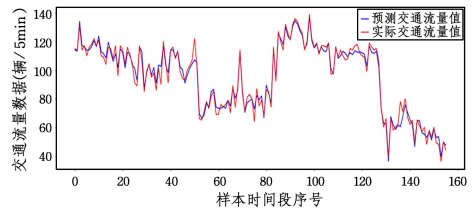


图 7 18 号检测器的交通流量预测结果

Fig. 7 Forecast results of traffic flow of No. 18 detector

由图 7 可知,DCGRU-RF 模型的拟合效果总体表现较好,虽然高峰期的流量变化波动较大,但模型能较好地捕捉到它的变化趋势。为了进一步表明模型的优越性,本文选取 GRU^[20],DCGRU^[14],LSTM-SVR^[21] 这 3 个模型分别对整个道路网中的检测器进行预测,并对上述模型的预测效果(见表 1)进行对比分析。

(1) GRU 网络模型中的隐藏层数为 2,隐藏层的节点数为 35,批大小为 64,迭代次数为 1400,学习率为 0.03。

(2) DCGRU 网络模型中的隐藏层数为 2,隐藏层的节点数为 30,批大小为 64,迭代次数为 1500,学习率为 0.03。

(3) LSTM-SVR 网络模型中,LSTM 网络的隐藏层数为 2,隐藏层节点数为 30,批大小为 64,迭代次数为 1500,学习率为 0.03;SVR 预测器的核函数为 rbf,迭代次数为 700,惩罚因子为 0.01。

本文选取 3 个时段对模型预测性能进行考查,分别是 8 点至 20 点,8 点至 10 点,14 点至 16 点。由表 1 可知,本文提出的模型的预测值与真实值之间的差值并不大,在 3 个时段中的预测效果优于其他常用的预测模型,预测准确度最高可达 95%。由于高峰期流量波动较大,因此该时段的预测精度略低于流量平稳时段。此外,4 种模型中 GRU 模型的预测效果最差,与 DCGRU-RF 模型的预测精度最多相差 3%,这可能与空间相关性特征的刻画不充分有关;LSTM-SVR 的表现效果略逊于本文提出的模型,而且 SVR 预测器的参数较多,调参过程较麻烦。通过与 3 个模型的预测误差进行对比可以发现,本文提出的 DCGRU-RF 模型具有更好的准确性。

表 1 模型预测误差对比

Table 1 Comparison of model prediction errors

时段	评价标准	GRU	DCGRU	LSTM-SVR	DCGRU-RF
8 点—20 点 (时段 1)	RMSE	8.87	7.43	6.14	5.27
	MAPE/%	6.91	5.26	4.66	3.79
8 点—10 点 (时段 2)	RMSE	9.61	7.65	6.13	5.91
	MAPE/%	7.05	5.92	4.89	4.58
14 点—16 点 (时段 3)	RMSE	8.22	7.32	5.92	4.64
	MAPE/%	6.39	5.41	4.58	3.29

结束语 本文提出了一种新的 DCGRU-RF 路网短时交通流预测模型,将扩散卷积算子应用于 GRU 网络,使其能够同时学习路网中多个节点间的动态流动性与交通流数据中存在的时间相关性,利用 RF 预测器处理交通流数据中的非线性特点,对 DCGRU 模型中学习好的特征进行继续训练,并最终输出预测结果。通过与多个现有模型的实验对比,证明了本文所提方法的有效性,预测准确度可达 95%。由于本文仅考虑了交通流数据中的时空相关性特点,并未对天气状况以及交通事故等影响因素进行考虑,因此今后的研究工作会将这些影响因素转化为模型可以学习的特征,以使模型的预测效果更佳。

参考文献

- [1] YANG D I, WU J P, ZHANG Q S. The development of intelligent transportation system (ITS) and its model research [J]. Journal of Beijing University of Aeronautics and Astronautics, 2000, 26(1): 22-25.
- [2] LV Y S, DUAN Y J, et al. Traffic Flow Prediction With Big Data: A Deep Learning Approach[J]. IEEE, 2015, 16(2): 865-873.
- [3] LEVIN M, TSAO Y D. On forecasting freeway occupancies and volumes[J]. Transportation Research Record, 1980, 773: 47-49.
- [4] VASANTHAKUMAR S, VANAJAKSHI L. Short-term traffic flow prediction using seasonal ARIMA model with limited input data[J]. European Transport Research Review, 2015, 7(3): 21.
- [5] GUO J, HUANG W, WILLIAMS B M. Adaptive Kalman filter approach for stochastic short-term traffic flow rate prediction and uncertainty quantification[J]. Transportation Research Part C: Emerging Technologies, 2014, 43: 50-64.
- [6] WU C H, HO J M, LEE D T. Travel-time prediction with support vector regression[J]. IEEE transactions on intelligent transportation systems, 2004, 5(4): 276-281.
- [7] HOU Y, EDARA P, SUN C. Traffic flow forecasting for urban work zones[J]. IEEE Transactions on Intelligent Transportation Systems, 2015, 16(4): 1761-1770.
- [8] Xiaolei M, Zhuang D, Zhengbing H, et al. Learning Traffic as Images: A Deep Convolutional Neural Network for Large-Scale Transportation Network Speed Prediction[J]. Sensors, 2017, 17(4): 818.
- [9] HU W, YAN L, LIU K, et al. A short-term traffic flow forecasting method based on the hybrid PSO-SVR[J]. Neural Processing Letters, 2016, 43(1): 155-172.

- [10] CHEN X B, LIU X, WEI Z J, et al. Short-term Traffic Flow Forecasting of Road Network Based on GA-LSSVR Model[J]. Journal of Transportation Systems Engineering and Information Technology, 2016, 17(1): 60-66.
- [11] WU Y, TAN H. Short-term traffic flow forecasting with spatial-temporal correlation in a hybrid deep learning framework[J]. arXiv: 1612. 01022.
- [12] BENGIO Y. Learning Deep Architectures for AI[J]. Foundations & Trends® in Machine Learning, 2009, 2(1): 1-127.
- [13] BREIMAN L. Random Forests [J]. Machine Learning, 2001, 45(1): 5-32.
- [14] LI Y, YU R, SHAHABI C, et al. Diffusion convolutional recurrent neural network: Data-driven traffic forecasting[J]. arXiv: 1707. 01926, 2017.
- [15] CHO K, MERRIENBOER B V, GULCEHRE C, et al. Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation[J]. arXiv: 1406. 1078.
- [16] LIPPI M, BERTINI M, FRASCONI P. Short-Term Traffic Flow Forecasting: An Experimental Comparison of Time-Series Analysis and Supervised Learning[J]. IEEE Transactions on Intelligent Transportation Systems, 2013, 14(2): 871-882.
- [17] PEARSON, KARL. The Problem of the Random Walk[J]. Nature, 1905, 72(1865): 294-294.
- [18] WANG D, ZHANG Q, WU S Y. Traffic Flow Forecast with Urban Transport Network [C] // 2016 IEEE International Conference on Intelligent Transportation Engineering. 2016: 139-143.
- [19] LUO X L, JIAO Q Q, NIU L Y, et al. Short-term traffic flow prediction based on deep learning[J]. Application Research of Computers, 2017, 34(1): 91-95.
- [20] BAI S, KOLTER J Z, KOLTUN V. An Empirical Evaluation of Generic Convolutional and Recurrent Networks for Sequence Modeling[J]. arXiv: 1803. 01271.
- [21] ZHENG Z, PAN L, PHOLSENA K. Mode Decomposition Based Hybrid Model for Traffic Flow Prediction [C] // 2018 IEEE Third International Conference on Data Science in Cyberspace (DSC). IEEE, 2018: 521-526.



XIONG Ting, born in 1993, master. Her main research interests include data mining, and traffic flow prediction.



QI Yong, born in 1970, Ph.D, professor, is a member of China Computer Federation. His main research interests include traffic big data and so on.