

用于异常交通流预测的时空生成对抗聚类图卷积网络

张红, 杨俊译

(兰州理工大学计算机与通信学院, 甘肃 兰州 730050)

摘要: 准确的交通流预测对城市交通至关重要, 而交通流中的异常事件(如恶劣天气、紧急道路状况等)导致的异常交通流, 对现阶段交通流模型预测的准确性提出严峻挑战。现有的交通流预测模型大多难以有效捕捉交通流数据的复杂动态模式, 且计算资源开销较大。为了解决这些问题, 本文提出一种用于异常交通流预测的时空生成对抗聚类图卷积网络(CG-STGAN)。构建一个基于时空生成对抗网络的模型, 其中, 生成器用于捕捉并模拟正常交通流数据, 鉴别器则对比生成数据与原始数据, 判断其是否异常。通过对抗训练, 生成器和鉴别器协同提升了异常交通流的检测能力。此外, 提出聚类图卷积网络, 将交通图分解为多个子图, 以限制图卷积的邻域扩张, 从而降低内存使用并提高预测准确性。最后, 结合图卷积门控循环单元挖掘交通流的短期时空特征, 并通过长短期记忆(LSTM)模块学习交通流的长期依赖。实验结果表明, 与基线模型相比, CG-STGAN 在异常交通流预测方面表现出最佳性能。

关键词: 交通流预测; 异常交通流; 时空动态特性; 生成对抗网络; 聚类图卷积网络

中图分类号: TP181

文献标志码: A

DOI: 10.19678/j.issn.1000-3428.0070180

Spatio-Temporal Generative Adversarial Clustered Graph Convolutional Network for Anomalous Traffic Flow Prediction

ZHANG Hong, YANG Junyi

(School of Computer and Communication, Lanzhou University of Technology, Lanzhou 730050, Gansu, China)

【Abstract】 Accurate traffic flow prediction is critical for urban transportation; however, abnormal events in traffic flow, such as severe weather or emergency road conditions, present significant challenges in maintaining the accuracy of current traffic flow prediction models. Most existing models fail to accurately capture complex dynamic patterns of traffic data and often require substantial computational resources. To address these issues, this paper proposes a Spatio-Temporal Generative Adversarial Clustering Graph Convolutional Network (CG-STGAN) for abnormal traffic flow prediction. The model uses a generator to capture and simulate normal traffic flow data, while a discriminator compares the generated data with the original data to detect abnormalities. Adversarial training helps improve the ability of the model to identify abnormal traffic flows. In addition, a Cluster Graph Convolutional Network (Cluster-GCN) is introduced to divide the traffic graph into multiple subgraphs, thereby limiting graph convolution neighborhood expansion, reducing memory usage, and improving prediction accuracy. A graph convolutional gated recurrent unit is employed to extract short-term spatiotemporal features, while a Long Short-Term Memory (LSTM) module learns long-term dependencies. The results of experiments demonstrate that CG-STGAN outperforms baseline models in abnormal traffic flow prediction.

【Key words】 traffic flow prediction; abnormal traffic flow; spatio-temporal dynamic features; Generative Adversarial Network (GAN); Cluster Graph Convolutional Network (Cluster-GCN)

0 引言

随着机动车数量的逐年增加, 交通拥堵和废气排放问题日趋严重^[1]。智能交通系统(ITS)可以实现交通控制和交通引导, 能够有效缓解交通拥堵, 提高道路的通行能力。然而, 由于突发的异常交通流情况等外部因素影响, 设计一种有效的交通流预测方法仍然面临诸多挑战。短期交通流量预测是基于

历史数据对未来一段时间内的交通流量状态进行预测。早期方法主要采用传统的统计模型进行交通流预测, 例如卡尔曼滤波^[2]和指数平滑^[3], 但这些模型在处理复杂数据分布和高维数据时效果不佳。研究人员提出了机器学习模型, 如支持向量回归(SVR)^[4]和 k 最近邻(k -NN)^[5]。然而, 这些方法存在对标记数据依赖较高、容易过拟合、在新数据上泛化能力差等问题。

基金项目: 甘肃省科技重大专项计划(25ZYJA037); 甘肃省重点人才项目(2024RCXM57)。

作者简介: 张红(CCF 会员), 女, 副教授、博士, 主研方向为机器学习、智能交通; 杨俊译, 硕士研究生。

收稿日期: 2024-07-26

修回日期: 2024-11-19

E-mail: zhanghong@lut.edu.cn

随着深度学习技术的发展,卷积神经网络(CNN)逐渐应用于交通流预测。然而,交通流数据通常具有复杂的非欧氏空间结构,传统的 CNN 在捕捉这些复杂关系时表现出一定局限性。为此,图卷积网络(GCN)被提出,其通过节点信息和连接关系进行特征变换,广泛应用于图数据的空间特征提取。随后,递归神经网络,如长短期记忆网络(LSTM)^[6]和门控循环单元(GRU)^[7],被引入交通流预测领域。这些方法有效捕捉了交通流数据中的时间依赖性,但在空间依赖性建模方面仍显不足。为此,基于图神经网络(GNN)的模型,如STGODE^[8],通过道路邻接关系和时间序列相似性构建邻接矩阵,试图同时处理时间和空间依赖性。然而,它们在动态空间依赖建模上仍存在一定局限性。BAO等^[9]提出结合先验知识的时变图卷积网络(PKET-GCN),将交通流数据划分为动态特征和静态特征,分别对其进行识别和特征提取,以提升道路交通流预测的准确性。HUANG等^[10]提出多视图动态图卷积网络(MVDGCN),捕获不同级别下的时空依赖关系。GSTPRN^[11]构建了位置图卷积,增强了空间位置和邻域信息。HGCN^[12]使用多个带有光谱池的 GCN,以分层方式进行时空学习。尽管上述方法讨论了交通流数据的时空特征,但普遍忽略了异常交通流对交通流预测准确性的影响,导致模型的预测性能不高。

交通流数据不仅和交通流量、速度、时间占用率和密度密切相关,还受突发恶劣天气和紧急道路状况等异常交通流情况的影响^[13]。结合深度学习的异常交通流预测模型能够利用自动特征提取能力更好地捕捉数据中的异常情况,有效提高模型预测的准确性和鲁棒性。LI等^[14]提出基于图时空模式学习网络,有效学习了道路的时空特征,并设计异常状态评估方法来预测异常,但该方法对外部因素考虑较少。CHEN等^[15]提出基于注意力的循环神经网络架构,其将异常交通流纳入考虑,但难以捕捉长时间依赖。JANG等^[16]提出考虑外部因素数据的 LSTM+GRU 组合模型,将外部因素与交通流数据结合,提高预测准确性,但模型参数数量多,训练和计算成本较高。尽管这些方法在时间序列异常预测方面取得了一定成果,但它们通常过于依赖现有的标记数据。当异常样本较少时,模型的预测准确性会受到明显影响。

针对上述问题,有研究者提出利用生成对抗网络(GAN)技术进行动态异常检测。GAN 通过生成正常交通流数据,并与原始数据进行对比,能够有效

检测异常交通流事件,从而提升模型在复杂动态交通场景下的异常检测能力^[17]。WU等^[18]将 GAN 与 LSTM 结合,提出基于顺序生成对抗网络的场景生成模型(LSTM-GAN),该模型能够有效捕捉交通流中存在的时间动态特征,但忽略了交通流数据的空间特性。BASHAR等^[19]提出基于 GAN 的时间序列异常预测方法(TAno-GAN),通过学习数据分布和生成高质量样本提升预测性能。ALEXANDER等^[20]提出基于 GAN 的事件序列异常预测模型(TadGAN),将 LSTM 作为基础模型与 GAN 结合,实现时间序列数据重建和异常判定。XU等^[21]提出基于 GAN 的时间序列数据异常预测算法,利用 GAN 生成特性,将生成的正常数据与原始数据对比以检测异常,并在模型中加入 Transformer 增强提取上下文能力。DENG等^[22]提出时空图卷积对抗神经网络模型(STGAN),通过捕捉道路交通时空特征,用对抗神经网络提升预测精确性,在异常交通流预测方面取得显著成效。然而,由于 GAN 本身的数据生成特性及交通流数据的复杂性,预测时计算资源开销较大,限制了其应用和模型准确性。

针对以上问题,本文提出了一种用于异常交通流预测的时空生成对抗网络的聚类图卷积(Cluster-GCN)模型(CG-STGAN)。CG-STGAN 基于 GAN,结合 Cluster-GCN、图卷积门控循环单元(GCGRU)和 LSTM。Cluster-GCN 能够将交通网络划分为多个子图,对不同的子图再使用图卷积操作去提取局部特征,同时在局部特征提取的基础上提高模型全局信息特征提取能力,从而提升模型捕捉异常交通流变化的敏感度和准确性。通过将外部因素数据与长短期数据融合后输入 Cluster-GCN 进行卷积,进一步优化了模型的内存占用率和计算效率。特别地,GCGRU 用于捕捉交通流的短期时空特征,LSTM 用于捕捉交通流的长期趋势特征。最终,模型显著降低了内存占用率,提高了异常交通流预测的准确性。

本文主要贡献如下:

1) 本文提出一个异常交通流预测模型 CG-STGAN,它能够很好地捕获节点之间的隐藏状态,充分发掘交通路网中的异常车流量信息,大幅提升模型异常交通流预测的准确性和泛化能力。

2) 本文构建了一个 Cluster-GCN 模块,该模块通过最大限度避免 GCN 节点之间邻域扩张,降低模型运行时的内存占用率,并进一步提升模型的交通流预测准确率。

3) 提出 GCGRU 模块以捕获交通流的长短期依赖特征, 将 GCGRU 模块和 Cluster-GCN 模块结合, 在降低模型内存占用率的基础上, 最大程度地提取异常交通流时空特征。

1 模型构建

本文中, CG-STGAN 模型的主要框架由鉴别器和生成器组成。在此基础上, 生成器由短期时间相关性、长期时间相关性、外部相关性模块以及 Cluster-GCN 构成。特别地, 短期时间相关性模块

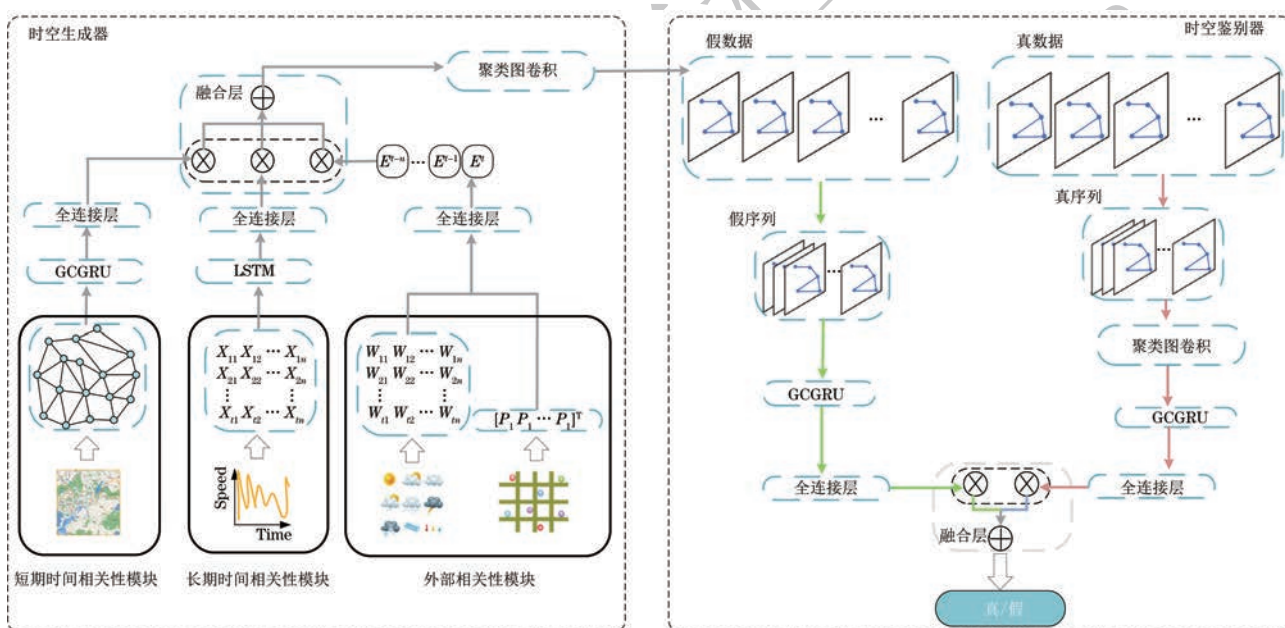


图 1 CG-STGAN 模型总体框架

Fig.1 Overall framework of CG-STGAN model

特别地, 本文将路网定义为图 G 。道路网公式表示为 $G=(V, E, A)$, 其中, V 代表了道路各个路段的集合 $V=\{v_1, v_2, \dots, v_n\}$, 即道路节点的集合; E 代表边的集合和路段之间的连通性; A 代表加权邻接矩阵。通过使用带高斯核的阈值方法来建立邻接矩阵。此外, 交通流预测问题旨在学习一个函数 $N(\cdot)$, 对于每个时间步长 t , 有指定的道路 V_n 的交通流为 X_n^t , 具体的函数映射关系如下:

$$[X_{T+1}^{S_1}, \dots, X_{T+H}^{S_T}] = N([X^{S_1}, \dots, X^{S_T}; G]) \quad (1)$$

将 $X=\{X^{S_1}, \dots, X^{S_T}\} \in \mathbb{R}^{T \times N}$ 作为函数的输入, 以此预测未来 H 个时间步长的交通流情况。

1.1 时空生成对抗网络

GAN 通过生成器和判别器的互相博弈产生良好输出效果。令 x 代表现实世界中真实的数据, z 代表输入 G 网络的噪声, $G(z)$ 表示生成的“假数据”, $D(\cdot)$ 表示判断生成数据是否真实的概率,

由 GCGRU 构成, 通过结合 GCN 和 GRU 模块以此捕捉异常交通流短期时空特征; 长期时间相关性模块通过引入 LSTM 模块来提取长时间的时间特征; 外部相关性模块包含了交通流数据集中的天气以及其他道路信息数据; 最后, 提出并使用 Cluster-GCN 将交通网络划分为多个子图并在这些子图上进行卷积操作, 同时避免图卷积过程中邻域扩展的问题。CG-STGAN 模型总体框架如图 1 所示(彩色效果见《计算机工程》官网 HTML 版, 下同)。

GAN 的公式如下:

$$\begin{aligned} \min_G \max_D V(D, G) = \\ \min_G \max_D E_{x \sim p_{\text{data}}(x)} [\ln D(x)] + \\ E_{z \sim p_z(z)} [\ln(1 - D(G(z)))] \end{aligned} \quad (2)$$

在该网络中, 生成器负责生成“假数据”, 而判别器则负责对生成的数据进行判别, 它的作用是判别数据是否是真实的。通过生成器与判别器的对抗博弈学习, 使得两者都得到最佳的模型, 以此输出最优结果。

1.1.1 时空生成器

在本节中, 主要介绍 CG-STGAN 生成器的主要组成。时空生成器具体由 3 个独立部分构成: 近期依赖模块、趋势依赖模块以及外部依赖模块。经过融合层后, 通过 Cluster-GCN 对上述 3 个部分进行合并输出与预测。

1) 短期时间相关性模块。

数据集中两个相邻的节点往往具有相同的交通

动态,具有高度的相关性,而距离较远的节点往往具有不同的交通动态。本文将 GRU 中的门控矩阵乘法替换为 GCN,得到 GCGRU 模块,通过 GRU 模块捕捉节点之间的时间相关性,再利用 GCN 模块捕捉节点之间的空间相关性。特别地,本文用 GCN 来更新 GRU 中的更新门、重置门和隐藏状态,该模块使得模型能够在初期获得更加全面的道路交通时空数据,具体定义如下:

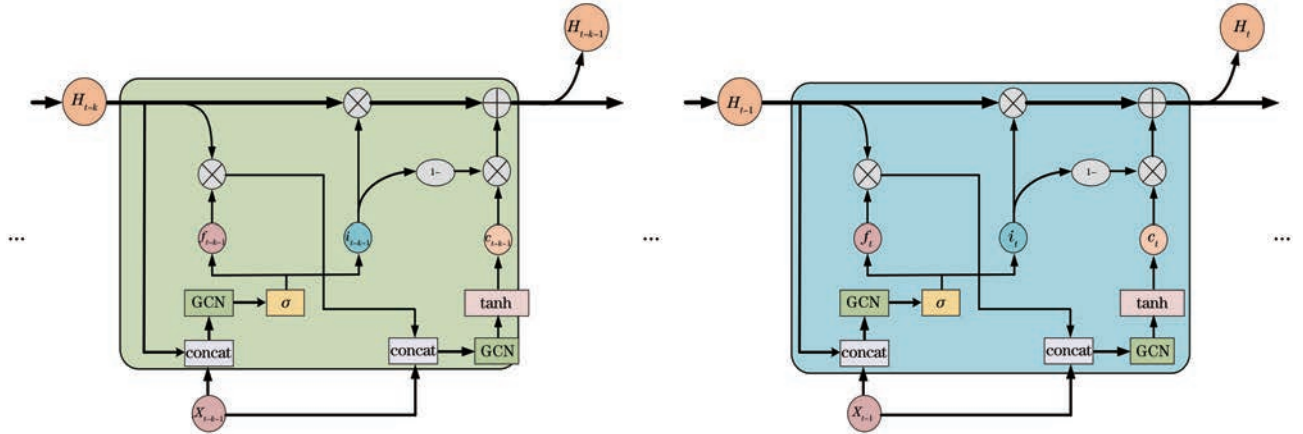


图 2 GCGRU 模块
Fig.2 GCGRU module

2) 长期时间相关性模块。

交通流的时空特征不仅与短期交通动态相关,还与其长期时间特征紧密相连。而其长期特征往往受到很多因素的影响,例如突发事件以及道路封锁,城市事件对交通数据有着重要影响。LSTM 可学习交通流的长期时间特征,因此模型采用 LSTM 来捕捉每个节点流量数据的趋势,本文通过 T_h 来表示趋势段的输入长度,假设当前时间为 t ,具体可表述为:

$$X_h = (x_{v,t-T_h}, x_{v,t-T_h+1}, \dots, x_{v,t-1}) \in \mathbb{R}^{T_h \times F} \quad (7)$$

式中: X_v 为 t 时刻 v 节点的特征矩阵。LSTM 模块具有捕获时间长期依赖的作用,主要用于将时变产生的表示形式输入对应的 LSTM 中,建模其时间依赖关系,并且以此推导出隐藏流量状态。

3) 外部相关性模块。

考虑到外部因素对交通数据预测的重要影响,本文将这些因素通过全连接层输入,构建外部相关性模块。而在现实的交通中,道路交通预测的准确性通常受到许多外部因素的共同影响。例如时间、天气。首先,交通流量在不同时间区间内具有不同的特点,比如交通流量在晚上会急剧下降,在清晨又会突然激增。其次,天气因素也影响道路交通流,例如,当天气为雨等恶劣天气时,道

$$f_t = \sigma(W_{f*G}([X_t, H_{t-1}] + b_f)) \quad (3)$$

$$i_t = \sigma(W_{i*G}([X_t, H_{t-1}] + b_i)) \quad (4)$$

$$\tilde{C}_t = \tanh(W_{c*G} \cdot (r \odot [X_t, H_{t-1}] + b_c)) \quad (5)$$

$$C_t = C_{t-1} \odot f_t + \tilde{C}_t \odot (1 - z_t) \quad (6)$$

式中: f_t 代表 t 时刻的更新门; i_t 代表 t 时刻的复位门; $*G$ 代表图卷积; X_t 代表 t 时刻的输入; H_{t-1} 代表 $t-1$ 时刻的输出; W_f, W_i, W_c 代表权重值; b_f, b_i, b_c 代表偏置项。GCGRU 如图 2 所示。

路车流量下降明显,与良好天气下的道路交通流形成鲜明对比。

本文模型将外部因素分解为道路时间因素和天气因素,其中道路时间因素记为 T ,天气因素记为 W ,其中有外部属性矩阵 $O = \{W, T\}$,将属性矩阵输入全连接层中,得到属性增强单元。其中,道路时间属性可细分为: $T_{\text{weekday}}, T_{\text{hour}} \in T$ 。与道路时间属性相似,天气属性同样是一组互斥的属性类别,每个样本仅归属于其中一个类别。因此,本文采用 one-hot 编码表示道路的外部相关性,以充分反映数据的稀疏性、缺乏距离信息和互斥性等特征。在实际应用中,具体的属性增强矩阵如下所示:

$$O = [T_{\text{weekday}}; T_{\text{hours}}; W] \quad (8)$$

4) 损失函数。

现实生活中的交通流不仅受天气影响,其他因素如时间同样影响道路交通预测的准确性。为了提高生成器在异常情况下生成真实交通流数据的能力,本文采用一个损失函数来缩小生成数据与真实数据之间的差距,从而使生成器能够生成极为真实的样本以欺骗鉴别器。因此,针对数据集中的一个数据点,具体的损失函数公式如下:

$$L_G(\sigma) = \sum_s -\ln(1 - D_\Phi(\hat{R}_{v,t})) + \epsilon_G \|G_\theta(v, t) - Q_{v,t}\|_2 \quad (9)$$

式中: G_θ 和 D_ϕ 分别代表了 GAN 的生成器函数和鉴别器函数; ϵ_G 是一个用于平衡鉴别器损失和预测误差的超参数; $\hat{R}_{v,t}$ 表示生成器生成的虚假序列值; $Q_{v,t}$ 表示 t 时刻 v 节点真实交通流数据。实验最后, 当生成器生成的实例越接近真实数据, ϵ_G 和 $\hat{R}_{v,t}$ 的值将会越小。

1.1.2 时空鉴别器

本节提出了一种灵活的异常评分判别机制。首先, 利用 GCGRU 层对当前时空特征进行处理, 以捕捉短期时间相关性。接着, 通过 Cluster-GCN 进行图卷积操作, 以提取交通流图的特征。最后, 经过全连接层和融合层, 模型对异常情况进行分类预测并输出结果。

1) 对抗损失。

鉴别器与生成器类似, 仍通过 GCGRU 模块来捕捉最近一段时间的时空特征, 并通过使用 Cluster-GCN 来捕获当前时间段的空间特征, 最后使用 Sigmoid 函数作为全连接层的激活函数, 以此判断当前输入序列的真假。基于此, 为了区分真假序列, 本文提出了一种能够区分真假序列异常的损失值, 具体公式如下:

$$L_D(\Phi) = \sum_s -\ln(D_\phi(\hat{S}_{v,t})) - \ln(1 - D_\phi(S_{v,t})) \quad (10)$$

式中: 下标 D 代表鉴别器; $S_{v,t}$ 和 $\hat{S}_{v,t}$ 分别代表了真、假序列值。在本文模型中, 鉴别器和生成器相互对抗, 其中鉴别器利用生成器生成的异常交通流的数据进行学习, 随着生成器生成的数据变得越来越真实, 鉴别器的鉴别能力也得到了大幅提升。

2) 异常分数。

在提出的模型中, 生成器和鉴别器均可独立用作异常交通流数据的检测器, 因此定义了一个异常分数值以比较异常交通流的检测效果。对于生成器而言, 异常分数可视为预测误差, 用于评估异常情况下的道路交通流变化; 对于鉴别器, 其输出直接反映异常得分, 用于检测偏离时空特征的异常情况, 具体公式如下:

$$s_D(v, t) = D_\phi(S_{v,t}) - D_\phi(\hat{S}_{v,t}) \quad (11)$$

$$s_G(v, t) = \|G_\theta(v, t) - X_{v,t}\|_2 \quad (12)$$

$$s_{\text{score}}(v, t) = s_G(v, t) + \theta s_D(v, t) \quad (13)$$

式中: 下标 D 和 G 分别代表 t 时刻 v 节点的鉴别器和生成器; θ 为平衡生成器和鉴别器的超参数; $s_D(v, t)$ 和 $s_G(v, t)$ 分别代表鉴别器对异常情况下交通流的判断得分和生成器生成数据与真实数据之间的差异得分。最终, 将两个异常得分相加得到综

合异常分数值 $s_{\text{score}}(v, t)$, 通过最小化异常分数, 可以有效判断异常情况下的交通流状况。

1.2 聚类图卷积

在 GCN 中, 通过使用小批量随机梯度下降 (SGD), 即每次更新是建立在一个 Mini-Batch 的基础上的, 使得其能对每个 Epoch 进行多次更新, 从而使得模型更快收敛。因此该方法可以有效地减少模型整体对于内存的需求。小批量损失函数为:

$$L = \frac{1}{Q} \sum_{i \in Q} \text{loss}(y_i, M_i^L) \quad (14)$$

式中: $Q \in |N|$ 代表了一个批量内节点的索引; y_i 代表数据的真实值; M_i^L 代表第 i 个节点的预测结果。但是传统的 SGD 往往面临着邻域扩展等问题, 进而导致 GCN 的复杂程度呈指数级增长, 造成了极大的计算量以及内存占用率。为了解决上述问题, CHIANG 等^[23] 提出了 Cluster-GCN。Cluster-GCN 是在 SGD 小批量基础上进行修改的, 其通过设计一个批量和计算子图来最大化嵌入利用率, 将嵌入利用率与聚类目标联系起来。首先, 将 SGD 更新效率与图聚类算法联系起来。对于图 G , 将其节点划分为 n 组子图, 分别为 $\tilde{G} = [G_1, G_2, \dots, G_n]$ 。在子图内, 对于每个节点 V , 有 $V = [V_1, V_2, \dots, V_n]$, 特别地, 本文的每个子图公式表示为:

$$\tilde{G} = [G_1, G_2, \dots, G_n] = [\{V_1, E_1\}, \{V_2, E_2\}, \dots, \{V_l, E_n\}] \quad (15)$$

式中: E_n 代表了每个节点对应的边。重建后的邻接矩阵可以表示为:

$$A = \bar{A} + \Delta = \begin{bmatrix} A_{11} & \dots & A_{1n} \\ \vdots & & \vdots \\ A_{n1} & \dots & A_{nn} \end{bmatrix} \quad (16)$$

式中, 有:

$$\bar{A} = \begin{bmatrix} A_{11} & \dots & 0 \\ \vdots & & \vdots \\ 0 & \dots & A_{nn} \end{bmatrix}, \Delta = \begin{bmatrix} 0 & \dots & A_{1n} \\ \vdots & & \vdots \\ A_{n1} & \dots & 0 \end{bmatrix} \quad (17)$$

A_{ii} 是一个维数为 $|V_i| \times |V_i|$ 的邻接矩阵, $\tilde{G}$ 的邻接矩阵为 $\bar{A}$, 所有非对角块组成的矩阵表示为 Δ 。节点 V_n 的特征矩阵 X 和训练标签 Y 也可进一步划分为 $[X_1, X_2, \dots, X_n]$ 和 $[Y_1, Y_2, \dots, Y_n]$ 。

将 GCN 的函数划分为不同的批次, 设 $\bar{A}'$ 为 $\bar{A}$ 的归一化版本, 得到嵌入矩阵如下:

$$\begin{aligned} \bar{A} = & \bar{A}' \sigma(\bar{A}' \sigma(\cdots \sigma(\bar{A}' X W^{(0)}) W^{(1)}) \cdots) W^{(k-1)} = \\ & \begin{bmatrix} \bar{A}'_{11} \sigma(\bar{A}'_{11} \sigma(\cdots \sigma(\bar{A}'_{11} X_1 W^{(0)}) W^{(1)}) \cdots) W^{(k-1)} \\ \vdots \\ \bar{A}'_{mm} \sigma(\bar{A}'_{mm} \sigma(\cdots \sigma(\bar{A}'_{mm} X_n W^{(0)}) W^{(1)}) \cdots) W^{(k-1)} \end{bmatrix} \end{aligned} \quad (18)$$

基于此, Cluster-GCN 的损失函数可以表示为:

$$\begin{aligned} L_{A'} &= \sum_i \frac{|V_i|}{N} L_{A'_{ii}}, \\ L_{A'_{ii}} &= \sum_{m \in V_i} \text{Loss}(y_m, O_m^{(L)}) \end{aligned} \quad (19)$$

式中: $L_{A'}$ 和 $L_{A'_{ii}}$ 为划分过程后的损失函数; $L_{A'_{ii}}$ 根据 V_b 的采样情况对 SGD 更新, 因此聚类图卷积的搜索过程比原图卷积更为高效。具体的 Cluster-GCN 原理如图 3 所示。

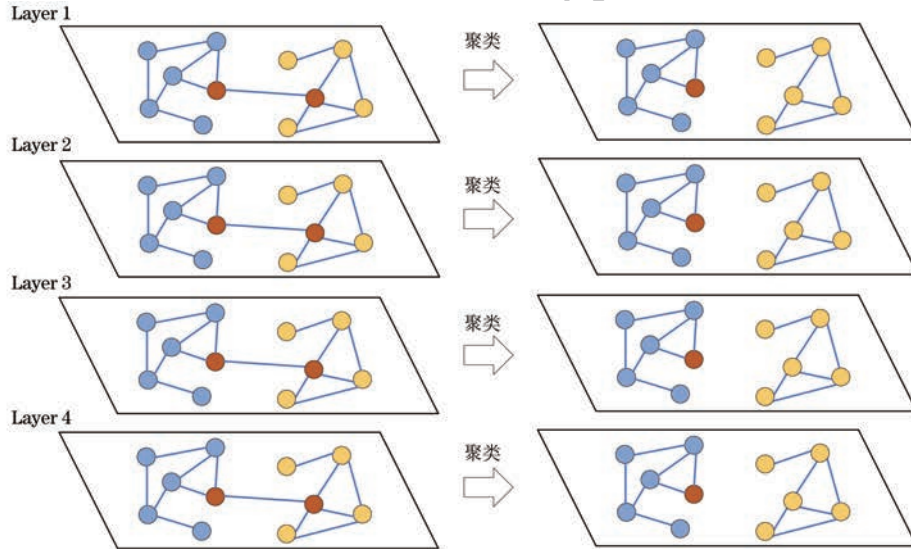


图 3 Cluster-GCN 原理图

Fig.3 Schematic diagram of Cluster-GCN

如图 3 所示, 通过使用 Cluster-GCN 可以让分类器专注于自身领域, 避免传统 GCN 复杂的邻域搜索, 极大地降低了模型的内存占用率, 提高了模型的计算效率。

1.3 融合层

为了将数据源更好地进行融合, 本文提出使用 Fusion 层来对数据进行多模态学习和信息整合。对于每个数据点 d , 有 $d = \langle V, T, X_{v,t} \rangle$ 。具体定义表示如下:

$$\hat{X}_{v,t} = \tanh(\Theta_{F*G} [X_V, X_T, X_O]) \quad (20)$$

式中: $\hat{X}_{v,t}$ 代表使用生成器预测的关于子图 G_v 的交通流预测; $*G$ 表示聚类图卷积; X_R, X_T, X_O 分

别代表短期相关性模块、长期相关性模块以及外部相关性模块的输出。

2 实验

2.1 数据描述

本文使用了美国加利福尼亚州(加州)和纽约市的公开数据集进行实验, 如表 1 所示, 数据集包括了道路交通信息、速度信息, 以及加州公路巡警事故报告、车道关闭系统报告。具体而言, PeMS 数据集包含了 2014 年 4 月—2014 年 5 月的加州公路巡警事故报告和车道关闭系统报告; NYC 数据集则收录了 2014 年 11 月纽约市的交通异常的相关事件记录报告。

表 1 实验数据集

Table 1 Experimental dataset

数据集	节点数	时间间隔/min	特征长度	采样日期
PeMS	862	5	4	2014 年 1 月 15 日—2014 年 11 月 30 日
NYC	365	30	12	2017 年 1 月 1 日—2017 年 5 月 6 日

1) PeMS(bay): 该数据收集了 2014 年加州旧金山湾区高速公路交通事故以及道路关闭报告数据, 该数据集详细记录了事故的开始时间、持续时间以及位置和类型, 因此使用本数据集进行训练会有更好的泛化性, 能够更好地对模型进行

训练。

2) NYC: 该数据集收集了纽约市出租车数据以及自行车租赁数据, 记录了由社会事件引起的道路交通异常情况, 且详细记录了异常事件的描述、发生时间以及其地点等信息。

为了更好地展示异常交通流情况,本文摘取了 NYC 数据集中的一部分进行深入分析。例如,当附

近举办了一场 20:00—22:00 的音乐会时,正常流量和异常流量对比如图 4 所示。

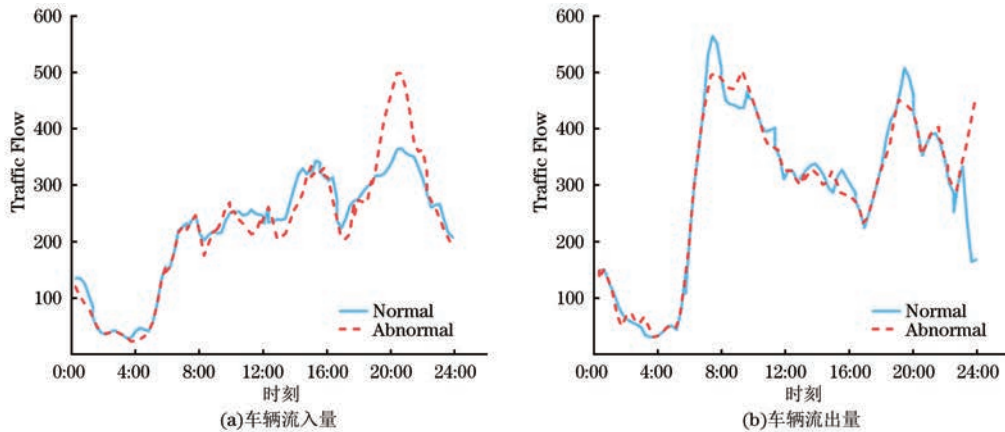


图 4 正常流量和异常流量对比

Fig. 4 Comparison between normal traffic and abnormal traffic

根据图 4 可以明显看出,演出开始和结束之前,整体交通流量保持正常状态。在音乐会开始后,交通流入量出现了显著的异常点。同时,音乐会结束时,车辆流出量也呈现出相应的异常波动。

2.2 实验设置

本文使用的实验环境主要是在 Linux 系统下进行搭建,使用到的主要硬件参数如下:CPU 为 Intel Xeon Silver 4214R Processor 16. 5M Cache 2.40 GHz, GPU 为 NVIDIA GeForce RTX 3090。本次模型的构建以及训练均在系统提供的终端内完成,使用基于 PyTorch 的机器学习框架。其次,将 Batch Size 大小设置为 256,使用学习率为 0.001 的 Adam 优化算法来训练模型。

2.2.1 异常评估指标

为了更加精准地评估异常检测的覆盖率和准确度,采用异常召回率来衡量异常检测的性能。在异常检测中,异常得分越高的将其视为异常,取 $C\%$ 异常分数值得分较高的节点标记为异常,用 Ω_c 表示,而真实数据中存在的异常用 Ω 表示,如果模型检测到的异常和数据集真实异常重叠,则记为一次有效的异常检测。为了充分验证模型异常检测的有效性以及交通流预测的精确性,使用异常召回率对异常交通流情况进行评估:

$$\text{Recall}@C\% = \frac{|\Omega \cap \Omega_c|}{|\Omega|} \quad (21)$$

2.2.2 交通流预测评估指标

为了更好地分析实验结果,评估模型对异常交通流的预测性能,提出以下评价指标来对真实值和预测值之间的误差情况进行评估:

1) 均方根误差 (RMSE, E_{RMSE}):

$$E_{\text{RMSE}} = \left[\frac{1}{n} \sum_{t=1}^n (y_t - \hat{y}_t)^2 \right]^{\frac{1}{2}} \quad (22)$$

2) 平均绝对误差 (MAE, E_{MAE}):

$$E_{\text{MAE}} = \frac{\sum_{t=1}^n |y_t - \hat{y}_t|}{n} \quad (23)$$

式中: n 表示交通路网上的节点数; y_t 和 $\hat{y}_t$ 分别为第 t 个时间步的真实交通信息和模型的预测值。

2.3 基线模型

为了验证本文模型的有效性,将相应的数据集放入以往的模型进行训练,进行比较实验的模型有:

1) 跨时空城市异常检测分解模型 (ST_decompn)^[24]: 一种基于分解的城市异常检测模型,通过学习道路时空特征,先建模正常道路模式,再通过局部异常因子 (LOF) 应用于异常分量以此进行异常检测。

2) 属性增强的时空图卷积网络 (ASTGCN)^[25]: 该模型以卷积网络为基础,引入时空注意力机制来捕获时空特征。

3) 时空对抗神经网络 (STGAN)^[22]: 通过捕捉道路的时空特征,并结合 GAN 的数据生成能力进行交通流预测。

4) 镜像时间图自编码器 (MTGAE)^[26]: 该方法用于捕捉交通网络中的异常和节点间的时空关联,增强了特征提取能力,并通过高维映射识别复杂交通网络中的异常行为。

2.4 实验结果

2.4.1 异常检测实验分析

不同基线模型下的召回率如表 2 和表 3 所示(最优值加粗标示,下同)。可以看出,CG-STGAN 在 PeMS 和 NYC 数据集上的异常检测性能表现出色。

表 2 PeMS 数据集上不同模型的召回率对比

Table 2 Recall comparison of different models on PeMS dataset

数据集	模型	1%异常情况	2%异常情况	3%异常情况	4%异常情况	10%异常情况
PeMS	ASTGCN	0.16	0.23	0.29	0.40	0.55
	ST-Decompn	0.20	0.34	0.43	0.52	0.75
	STGAN	0.22	0.40	0.52	0.56	0.83
	MTGAE	0.24	0.43	0.55	0.60	0.85
	CG-STGAN	0.27	0.48	0.61	0.71	0.88

表 3 NYC 数据集上不同模型的召回率对比

Table 3 Recall comparison of different models on NYC dataset

数据集	模型	0.2%异常情况	0.4%异常情况	0.6%异常情况	0.8%异常情况	1%异常情况
NYC(Hit Events)	ASTGCN	9	13	14	14	16
	ST-Decompn	2	7	16	17	17
	STGAN	9	15	17	18	19
	MTGAE	9	14	16	18	18
	CG-STGAN	10	15	18	19	19

对于 PeMS 数据集,为了更好地判断模型的异常检测能力,本文取出 1.5%的异常情况进行检测,并设置了 10%的召回率,表 2 是不同模型的召回率具体情况。对于 NYC 数据集,用基线模型 ASTGCN 进行检测,由表 3 可以看出,在 1%的异常情况下,ASTGCN 命中了 16 次异常事件,而 CG-STGAN 命中了 19 次异常事件。可见在同样的数据集和相同的实验参数下,在原有模型基础上加入 Cluster-GCN 后的模型在召回率方面有明显的上升,幅度约为 6%。同时,在 NYC 数据集下,异常检测的准确率也有一定的提升。可以看出,本文提出

的方法在不同的情况下均取得了较高的性能。

实验结果进一步证实了模型在复杂异常交通流数据中使用图卷积聚合方法的实用性。特别地,结果表明 Cluster-GCN 对于提高异常交通流预测起到了积极的作用。可以明显看出 CG-STGAN 模型能够有效地提高模型的异常检测效率,表明其能够保持良好的异常检测性能。

2.4.2 交通流预测实验分析

如表 4 所示,CG-STGAN 在 PeMS 和 NYC 数据集上的预测性能表现出色,在对应的预测时间范围内,CG-STGAN 都取得了优异的结果。

表 4 CG-STGAN 与基线模型在不同预测层(horizons)的性能比较

Table 4 Performance comparison between CG-STGAN and baseline models in different prediction horizons

模型	指标	PeMS				NYC			
		15 min	30 min	45 min	60 min	30 min	60 min	90 min	120 min
ST_decompn	RMSE	25.02	25.07	25.11	25.14	14.65	14.84	14.92	14.97
	MAE	12.38	12.41	12.45	12.48	3.55	3.63	3.66	3.68
ASTGCN	RMSE	7.35	7.37	7.44	7.48	6.40	6.61	6.65	6.68
	MAE	3.52	3.54	3.55	3.57	1.95	2.04	2.10	2.14
STGAN	RMSE	5.01	5.03	5.11	5.16	5.12	5.25	5.31	5.38
	MAE	2.09	2.11	2.13	2.14	1.18	1.26	1.29	1.32
MTGAE	RMSE	4.97	4.99	5.07	5.13	4.85	5.10	5.19	5.24
	MAE	2.02	2.04	2.08	2.11	1.18	1.23	1.26	1.28
CG-STGAN	RMSE	4.83	4.88	4.97	5.03	4.65	4.82	4.95	5.12
	MAE	1.95	2.01	2.04	2.06	1.07	1.15	1.19	1.21

在 PeMS 数据集中,在 60 min 时长下,相较于传统模型如 MTGAE,CG-STGAN 模型的 RMSE 和 MAE 分别改进了约 1.9%和 2.4%。在 NYC 数据集中,在 120 min 时长下,CG-STGAN 模型与其

他机器学习模型如 ST-Decompn、ASTGCN 和 STGAN, RMSE 下降了约 2.3%, MAE 下降了约 5.5%。对表 4 的详细分析表明,CG-STGAN 模型在性能上明显优于大多数基线模型。这种优势源于

CG-STGAN 结合了 GAN 生成异常交通流数据的能力和 Cluster-GCN 节省内存占用的优点,使得该方法能够更好地拟合异常交通流数据的复杂性和非线性特征。实验结果充分证明了 CG-STGAN 在异常交通流预测上的可行性。该模型不仅在预测精度上有显著提升,还能更好地适应复杂的数据特征和非线性关系,为异常交通流预测提供了更加可靠和准确的解决方案。

交通流的异常预测相比正常情况下更具不确定性,增加了预测难度,导致大多数现有模型表现不佳。然而,CG-STGAN 在异常交通流预测中表现出色。它利用 GAN 解决了异常交通数据稀少的问题,

并通过 Cluster-GCN 优化内存占用。CG-STGAN 的主要优势在于结合了 GAN 和 Cluster-GCN,使模型能更准确地反映异常情况对交通流的影响,从而提高预测性能。

为了更直观地展示 CG-STGAN 模型,将实验结果进行可视化。图 5 展示了模型在 PeMS 数据集上的表现,其中图 5(a)为模型的 RMSE 结果,图 5(b)为 MAE 结果。此外,图 6(a)和图 6(b)分别展示了模型在 NYC 数据集上的 RMSE 和 MAE 表现。如图 5、图 6 所示,随着时间的增加,CG-STGAN 模型的预测误差明显低于其他基线模型,取得了良好的实验效果,充分证明了本文模型预测的有效性。

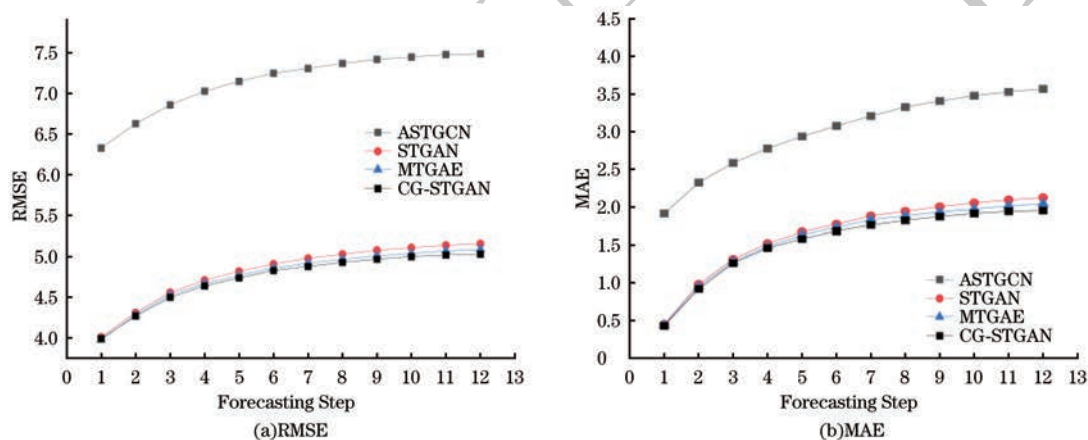


图 5 PeMS 数据集上的性能比较

Fig. 5 Performance comparison on PeMS dataset

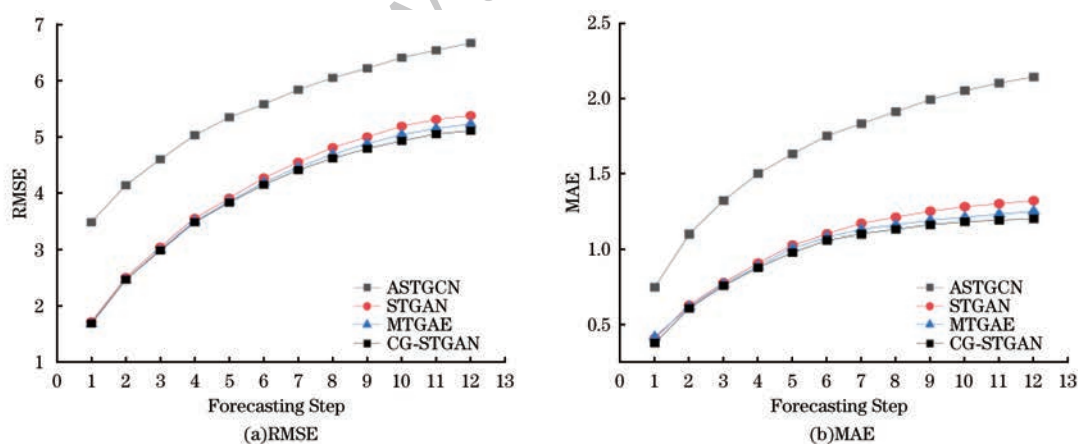


图 6 NYC 数据集上的性能比较

Fig. 6 Performance comparison on NYC dataset

2.4.3 不同数据集上的内存使用情况

通过表 5 展示了 CG-STGAN 模型在两个不同数据集上的内存使用情况,突出其即使在处理大规模图数据时也能维持较低的内存占用。CG-STGAN 在降低内存使用方面的卓越性能,在两种数据集上均取得了显著成果。

表 5 显示,与 MTGAE 相比,使用 Cluster-GCN

显著降低了 CG-STGAN 的内存占用,最多减少 33%。在深入分析改进后的模型时,观察到内存占用显著降低,并且当模型架构中增加新层时,主要引入的变量是权值矩阵。Cluster-GCN 通过聚类技术优化权值矩阵的存储和计算,有效控制每个新增层的内存需求。与传统 GCN 相比,Cluster-GCN 减少了冗余连接和参数存储,从而优化了内存使用并提

高了整体算法的运行效率。

表 5 不同数据集上的内存使用情况比较

Table 5 Comparison of the memory usage on different datasets

数据集	层数	STGAN/MB	MTGAE/MB	CG-STGAN/MB
PeMS	2	5 982	3 735	2 503
	3	6 138	4 021	2 945
	4	6 221	4 348	3 305
NYC	2	5 406	3 560	2 428
	3	5 575	3 775	2 534
	4	5 630	3 991	2 668

综上所述, Cluster-GCN 的创新之处在于其有效避免了图卷积邻域扩张带来的开销,从而节省了内存并提升了模型性能。这一特性在处理大规模交通流数据时尤为重要。

模型内存占用比较如图 7 所示,可以看出,本文提出的 CG-STGAN 模型相较于 MTGAE 模型而言能够有效降低模型的内存占用率。

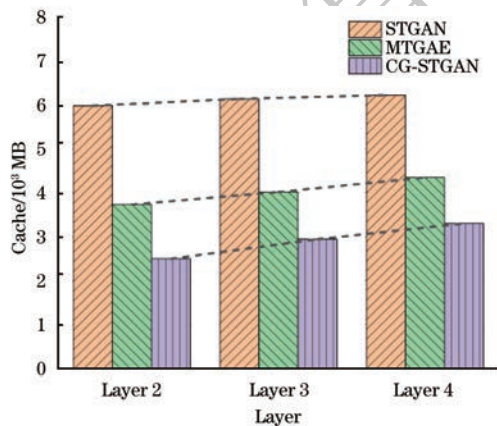


图 7 模型内存占用比较

Fig.7 Model memory usage comparison

2.5 消融实验

由于模型由 3 个独立部分组成,本文在 PeMS 和 NYC 数据集上进行了消融实验,以验证不同组合对模型内存占用和交通流预测的影响。使用多个生成器对现有数据进行测试,表 6 展示了模型各变体的预测误差,其中, No-Recent 表示不含短期时间相关性模块生成器测试; No-Trend 表示不含长期时间相关性模块生成器测试; No-External 表示不含外部相关性模块生成器测试; No-Cluster-GCN 表示不含聚类图卷积测试。

通过消融实验能够观察到, CG-STGAN 能够准确地捕捉到动态交通特征的变化趋势。特别地,当缺失短期时间相关性模块时,模型将无法稳定建立起时空特征,致使模型预测精度下降。而在长期预测中,当道路趋势信息不存在时, CG-STGAN 将无

表 6 模型在数据集上的表现

Table 6 Performance of the models in the dataset

数据集	模型	RMSE	MAE
PeMS	No-Recent	16.78	9.57
	No-Trend	5.11	2.15
	No-External	5.06	2.09
	No-Cluster-GCN	5.16	2.24
	Proposed Model	5.03	2.06
NYC	No-Recent	19.10	4.01
	No-Trend	5.39	1.35
	No-External	5.33	1.30
	No-Cluster-GCN	5.46	1.48
	Proposed Model	5.12	1.21

法学习到长期依赖信息,最终也将导致模型的预测误差增大。与此同时,可以观察到如果没有外部信息模块,模型也面临着较高的预测误差。特别是在对应的时间步长的预测中,相较于仅考虑单一模块的情况,其在 RMSE 和 MAE 方面平均下降了 1.5%。最后,根据表 6 显示的结果,模型中缺失 Cluster-GCN 时,模型的内存使用增加,且由于无法充分利用节点间的分组信息,模型的预测精度有所下降, MAE 和 RMSE 平均上升了 4.6% 和 16%。

综上,当 3 个模块共同作用且加入聚类图卷积时,周边道路信息将会提供更丰富的上下文,进而模型可以更精准地选择存储和处理哪些信息,这将有助于优化内存占用。通过将周边道路外部信息进行特征融合,可以显著增强模型的感知能力,通过融合不同来源的信息,可以更有效地利用内存资源。因此,通过实验不难发现,当加入周边道路信息后模型的预测性能得到显著提升,进一步也验证了模型在异常交通流预测的过程中的有效性。

3 结束语

本文提出的 CG-STGAN 模型在异常交通流预测中取得了显著成果,解决了现有模型难以捕捉复杂动态模式以及计算资源消耗过大的问题。通过结合 GAN 和 Cluster-GCN, CG-STGAN 显著提升了异常交通流预测的准确性,并降低了内存需求。实验结果表明,该模型在大规模数据处理中的优势显著,不仅提高了预测精度,还优化了资源使用效率。未来将进一步探索复杂时空依赖关系和多样化交通影响因素,并尝试引入多图卷积结构以提升模型性能。

参考文献

- [1] 郭瑞, 胡国梁, 王俊茗. 车联网中可匿名的无证书聚合签名方案[J]. 计算机工程, 2024, 50(11): 207-222.

- GUO R, HU G L, WANG J M. Anonymous certificateless aggregate signature scheme in VANETs [J]. Computer Engineering, 2024, 50(11): 207-222. (in Chinese)
- [2] DESHPANDE, N, HYOSHIN J. Physics-informed deep learning with Kalman filter mixture for traffic state prediction [J]. International Journal of Transportation Science and Technology, 2025, 17: 161-174.
- [3] LI Z X, CAO J D, SHI X L, et al. QPSO-AHES-RC: a hybrid learning model for short-term traffic flow prediction [J]. Soft Computing, 2023, 27(14): 9347-9366.
- [4] LIN G C, LIN A J, GU D L. Using support vector regression and K-nearest neighbors for short-term traffic flow prediction based on maximal information coefficient [J]. Information Sciences, 2022, 608: 517-531.
- [5] ZHUANG W Q, CAO Y B. Short-term traffic flow prediction based on a K-nearest neighbor and bidirectional long short term memory model [J]. Applied Sciences, 2023, 13(4): 2681.
- [6] NARMADHA S, VIJAYAKUMAR V. Spatio-Temporal vehicle traffic flow prediction using multivariate CNN and LSTM model [J]. Materials Today: Proceedings, 2023, 81: 826-833.
- [7] CHAUHAN N S, KUMAR N, ESKANDARIAN A. A novel confined attention mechanism driven Bi-GRU model for traffic flow prediction [J]. IEEE Transactions on Intelligent Transportation Systems, 2024, 25(8): 9181-9191.
- [8] FANG Z, LONG Q Q, SONG G J, et al. Spatial-temporal graph ODE networks for traffic flow forecasting [C] // Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery and Data Mining. New York, USA: ACM Press, 2021: 364-373.
- [9] BAO Y X, LIU J L, SHEN Q Q, et al. PKET-GCN: prior knowledge enhanced time-varying graph convolution network for traffic flow prediction [J]. Information Sciences, 2023, 634: 359-381.
- [10] HUANG X, YE Y, YANG X, et al. Multi-view dynamic graph convolution neural network for traffic flow prediction [J]. Expert Systems with Applications, 2023, 222: 119779.
- [11] CHEN Y B, LI K L, YEO C K, et al. Traffic forecasting with graph spatial-temporal position recurrent network [J]. Neural Networks, 2023, 162: 340-349.
- [12] GUO K, HU Y L, SUN Y F, et al. Hierarchical graph convolution network for traffic forecasting [J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2021, 35(1): 151-159.
- [13] 翟志鹏, 曹阳, 沈琴琴, 等. 基于多时空图融合与动态注意力的交通流预测 [J]. 计算机工程, 2025, 51(9): 139-148.
- ZHAI Z P, CAO Y, SHEN Q Q, et al. Traffic flow prediction based on multiple spatio-temporal graph fusion and dynamic attention [J]. Computer Engineering, 2025, 51(9): 139-148. (in Chinese)
- [14] LI H T, MA Y J, WANG X, et al. Graph spatiotemporal pattern learning network for real-time road network traffic abnormal incident detection [J]. Transportation Research Record: Journal of the Transportation Research Board, 2023, 2677(12): 815-829.
- [15] CHEN Q, WANG W, HUANG X, et al. Attention-based recurrent neural network for traffic flow prediction [J]. Journal of Internet Technology, 2020, 21(3): 831-839.
- [16] JANG H C, CHEN C A. Urban traffic flow prediction using LSTM and GRU [J]. Engineering Proceedings, 2024, 55(1): 86.
- [17] 陈亮, 吴攀, 刘韵婷, 等. 生成对抗网络 GAN 的发展与最新应用 [J]. 电子测量与仪器学报, 2020, 34(6): 70-78.
- CHEN L, WU P, LIU Y T, et al. Development and application of the latest generation against the network of GAN [J]. Journal of Electronic Measurement and Instrumentation, 2020, 34(6): 70-78. (in Chinese)
- [18] WU C, CHEN L, WANG G B, et al. Spatiotemporal scenario generation of traffic flow based on LSTM-GAN [J]. IEEE Access, 2020, 8: 186191-186198.
- [19] BASHAR M A, NAYAK R. TAnoGAN: time series anomaly detection with generative adversarial networks [C] // Proceedings of the IEEE Symposium Series on Computational Intelligence (SSCI). Canberra, Australia: IEEE Press, 2020: 1778-1785.
- [20] ALEXANDER G, LIU D Y, ALNEGHEIMISH S, et al. TadGAN: time series anomaly detection using generative adversarial networks [C] // Proceedings of the IEEE International Conference on Big Data. Atlanta, USA: IEEE Press, 2021: 33-43.
- [21] XU L Y, XU K, QIN Y C, et al. TGAN-AD: Transformer-based GAN for anomaly detection of time series data [J]. Applied Sciences, 2022, 12(16): 8085.
- [22] DENG L Y, LIAN D F, HUANG Z Y, et al. Graph convolutional adversarial networks for spatiotemporal anomaly detection [J]. IEEE Transactions on Neural Networks and Learning Systems, 2022, 33(6): 2416-2428.
- [23] CHIANG W L, LIU X Q, SI S, et al. Cluster-GCN: an efficient algorithm for training deep and large graph convolutional networks [C] // Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. New York, USA: ACM Press, 2019: 257-266.
- [24] ZHANG M Y, LI T, SHI H Z, et al. A decomposition approach for urban anomaly detection across spatiotemporal data [C] // Proceedings of the 28th International Joint Conference on Artificial Intelligence. Macao, China: International Joint Conferences on Artificial Intelligence Organization, 2019: 6043-6049.
- [25] GUO S N, LIN Y F, FENG N, et al. Attention based spatial-temporal graph convolutional networks for traffic flow forecasting [J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2019, 33(1): 922-929.
- [26] REN Z Y, LI X J, PENG J, et al. Graph autoencoder with mirror temporal convolutional networks for traffic anomaly detection [J]. Scientific Reports, 2024, 14: 1247.

文字编辑 金胡考
栏目编辑 赖玉玲