

摘要

近年来，随着城市化进程的加速推进，我国汽车保有量日益增长，给居民出行带来了便利。同时，交通拥堵、交通噪声、空气质量恶化等问题严重制约了城市的发展，成为全国各个城市普遍面临的难题。智能交通及各种现代技术的出现，为有效改善交通状况提供了新的思路。交通流预测作为智能交通系统的关键核心技术，已成为交通领域研究的热点。本文基于神经网络基础和图神经网络的理论与方法，建立了两种混合深度学习模型的交通流量预测，具体的研究内容包括以下两个方面：

(1) 针对交通流的非线性、时间动态性、空间相关性等特征，提出一种基于注意力机制的 CNN-BiGRU 的交通流量预测模型。该模型中，卷积神经网络(CNN)用于捕获交通流量的空间相关特征。双向门控循环单元(BiGRU)分别从前向和后向传播中都获取信息，充分提取交通流的时间相关特征。注意力机制的引入使得交通流最重要的特征可分配到最大程度的关注，从而更好的提升 CNN 网络和 BiGRU 网络提取时空特征的能力，达到提升模型的预测精度的目的。在真实数据集上的对比实验和消融实验的结果都表明：基于注意力机制的 CNN-BiGRU 的交通流预测模型在捕获时空相关性方面具有一定的优势，较其他的基线模型具有提高预测性能的优势。

(2) 针对交通流的空间动态性和多尺度时间相关性，提出一种基于时空注意力的 ChebNet 和 BiGRU 相结合的交通流量预测模型(STAM-CN-BiGRU)，该模型由多层时空残差模块组成。首先，利用 ChebNet 学习路网拓扑引起的交通流的空间依赖性，并将 BiGRU 与 CNN 相结合，获取交通流的多尺度时间相关性。然后，在交通流预测模型中引入混合时空注意模块，刻画每个节点的时空依赖关系的不同重要程度。把三个不同时间粒度的交通流量（临近序列、日周期序列、周周期序列）的预测结果进行加权融合，得到最终预测结果。最后，在真实数据集上，对所提模型的一步预测、多步预测的预测性能进行了广泛评估，并且与其他基线模型进行对比，结果表明，该模型较其他基线模型可以显著提高预测精度，特别是在多步预测方面优于其他模型。

关键词：智能交通；交通流预测；卷积神经网络；双向门控循环单元；图神经网络；残差连接；注意力机制

Abstract

In recent years, with the acceleration of urbanization, the total quantity of vehicle in our country is increasing day by day, which has brought the convenience for people to travel. At the same time, traffic congestion, traffic noise, air quality deterioration and other problems have seriously restricted the development of cities, and become a common problem faced by all cities in the country. The emergence of intelligent transportation and various modern technologies provide a new idea for improving traffic conditions effectively. Traffic flow prediction, as the key core technology of intelligent transportation system, has become a hot research topic in traffic field. Based on the theory and method of neural network and graph neural network, this paper establishes two kinds of mixed deep learning models for traffic flow prediction. The specific research contents include the following two aspects:

(1) Aiming at the characteristics of traffic flow such as nonlinearity, temporal dynamics and spatial correlation, a CNN-BiGRU model based on attention mechanism traffic flow prediction model is proposed. In this model, convolutional neural network (CNN) is used to capture the spatial correlation features of traffic flow. The bidirectional gated cycle unit (BiGRU) acquires information from both forward and backward propagation, and fully extracts the time-dependent characteristics of traffic flow. The introduction of the attention mechanism enables the most important features of traffic flow to be allocated to the maximum degree of attention, so as to better improve the ability of CNN network and BiGRU network to extract temporal and spatial features and achieve the purpose of improving the prediction accuracy of the model. The results of comparison and ablation experiments on real data sets show that the CNN-BiGRU traffic flow prediction model based on attention mechanism has certain advantages in capturing temporal and spatial correlation, and has the advantage of improving the prediction performance compared with other baseline models.

(2) Aiming at the spatial dynamics and multi-scale temporal dependence of traffic flow, a traffic flow prediction model combining ChebNet and BiGRU based on spatio-temporal attention (STAM-CN-BiGRU) is proposed, which is composed of multi-layer spatio-temporal residual modules. Firstly, ChebNet was used to learn the spatial dependence of traffic flow caused by road network topology, and BiGRU

Abstract

combined with CNN to obtain the multi-scale temporal correlation of traffic flow. Then, the mixed spatio-temporal attention module is introduced into the traffic flow prediction model to describe the different importance of the spatio-temporal dependence of each node. The forecast results of three different time granularity traffic flows (the recent segment, the daily-periodic segment, the weekly-periodic segment) were weighted and fused to get the final forecast results. Finally, on the real data set, the prediction performance of one-step prediction and multi-step prediction of the proposed model is extensively evaluated, compared with other baseline models. the results show that the proposed model can significantly improve the prediction accuracy compared with other baseline models, especially in the aspect of multi-step prediction.

Keywords: Intelligent transportation; Traffic flow prediction; Convolutional neural network; Bidirectional gated cycle unit; Graph neural network; Residual connection; Attention mechanism

目录

第 1 章 绪论	1
1.1 选题背景和研究意义	1
1.1.1 选题背景	1
1.1.2 研究意义	2
1.2 国内外研究现状	2
1.2.1 基于时间序列的预测方法	2
1.2.2 基于传统机器学习的预测方法	3
1.2.3 基于深度学习的预测方法	3
1.2.4 基于强化学习的预测方法	5
1.3 论文的研究内容与主要工作	5
1.4 论文的创新点	6
1.5 论文的组织结构	7
第 2 章 基础理论与方法	8
2.1 时间序列及时间序列的单步、多步预测	8
2.2 交通流的基础知识	9
2.2.1 交通流的参数	9
2.2.2 交通流数据特征	9
2.3 数据预处理	10
2.3.1 数据清洗	11
2.3.2 数据集成	12
2.3.3 数据变换	12
2.3.4 数据归约	12
2.4 神经网络基础	13
2.4.1 卷积神经网络	13
2.4.2 双向 GRU 模型结构	14
2.4.3 注意力机制	15
2.4.4 残差连接	16
2.5 图神经网络	17

2.5.1 图卷积神经网络	17
2.5.2 图注意力网络.....	19
2.6 本章小结	19
第 3 章 基于 CNN-BiGRU-Attention 的交通流量预测.....	20
3.1 引言.....	20
3.2 交通流量预测模型框架.....	20
3.3 数据集和实验参数设置.....	21
3.3.1 数据集和实验参数设置.....	21
3.3.2 实验参数设置及实验环境配置.....	21
3.3.3 评价指标.....	22
3.4 实验结果分析	22
3.5 本章小结	24
第 4 章 基于 STAM-CN-BiGRU 的交通流量预测	25
4.1 引言	25
4.2 构建 STAM-CN-BiGRU 混合神经网络模型	25
4.2.1 空间依赖性建模	26
4.2.2 时空注意力模块	28
4.2.3 多组件融合	29
4.3 数据集及参数设置	31
4.3.1 实验数据集	31
4.3.2 实验参数设置	32
4.4 PEMS04 实验结果和分析	32
4.4.1 基线模型.....	32
4.4.2 PEMS04 数据集的一步预测	32
4.4.3 PEMS04 数据集的多步预测.....	33
4.4.4 PEMS04 数据集的消融实验.....	34
4.5 PEMS08 实验结果和分析	35
4.5.1 PEMS08 数据集的一步预测	35
4.5.2 PEMS08 数据集的多步预测.....	36
4.5.3 PEMS08 数据集的消融实验	37
4.6 结论.....	38
第 5 章 总结与展望	40

5.1 研究工作总结	40
5.2 未来工作展望	41
参考文献	42
致谢	47
攻读硕士期间发表的学术论文及参与的科研项目	49

第 1 章 绪论

1.1 选题背景和研究意义

1.1.1 选题背景

随着科学技术的快速发展和人均国内生产总值（Gross Domestic Product, GDP）的提高，城市道路交通系统所面临的形势愈加严峻。城市经济飞速发展、工业化城市化的加速推进、私家车的普及、道路施工、公交停泊时占道、交通规划的滞后等导致了严重的交通拥堵，交通事故的频发，造成了巨大的经济损失，使得居民的日常出行和身体健康也受到了严重的影响。交通问题是伴随着科学技术的进步和人类社会的向前发展所必须要亟待解决的一个重大问题。由此，智能交通系统(Intelligent Traffic System, ITS)火速发展起来。

智能交通系统的迅速发展得益于现代信息技术的不断革新为其提供了重要的技术支撑。智能交通系统的相关研究受到学者的广泛关注，尤其是在交通领域的应用推动了技术的改进、提高了道路通行效率、提高了城市居民的出行质量、保障了车辆的安全。在面对资源有限、人口众多、车辆激增的现状下，智能交通管理调度系统的建立和完善成为减少交通拥堵的必要手段。实时、准确的交通流量预测成为建立智能调度系统的核心，交通流量预测不仅可以帮助有关部门识别城市交通状态，而且可为解决交通拥堵问题和居民出行提供支持 with 指导^{[1][2]}。

近年来，随着交通行业的快速发展，许多信息采集设备在道路上得到了广泛的部署，如线圈检测仪、行驶记录仪、超声波检测仪等。这些检测设备产生各种形式的海量交通数据，这些数据能够反映交通基本状况。基于海量数据的交通流量预测是交通管理部门采取相应措施的重要依据^[3]，同样也是交通领域的重要研究课题之一。

交通流量预测根据预测时间的长短分为长时预测、中时预测和短时预测。通常情况下，长时预测一般在一天至一年之间。中时预测一般为 30 分钟以上，但是在数小时以内。短时预测一般不超过 15 分钟。交通流预测问题作为 ITS 的重要组成部分之一，其研究和相应的应用也受到业界的广泛关注与讨论^{[4][5][6]}。交通流量数据本身也是时空数据^[7-11]，在时间方向上是动态变化的，即某一条道路上的过去时刻的流量值影响自身道路将来时刻的流量值，该道路某一位置过去的历史流量值将影响其有空间位置相关性的道路的未来流量值。同时交通流量数据又具

有极其复杂的空间依赖关系，处在不同位置的车流量彼此之间互相影响。其次，交通流数据的非线性、多尺度时间相关性、随机性等特征，交通流预测的现有研究虽然取得了初步的成效，但仍然未达到让人满意的成果，因此，交通流量预测研究成为学者们的研究热点和难点，仍然是一项巨大的挑战。

1.1.2 研究意义

交通流预测作为当今智慧城市研究中最关键的问题之一，在交通领域受到了极大的关注。交通预测是有效缓解交通拥堵和解决其相关问题的关键，特别是对于人口快速增长的一线城市来说，准确的流量预测(如交通流量、人流)变得尤为重要。

本文通过构建基于两种不同类型的混合深度学习模型的交通流量预测，帮助交通管理部门实时地掌控交通实时状态。ITS 中呈爆炸式增长的数据对学者的研究和硬件设备带来了极大的考验，借助于当代信息技术和分析方法，能够为其研究提供有效的、可靠的数据完整性支撑。同时也为疏导和控制交通给予根本的判断依据。准确、实时的交通流预测是 ITS 的重要组成部分，也是实现城市道路交通规划与控制的基础，对提高道路通行效率和智能交通的长足发展等具有重要的实际意义。

1.2 国内外研究现状

通过阅读大量的文献和近年来既有的研究成果，总结出交通流预测方法大致归纳为四类：基于时间序列的预测方法、基于传统机器学习的预测方法、基于深度学习的预测方法、基于强化学习的预测方法。

1.2.1 基于时间序列的预测方法

早期基于时间序列的交通流预测方法主要包括历史平均(History Average, HA)模型^[12]、自回归移动平均(Autoregressive Moving Average, ARIMA)模型、季节性 ARIMA (seasonal ARIMA model)模型^[13]、及其它变体模型^[14-15]。Zhang 等^[16]提出了一种新的多步交通流预测的混合模型，该模型将交通流的输入部分分为周期组件、确定组件、波动性组件三部分，用频谱分析技术、ARIMA 模型、带条件方差公式的广义自回归条件异方差模型分别来捕获周期性规律、建模确定性组件、估计波动性。用德克萨斯州休斯顿的交通流数据验证所提模型的预测性能，结果表明，该模型在多步预测方面具有明显的优势。李炜聪等^[17]针对公交客流季节性

波动和周期性变化的特征，构建了一种基于季节性差分整合移动平均自回归模型交通客流量预测模型，并以青岛市的数据集进行验证，结果表明该模型预测精度高。杨紫煜等^[18]在传统的卡尔曼滤波基础之上引入混沌理论，构建短时交通流预测模型，研究表明，该模型预测周期越长，精准度越高，且小样本量的预测精准度不如大样本的预测精度。基于时间序列的方法其模型复杂度低，交通流量的非线性特征属性也能很好地刻画，并且算法和模型易于部署，简单明了，但前提是时间序列数据平稳，且仅考虑交通流的时间关系，它们无法测量复杂的空间依赖关系，因而适用范围比较有限。

1.2.2 基于传统机器学习的预测方法

传统的机器学习方法包括：支持向量回归(support vector regression, SVR)^[19]、相关向量机(Relevance Vector Machine, RVM)^[20]、决策树(Decision Tree, DT)、随机森林(Random Forest, RF)^[21]、聚类(Clustering)、最大期望算法(expectation-maximization algorithm, EM)、贝叶斯方法(Bayesian Analysis, BA)^[22]等和其他混合流量预测模型。此外还有卡尔曼滤波(Kalman Filtering, KF)^[23]、逻辑回归模型(Logistic Regression, LR)^[24]等。Zhu 等^[25]和 Ryu 等^[26]分别用线性条件下的贝叶斯网络模型和 K-近邻模型来捕获时空属性特征进行短期交通流预测。Tang 等^[27]等从数据源头出发，研究去噪方案，并与支持向量机算法相结合，进行交通流多步预测，利用明尼阿波利斯市高速公路上的 3 个环路检测器采集的交通量数据验证了模型预测的有效性，进一步表明了去噪算法的结果优于未去噪的算法。为了删除输入数据之间的冗余信息，Hinton 等^[28]提出一种利用神经网络对多维数据进行降维的方法。Li 等^[29]建立一种基于季节性 ARIMA 和 SVR 算法的预测模型，以北京市的数据集验证该模型，结果表明，该模型可以很好地反应客流的变化规律，能够提高预测精度。大量的实验研究发现传统的机器学习方法对于短期交通流预测很友好，有较好的预测精度。但由于传统的机器学习对输入数据有很强的依赖性，需要大量的专家经验，并且这些方法没有考虑路网的空间影响，故对于复杂的交通流预测仍然是一个巨大的挑战。

1.2.3 基于深度学习的预测方法

近些年以来，由于计算机技术以及 GPU 云服务器(GPU Cloud Computing, GPU)强大的计算能力使得深度学习在诸多领域(如图像、自然语言、目标检测等)取得明显优异的成果。深度学习就是通过监督学习或者无监督学习，对所输入的

信息和海量数据无需人工干预,自动学习复杂数据中所隐含的模式,并将隐藏的数据特征提取出来。近年来深度学习在各个领域取得了突破性的进展也激发了学者将其应用于交通领域的研究中。最初,卷积神经网络(Convolutional Neural Networks, CNN)和循环神经网络(Recurrent Neural Network, RNNs)为两大主要模型,杜圣东等^[30]考虑到与交通流相关的多通道和多变量序列中存在潜在的隐含时空依赖关系,提出了一种新的时空注意交通预测(Spatial-Temporal Attention Traffic Forecasting, STATF)模型,用3个交通流量数据集验证了该模型在挖掘数据中隐含的时空依赖性和非线性相关特征方面具有很大的优势。卢生巧等^[31]用CNN提取空间相关性,用门控循环单元(Gate Recurrent Unit, GRU)神经网络提取时间维度上的特征,提出了一种短时交通流预测组合模型。但是单向的循环神经网络忽略了后续信息对当前时刻的反馈影响,特别地,对于长时间序列,会丢失重要信息。孙加新等^[32]提出了一种由CNN网络、双向长短期记忆网络(Bidirectional Long Short Term Memory Network, BiLSTM)、注意力机制(Attention Mechanism, AM)组成的融合多特征的组合模型。与以往研究不同的是,该模型充分地挖掘原始数据中所包含的时间、天气、温度等关键信息,使得模型具有更好的特征感知能力。通过实验对比表明,该模型预测结果的准确度以及稳定性较高。王博文等^[33]运用编码器-解码器(Encoder-Decoder, ED)框架解决时间序列过长所引起的误差积累问题,构建ED-LSTM模型进行交通流的多步预测,在PEMS-04、Xi-an两个数据集上进行验证,表明了该模型对单因素和多因素的多步交通流时间序列是适用的。为了更全面地考虑,Zou等^[34]提出了可以同时处理多个时间维度的一种新型的RNN结构单元来预测特征演变,并与CNN结合的联合模型来实现交通预测,实验表明,该模型能显著提高预测精度。Fouladgar等^[35]为避免不平衡训练数据集的影响引入了一个正则化的欧几里得损失函数,提出了一种基于去中心化深度学习的可扩展的分散交通流预测方法。在数据集上进行的大量实验,结果表明,该方法成功地预测了拥塞。Du等^[36]针对城市客流量的多通道以及不规则的特性,构建一种新的深度不规则残差LSTM模型。在不同类型的数据集上进行了广泛的实验,结果表明在城市客流量预测方面,该模型比传统方法和基于深度学习的方法具有一定的优势。此外,还有门控循环单元(Gated Recurrent Unit, GRU)^[37]、时域卷积网络(Temporal Convolutional Network, TCN)^[38]等深度学习模型,也广泛应用于交通领域。

图卷积网络(graph convolutional networks, GCN)^[39]凭借其强大的时空交通数据处理能力,取得了长足的发展,已成为交通领域最活跃和研究的热点。在[40][41][42]中对GCNs及其各种变体进行了详细的介绍。Yu等^[43]开发了一种新的

时空图卷积 (Spatio-Temporal Graph Convolutional Networks, STGCN) 网络的交通流量预测模型, 实验证明该模型训练时间短、收敛快、参数少。Wu 等^[44]采用图划分的方法, 利用扩散卷积递归神经网络对交通进行预测。Qiu 等^[45]提出了一个可以准确预测流量和密度的全新系统, 引入一种拓扑图卷积网络 (Topological Graph Convolutional Network, ToGCN) 和 Seq2Seq 的模型框架对交通流和密度做出预测, 用出租车轨迹大数据验证了该模型的有效性。Guo 等^[46]首先, 将时间维度划分为三个不同的时间粒度 (临近序列、日周期序列、周周期序列) 的交通流量。然后, 分别将三个不同的时间序列输入到基于注意力的时空图卷积网络 (Attention based Spatial-temporal graph convolutional, ASTGCN) 模型中。最后, 将三个不同时间序列的输出结果进行加权融合, 得到的融合结果为该模型的最终预测结果。在公开数据集上验证了该模型的优势。Song 等^[47]为了捕获交通数据的时空相关性和异质性, 提出一种时空同步图卷积网络 (Spatial-Temporal Synchronous Graph Convolutional, STSGCN) 模型。结果表明了该模型优于基线模型, 但是没有考虑时空局部图之间对于预测的权重。Wu 等^[48]在图卷积的基础上提出了动态自适应的邻接矩阵用来提取隐藏的图结构关系, 由时间卷积层、图卷积层和残差连接组成的 Graph Wave Net 模型, 但该模型忽视了邻接节点之间的权重分配。深度学习具备强大的学习能力、覆盖面广、适应性强、可移植性好等优点。大量实验也表明, 深度学习的预测性能好, 但是对硬件设备要求高、计算复杂、超参数调整耗时耗精力, 因此, 隐藏层神经元个数的确定和网络层数的确定, 是接下来要重点关注和研究的方向。

1.2.4 基于强化学习的预测方法

Qu 等^[49]为了降低交通流量的非平稳性特征和充分捕获交通流的特征, 构建了一种强化学习最优组合的交通流量预测模型。首先, 将流量分解为包含不同特征的多个分量, 对其中波动频率较大的分量采用变分模态分解 (Variational Mode Decomposition, VMD)。然后, 对分量进行重构和对流量特征进行筛选。最后, 根据各个模型的优势用强化学习来确定其权重以得到满意的结果。实验表明, 模型在捕获复杂交通流的特征方面具有优势。目前, 由于强化学习不需要环境的先验知识, 因此, 强化学习在路径规划算法中得到了很好的应用。此外, 在城市交通应用方面, 强化学习主要解决交通控制问题, 如交通系统中客流量控制策略^[50]。

1.3 论文的研究内容与主要工作

以上内容仅为本文档的试下载部分，为可阅读页数的一半内容。如要下载或阅读全文，请访问：<https://d.book118.com/657021114056006142>