DOI:10.11918/202203059

# 基于时空特征融合的网络流量预测模型

薛自杰1,卢昱妃1,宁 芊1,黄霖宇1,陈炳才2

(1. 四川大学 电子信息学院,成都 610065;2. 大连理工大学 计算机科学与技术学院,辽宁 大连 116024)

摘 要:随着网络规模的日益增大,实时准确的网络流量预测对流量调度、路由设计等工作至关重要。由于网络流量数据的 非线性和不确定性,一些传统方法无法取得较好的预测精度。针对网络流量复杂的时空特征,本文提出一种基于时空特征融 合的神经网络(ST-Fusion)进行流量预测。该模型采用编码器 – 解码器结构。首先,编码器具有时间和空间两个并行的特征 通道,联合门限循环网络和自注意力机制提取流量的时序特征,采用图卷积神经网络提取流量的空间特征;然后,将编码器提 取的时空特征使用双边门限机制进行特征融合;最后,将融合的结果输入到基于门限循环网络的解码器中依次生成预测结 果。本文在3个公开的网络流量数据集(GEANT、ABILENE、CERNET)上进行实验,其评价指标选用 MAE、RMSE、ACCURACY、 VAR。实验结果表明 ST-Fusion 方法能够取得更好的预测效果。

**关键词:** 网络流量预测;特征融合; 双边门限机制; 图卷积神经网络; 门限循环网络; 自注意力机制 中图分类号: TP399 **文献标志码:** A **文章编号:** 0367 - 6234(2023)05 - 0030 - 09

# Network flow prediction based on spatial-temporal features fusion

XUE Zijie<sup>1</sup>, LU Yufei<sup>1</sup>, NING Qian<sup>1</sup>, HUANG Linyu<sup>1</sup>, CHEN Bingcai<sup>2</sup>

(1. College of Electronics and Information Engineering, Sichuan University, Chengdu 610065, China;

2. School of Computer Science and Technology, Dalian University of Technology, Dalian 116024, Liaoning, China)

**Abstract**: With the increasing scale of network, the accurate and real-time prediction of network flow is essential for traffic scheduling and routing design. However, due to the nonlinearity and uncertainty of network flow data, some traditional methods fail to achieve good prediction accuracy. Considering the complex spatial – temporal features of network flow, a novel network flow prediction method based on spatial-temporal features fusion (ST-Fusion) was proposed, combined with encoder-decoder architecture. First, the encoder was designed with two parallel feature channels: temporal and spatial channels. The temporal features were extracted by integrating gated recurrent unit (GRU) and self-attention mechanism, and the graph convolutional network (GCN) was used to extract the spatial features. Then, the temporal and spatial features were input into the GRU-based decoder to generate prediction results. Experiments were conducted on three public datasets (GEANT, ABILENE, and CERNET) using evaluation metrics including MAE, RMSE, ACCURACY, and VAR. Experimental results showed that the ST-Fusion method achieved better performance in network flow prediction.

Keywords: network flow prediction; features fusion; bilateral gated mechanism; graph convolutional network (GCN); gated recurrent unit (GRU); self-attention mechanism

随着互联网的不断发展,网络提供的服务日益 多样化,网络环境越来越复杂。如何优化网络结构、 保障网络稳定运行成为网络领域中急需解决的问 题。传统网络管理模式由于架构的限制对网络进行 调度优化越来越困难。而软件定义网络(softwaredefined networking,SDN)将网络中的控制平面和数 据平面分开,控制平面可以采集数据平面的数据从 而进行预测,然后基于链路状态生成控制策略。控 制策略通过流表下发至交换机,交换机根据流表信 息控制数据转发最终实现网络调度优化。流量预测 被部署于控制平面内,是网络调度的基础。研究精 准实时的流量预测算法对路由调度、优化网络设计、 保证网络服务质量、提升用户体验等具有重要意义。

一般而言,网络流量使用流量矩阵表示。流量 OD(origin destination)矩阵表示网络节点之间的流 量,即特定时间点网络中所有的原始和目的节点对

收稿日期:2022-03-18;录用日期:2022-06-24;网络首发日期:2022-08-29 网络首发地址: https://kns.cnki.net/kcms/detail/23.1235.T.20220826.1621.022.html 基金项目:四川省科技厅重点研发项目(2021YFQ0011);国家自然科学基金(61961040) 作者简介:薛自杰(1998—),男,硕士研究生;陈炳才(1976—),男,教授,博士生导师 通信作者:宁 芊,ningq@ scu.edu.cn

之间的流量。流量预测方法分为模型驱动和数据驱动两类。模型驱动是根据先验知识对网络进行系统 建模从而实现预测,但是其无法应对复杂的动态网络,所以数据驱动是目前研究的热点。其中基于统 计学的方法是根据已知数据得到最优回归函数实现 预测,最常用的是整合移动平均自回归(autoregressive integrated moving average, ARIMA)模型。文献[1-3]将 ARIMA 方法用于流量预测。但 ARIMA 模型 要求数据具有平稳性,不适用于快速变化的网络。 而基于数据特征进行预测的方法如有支持向量回归 (support vector regression, SVR)、K-最近邻算法 (K-nearest neighbors, KNN)、贝叶斯网络模型、神经 网络模型等具有更好的适应性。文献[4]第一次提 出使用 SVR 进行链路负载的预测。

近些年,随着深度学习的快速发展,越来越多的 神经网络模型被应用于流量预测并且取得较好的预 测效果。文献[5-6]将 BP 神经网络和深度置信网 络应用于预测任务中,证明两者在网络流量预测中 优于传统方法。循环神经网络(recurrent neural network, RNN)和基于 RNN 的长短期记忆网络(long short-term memory, LSTM)和门控循环神经网络 (gated recurrent unit, GRU)都使用递归循环机制执 行时间序列任务。文献[7-9]将 RNN、LSTM、GRU 等模型应用于 GEANT、ABILENE、Ties-1 等真实网 络流量中,证明 LSTM 和 GRU 比 RNN 更加适合流 量预测任务。还有一些工作是基于 RNN 的改进,例 如文献[10]在 LSTM 中增加星期、月份等额外的时 间信息,文献[11]将注意力机制应用于 GRU 中,文 献[12]将 GRU 单元进行堆叠和增加短连接结构, 文献[13]使用遗传算法对 LSTM 超参数进行优化, 文献[14-16]对流量数据进行小波分解等预处理 后再进行预测。这些工作通过缓解梯度问题、优化 超参数、进行数据预处理等方法提高预测精度。

上述方法只考虑到流量数据的时间特征,却忽略了邻近网络节点之间会互相影响,即忽略了流量的空间特征。所以文献[17]等工作使用卷积神经网络(convolutional neural network, CNN)或图卷积神经网络(graph convolution neural network, GCN)提取数据空间特征以提升预测效果。例如,文献[18-20]将GRU与GCN相结合提取时空特征从而进行流量预测。还有一些工作是基于GCN的改进。文献[21]将邻接矩阵加权来表示实际节点间的距离。文献[22]将注意力机制应用于GCN中,使模型关注于某些重要节点。还有工作利用其他方式进行空间特征提取。文献[23-24]使用CNN提取空间特征。文献[25]利用扩散卷积对空间特征

进行建模提取。文献[26]通过时空注意力使模型 关注到重要的节点和时刻。

现有方法虽然取得一定成果,但仍存在一些不 足。例如,基于 RNN 的模型相较于 BP 神经网络有 更强的时序特征提取能力,但 RNN 本身的递归循环 结构使模型很难对较长历史时刻的特征进行提取, 导致模型对长期依赖的建模能力不足。同时流量数 据具有时间和空间两种特征,但现有方法基本上采 用串行结构依次提取。这类方法在提取空间特征的 同时,数据的时间特征可能受到影响,这不利于时空 特征的完整提取。且串行结构在实际设备中运行速 度较慢,难以满足预测控制任务的实时要求。

为了解决上述问题,提升模型的预测效果和运行速度。本文提出一种新型的基于时空特征融合的神经网络(spatial-temporal features fusion,ST-Fusion)进行流量预测。模型采用编码器 – 解码器结构,其编码器由 GRU 和 self-attention 组成的时间编码器和基于 GCN 的空间编码器组成。它能够在时间、空间两个特征上同时捕捉流量的时空特征,然后通过双边门限机制进行自适应地特征融合。解码器则使用GRU 进行递归循环生成预测结果。

## 1 前 提

#### 1.1 GRU 结构

GRU 在 RNN 的基础上增加门控结构,使信息 能够选择性地在隐藏层传递,通过只保留重要信息 来解决长序列训练过程中可能出现的梯度消失和梯 度爆炸问题。GRU 有重置门和更新门两个门控结 构,相较于 RNN 具有参数少且收敛快等特点。

GRU 的数学表达式为

$$u_{\iota} = \sigma(W_{u}[X_{\iota}, h_{\iota-1}] + b_{u})$$
(1)

$$r_{t} = \sigma(W_{r}[X_{t}, h_{t-1}] + b_{r})$$

$$(2)$$

$$c_{t} = \tanh(\mathbf{W}_{c}[\mathbf{X}_{t}, r_{t} \cdot h_{t-1}] + b_{c})$$
(3)

$$h_{t} = u_{t} \cdot h_{t-1} + (1 - u_{t}) \cdot c_{t}$$
(4)

式中: $h_t$  为时刻 t 的隐层状态; $X_t$  为时刻 t 的流量信息; $r_t$  为重置门,表征忽略前一时刻信息的程度;  $u_t$  是更新门,表征上一时刻信息影响这一时刻的程度; $c_t$  是时刻 t 的记忆内容; $H_{t+1}$ 是时刻 t+1 的输出状态。GRU 利用时刻 t 的隐藏状态和当前输入去预测 t+1 时刻的流量状态。

#### 1.2 自注意力机制

循环神经网络常用于解决非线性时间序列的预 测问题。但 RNN 采用递归结构导致无法并行计算, 且较难捕获序列的长期依赖关系。因此,越来越多 的研究者采用注意力机制进行序列建模。其中基于 自注意力机制的 Transformer 模型以并行方式输入 数据,且并行执行所以不存在长期依赖问题。

由于 Transformer 采用并行注意力机制,这种机 制会使数据失去序列信息。而在时序任务中序列的 时序信息至关重要。所以在执行自注意力机制之 前,首先对数据进行位置编码,即将时序信息注入到 输入序列当中,如下:

$$u_{p,2i} = \sin(p/10\ 000^{2i/d_{\text{model}}}) \tag{5}$$

$$\boldsymbol{u}_{p,2i+1} = \cos(p/10\ 000^{2i/d_{\text{model}}}) \tag{6}$$

式中:**u** 为嵌入向量,p 为特征向量的位置;d<sub>model</sub>为特征向量的维度,每一个特征向量的位置编码都是由不同频率的余弦正弦函数组成,波长逐渐由 2π 增长到 10 000 · 2π。

多头注意力主要由单头注意力组成,单头注意 力是一种点积注意力,其通过缩放点积注意力来计 算特征矩阵的注意力值。点积注意力的输入由一组 查询、键、值组成,其中 d 代表输入数据特征维度,计 算查询和键的点积后除以√a 后得到值的加权权重, 点积注意力的计算公式如下所示:

Attention 
$$(\boldsymbol{Q}, \boldsymbol{K}, \boldsymbol{V}) = \operatorname{softmax}\left(\frac{\boldsymbol{Q}\boldsymbol{K}^{\mathrm{T}}}{\sqrt{d}}\right)\boldsymbol{V}$$
 (7)

式中: $Q = [q_1, q_2, \dots, q_n] \in R^{n \times d}$ 、 $K = [k_1, k_2, \dots, k_n]$  $\in R^{n \times d}$ 、 $V = [v_1, v_2, \dots, v_n] \in R^{n \times d}$ 分别为查询、键和 值, n 为个数。其中利用线性映射将节点的特征矩 阵映射到查询、键、值3个向量上,即: $Q = XW^{\circ}$ 、 $K = XW^{\kappa}$ 、 $V = XW^{\vee}$ 其中 $W^{\circ}$ ,  $W^{\kappa}$ ,  $W^{\vee} \in R^{d \times d_h}$ ,则自注意 力可以表示为

Attention 
$$(\boldsymbol{Q}, \boldsymbol{K}, \boldsymbol{V}) = \operatorname{softmax}\left(\frac{(\boldsymbol{X}\boldsymbol{W}^{\mathsf{Q}})(\boldsymbol{X}\boldsymbol{W}^{\mathsf{K}})^{\mathrm{T}}}{\sqrt{d}}\right)(\boldsymbol{X}\boldsymbol{W}^{\mathsf{V}})$$
  
(8)

自注意力机制可以使模型学习注意到输入特征 的某一方面的信息,但是单一的注意力机制只能注 意到单一信息,为了使模型同时注意多个方面的特 征,一般采用多头自注意力机制来进行注意学习。 多头自注意力机制由多个单头自注意力机制拼接后 线性映射得到。多头自注意力公式如下所示:

 $MultiHead(\boldsymbol{Q},\boldsymbol{K},\boldsymbol{V}) = Concat(h_1,h_2,\cdots,h_h)\boldsymbol{W}$ (9)

 $h_i$  = Attention( $XW_i^0, XW_i^K, XW_i^V$ ) (10) 式中: $W_i^0, W_i^K, W_i^V \in R^{d \times d_h}$ 是第 i 个注意力头  $Q_{\mathcal{K}} \mathcal{K} \mathcal{V}$ 的权重矩阵; $W \in R^{h \times d_h \times d}$ 是多头注意力权重矩阵;h为注意力头的数量; $d_h$  为每一个头的输出特征维 度;Concat 函数为拼接,用于拼接单头注意力的 结果。

#### 1.3 图卷积神经网络

给定一个邻接矩阵 A 和特征矩阵 X,GCN 在傅 里叶域内构造一个滤波器,该滤波器作用于图中的 节点,通过对其一阶领域进行卷积从而捕获节点之间的空间特征,可表示为

$$\boldsymbol{H}^{(l+1)} = \boldsymbol{\sigma} \left( \boldsymbol{D}^{-\frac{1}{2}} \hat{\boldsymbol{A}} \boldsymbol{D}^{-\frac{1}{2}} \boldsymbol{H}^{(i)} \boldsymbol{\theta}^{(i)} \right)$$
(11)

式中: $\hat{A} = A + I_N$  为增加自连接的单位矩阵; $\hat{D}$  为度 矩阵, $\hat{D} = \sum_{j} \hat{A}_{ij}$ ;  $H^{(i)}$  为第 *i* 层的矩阵; $\theta^{(l)}$  为该层 的参数; $\sigma(\cdot)$  为非线性激活函数。

# 2 基于 ST-Fusion 的流量预测

#### 2.1 问题描述

网络流量预测任务可以看作是一个图结构数据 的预测任务。本文使用非加权图 G = (V, E)来描述 网络拓扑结构,其中 V 为顶点集,即网络中路由交 换机等网络节点的集合;E 为边集,即节点之间链 路的集合;N 为节点总数。网络拓扑结构的邻接矩 阵定义为 $A \in R^{N \times N}$ ,其中元素的取值为0或1。若 节点之间没有链路连接,其该位置元素为0,否则 为1。网络 OD 流量矩阵  $X_t \in R^{N \times N}$ 是 t 时刻节点的 特征矩阵,将网络中的节点的流量信息作为节点的 特征。

综上所述,流量预测任务被描述为:基于拓扑 G 和若干特征矩阵 X,训练得到一个的映射 f,去预测 得到未来 T 个时间步的流量矩阵,即

 $[X_{i+1}, X_{i+2}, \cdots, X_{i+T}] = f[G, (X_{i-n+1}, X_{i-n+2}, \cdots, X_{i})]$ (12)

式中:n为输入时间序列的长度,T为输出时间序列的长度。

#### 2.2 ST-Fusion 模型框架

在本文中, ST-Fusion 模型如图 1 所示, 其由时 间编码器、空间编码器、融合单元、预测解码器组成。 首先将历史 N 个网络流量矩阵作为输入, 分别通过 空间编码器和时间编码器提取流量矩阵的空间特征 和时间特征。其中时间编码器使用 GRU 和自注意 力机制依次提取时间特征的短期和长期依赖, 空间 编码器使用 GCN 来捕获网络流量的空间特征, 从而 获得空间依赖。然后融合单元采用双边门控注意融 合机制对编码器得到的特征进行融合, 最终使用 GRU 递归解码输出预测结果。训练过程中使用平 均绝对误差优化整个网络。

#### 2.3 编码器

由于网络流量数据是典型的时空数据,即流量 具有时间和空间两个维度的特征。本文设计一个双 路编码器,即时间编码器和空间编码器,来并行提取 流量的时间特征和空间特征。



图 1 ST-Fusion 网络流量预测模型



## 2.3.1 时间编码器

时间编码器用于提取网络流量矩阵的时间特征。由于 GRU 网络是一种递归结构,这种结构可以 较好地捕捉短期流量特征,但无法有效捕捉长期流 量特征,所以本文采用自注意力机制再次捕捉 GRU 输出状态的时间特征,从而增强模型捕捉流量数据 长期特征的能力。

时间编码器的输入数据为X,定义为

$$\boldsymbol{X} = (N, H) \tag{13}$$

式中:N为网络中的节点个数,H为历史流量数据的 个数。

首先,流量数据被送入 GRU 层,其初步提取流量的时间特征,其中 GRU 的计算步骤见式(1)~(4)。

$$\boldsymbol{F}_1 = \mathrm{GRU}(\boldsymbol{X}) \tag{14}$$

然后将初步提取到的特征数据输入到基于自注 意力机制的 Transformer 中。Transformer 部分由基 本块堆叠而成,可再次进行特征提取。本文仅以第 *l* 个块的输出为例。首先输入到 *l*+1 模块中的多头 自注意力模块

 $M^{l+1} = MultiHead(O^{l}, O^{l}, O^{l})$  (15) 式中: $M^{l+1}$ 为多头注意机制得到的隐藏状态,查询矩 阵、键矩阵和值矩阵具有相同的状态, $O^{l}$ 为嵌入层 的输出。

随后的子层是位置前馈网络,由两个线性变换 组成,其中使用 ReLU 函数进行激活。

$$FFN(\boldsymbol{X}) = \boldsymbol{W}_2 \cdot ReLU(\boldsymbol{W}_1 \cdot \boldsymbol{x} + \boldsymbol{b}_1) + \boldsymbol{b}_2 \qquad (16)$$

$$\boldsymbol{O}^{l+1} = \mathrm{FFN}(\boldsymbol{M}^{l+1}) \tag{17}$$

式中 $W_1$ 、 $W_2$ 、 $O^{l+1}$ 为第l+1块的输出。在子层后使用残差连接和归一化。

2.3.2 空间编码器

空间编码器用于提取流量数据的空间特征。空间编码器的输入数据为*X*,定义为

$$\boldsymbol{X} = (N, H, \boldsymbol{F}) \tag{18}$$

式中:N为网络中的节点个数,H为历史流量数据的 个数,F为网络节点的特征。

本文使用两层 GCN 进行空间特征提取,其可表 示为

$$f(\boldsymbol{X}, \boldsymbol{A}) = \boldsymbol{\sigma}(\boldsymbol{A} \operatorname{ReLU}(\boldsymbol{A} \boldsymbol{X} \boldsymbol{W}_0) \boldsymbol{W}_1) \qquad (19)$$

式中: $\hat{A} = D^{-\frac{1}{2}}AD^{-\frac{1}{2}}$ 为数据预处理, $W_0 \in \mathbb{R}^{P \times H}$ 为输 入层到隐藏层的权重矩阵,其中 P 为特征矩阵的长 度,H 为隐藏单元数, $W_1 \in \mathbb{R}^{N \times T}$ 为从隐藏层到输出 层的权重矩阵,ReLU 为非线性激活函数。

本文使用上述 GCN 模型从流量数据中学习空间特征,GCN 可以提取路由与周围链路之间的拓扑 关系,从而得到流量的空间依赖性。

#### 2.4 时空特征融合层

由于网络场景的复杂性,流量的时间和空间特 征对未来流量的预测有不同的贡献度。所以动态调 整流量的时空特征进入编码器至关重要。本文采用 双边门控机制自适应动态调整这一过程。

将双路特征提取器中得到的时间特征 $f_{\text{temporal}}$ 和 空间特征 $f_{\text{spatial}}$ 输入到融合单元中。门限值c由时间 特征 $f_{\text{temporal}}$ 和空间特征 $f_{\text{spatial}}$ 决定。

$$c = \sigma(W_{\text{control}} \cdot [f_{\text{temporal}}, f_{\text{spatial}}])$$
(20)

式中: $f_{\text{temporal}} f_{\text{spatial}} \in R^{d \times d}, W_{\text{control}} \in R^{d \times 2d}, \sigma(\cdot)$ 是 sigmoid 激活函数,[]为拼接。

然后使用 *c* 引导时间特征,使用(1-*c*)引导空间特征。

 $f_{\text{fusion}} = f_{\text{temporal}} \odot c + f_{\text{spatial}} \odot (1 - c)$  (21) 式中: ①为哈代码积,  $f_{\text{fusion}}$ 为融合之后得到的流量的 时空特征。

#### 2.5 预测解码器

预测解码器基于融合后的时空特征进行流量预测。考虑到未来流量本身也是时序数据,所以本文 采用基于 GRU 结构的解码器递归进行预测。其中 GRU 结构的初始值为输入流量数据的最后一个值。

# 3 实验

在本节中通过实验验证模型的有效性。首先介 绍实验数据集、模型评价指标,然后介绍实验环境和 实验设置,最后是实验结果及其分析。

#### 3.1 网络流量数据集

本节评估了 ST-Fusion 模型在 3 个真实网络流 量数据集中的预测性能。GEANT 是面向研究和教 育界的泛欧数据网络,它将整个欧洲的国家研究和 教育网络(NREN)互相连接起来。其主要由 23 个 路由器组成,具有 38 个链路连接。ABILENE 主干 网位于北美,拓扑由 12 个节点、30 条无向链路组 成。网络流量数据是从 2004 年 3 月 1 日至 9 月 11 日每隔 5 min 采样得到的。中国教育和科研计算 机网(CERNET)主要连接中国国内高校,其主要包 含 14 个路由器节点和 32 条链路。

#### 3.2 评价指标

为了评价 ST-Fusion 模型在网络流量预测任务 中的性能,本文使用以下4种指标进行性能评价。

均方根误差(root mean squared error, RMSE)定 义为预测值与真实值之间的开方平均平方差,反映 测量极值效应和预测值的误差范围。

$$Y_{\rm rmse} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (Y_i - \hat{Y}_i)^2}$$
(22)

平均绝对误差(mean absolute error, MAE)定义 为平均误差绝对值,反映平均预测值的特异性。

$$Y_{\text{mae}} = \frac{1}{N} \sum_{i=1}^{N} |Y_i - \hat{Y}_i| \qquad (23)$$

准确率(accuracy)代表预测的准确率,定义如下:

$$Y_{\text{accuracy}} = 1 - \frac{\|Y - \hat{Y}\|_{\text{F}}}{\|Y\|_{\text{F}}}$$
(24)

Explained VARiance Score 为解释回归模型的方 差得分,其值取值范围是[0,1],越接近于1说明自 变量越能解释因变量。

$$Y_{\rm var} = 1 - \frac{\operatorname{var}(Y - Y)}{\operatorname{var}(Y)}$$
(25)

具体而言,对于 MAE、RMSE 指标,数值越小代 表模型效果越好。对于 ACCURACY、VAR 指标而 言,数值越大代表模型效果越好。

#### 3.3 实验环境和实验设置

在 Ubuntu16.04 系统环境下,使用 Pycharm 集成开发环境,使用基于 python3.8 的 Pytorch1.7 进行 实验。服务器配置如下: CPU 为 Intel Core i7-10700,GPU 为 Nvidia RTX 2070 SUPER。

训练模型前,需要对数据集进行预处理,将流量 矩阵中的每个元素进行归一化,使流量矩阵元素的 取值范围为[0,1]。

本文将数据集的 80% 作为训练数据,剩余 20% 作为测试数据。通过采用一个滑动窗口来生成序列 样本。每一组训练样本由 13 个时间步长组成,时间 步长为 5 min,其中前 10 个作为模型的输入,后 3 个 作为模型的预测输出。所有的基线模型都遵循其文 献中最佳参数和结果。实验中将 batch size 设置为 32,epoch 设置为 300,优化器采用 Adam,loss 函数为 MAE。

#### 3.4 基线模型

为了证明提出模型的有效性,本文选择一些基 于参数类和非参数类的经典方法进行比较,其基本 情况如下:

1) ARIMA,带有卡尔曼滤波的 ARIMA 被广泛 应用于时序建模任务中,通过拟合时序数据进行 预测。

2)SVR,采用拟合函数对时序进行回归。

3)GRU,通过重置门和更新门来解决 RNN 在时 序预测过程中梯度爆炸和长期依赖问题,进而提高 预测模型的预测效果。

4)GCN,通过考虑网络的拓扑结构来提高网络 流量预测的准确性。

5)T-GCN(Temporal GCN),通过融合图卷积和 带有门限机制的 RNN 来进行网络流量预测。

#### 3.5 训练初始化

在训练过程中,本文将历史数据的部分设置为 训练集  $y_{train}$ 、测试集  $y_{test}$ 、和测试过程中的验证集  $y_{valid}$ ,在训练集上采用 T + T'时间滑动窗口生成时间 训练序列,其中 T 作为输入,T'作为输出验证,使用 MSE 绝对平方误差作为损失函数,使用公式表示 如下:

loss = 
$$\sum_{t=0}^{y_{\text{train}}-T'} \left( \sum_{i=1}^{T'} | \hat{Y}_{t+T+i} - Y_{t+T+i} | \right)$$
 (26)

 $(Y_{i+T+1}, \dots, Y_{i+T+T}) = f(Y_{i+1}, \dots, Y_{i+T})$  (27) 式中f代表本文提出的预测模型,它通过将 MAE 作 为损失函数进行优化。

# 4 实验结果及分析

表1对比了 ST-Fusion 模型和其他基线模型在 GEANT、ABILENE、CERNET 3 个数据集中的预测性 能。本文将未来15 min、30 min、60 min 的3次预测 结果取平均值。其中,RMSE、MAPE 反映预测误差, 其值越小代表预测效果越好,而 ACCURACY、VAR 分别体现预测精度和自变量解释因变量的方差变化 的能力,其值越大代表预测效果越好。根据评价指标可得,ST-Fusion模型预测结果在所有的评价指标中取得最好成绩,既证明ST-Fusion模型较单纯时间

预测或空间预测具有较好的预测能力,同时也证明 其比时间、空间特征串行提取的时空预测模型 T-GCN有更好的预测效果。

表1 基于 GEANT、ABILENE、CERNET 数据集不同模型的流量预测结果

Tab.1 Prediction results of different models on GEANT, ABILENE, and CERNET datasets

数据集	模型 _	15 min				30 min				60 min			
		RMSE	MAE	ACCURACY	VAR	RMSE	MAE	ACCURACY	VAR	RMSE	MAE	ACCURACY	VAR
GEANT	ARIMA	0.037 98	0.021 24	0. 101 67	0.010 36	0.037 98	0.021 24	0.101 66	0.009 98	0.037 98	0.021 24	0.101 58	0.009 97
	SVR	0.088 35	0.084 31	0.314 47	0.686 12	0.088 71	0.084 59	0.319 92	0.67042	0.088 96	0.084 76	0.323 52	0.660 04
	GRU	0.024 07	0.007 00	0.641 91	0.860 22	0.026 96	0.007 72	0.598 91	0.824 69	0.027 53	0.007 99	0.59042	0.817 12
	GCN	0.024 36	0.008 96	0.63744	0.856 70	0.027 92	0.010 23	0.584 46	0.811 77	0.029 78	0.010 97	0.55676	0.785 85
	T-GCN	0.021 07	0.009 07	0.686 51	0.898 24	0.025 29	0.010 25	0.623 71	0.850 12	0.026 73	0.009 36	0.602 24	0.818 40
	ST-Fusion	0.02040	0.005 75	0.69638	0.89958	0.024 98	0.006 80	0.628 26	0.850 23	0.027 04	0.007 75	0.597 64	0.824 21
CERNET	ARIMA	0.070 94	0.046 83	0.319 16	0.000 28	0.070 94	0.046 83	0.319 16	0.000 28	0.070 94	0.046 83	0.319 15	0.000 27
	SVR	0.073 81	0.061 53	0.483 32	0.718 23	0.075 25	0.062 68	0.473 28	0.699 70	0.076 14	0.063 39	0.467 04	0.687 96
	GRU	0.038 09	0.019 09	0.733 36	0.874 80	0.041 12	0.021 05	0.712 11	0.854 03	0.042 17	0.021 15	0.704 84	0.846 14
	GCN	0.035 42	0.020 13	0.752 00	0.891 66	0.046 75	0.026 13	0.672 76	0.811 36	0.047 85	0.027 02	0.665 07	0.802.38
	T-GCN	0.030 25	0.016 50	0.788 22	0.921 00	0.035 75	0.019 39	0.749 92	0.889 82	0.038 78	0.021 21	0.728 52	0.870 16
	ST-Fusion	0.028 96	0.014 10	0.797 25	0.928 40	0.035 15	0.017 28	0.753 96	0.89436	0.038 36	0.019 32	0.731 45	0.873 21
ABILENE	ARIMA	0. 155 31	0.129 13	0.408 61	0.006 52	0. 155 31	0.129 13	0.408 58	0.006 48	0.155 34	0.129 16	0.40848	0.006 14
	SVR	0.056 94	0.039 67	0.792 60	0.940 06	0.065 19	0.046 04	0.762 57	0.917 54	0.071 02	0.050 05	0.741 37	0.899 81
	GRU	0.037 34	0.019 61	0.864 06	0.968 86	0.050 93	0.027 46	0.814 56	0.942 06	0.059 24	0.033 02	0.784 34	0.921 62
	GCN	0.048 26	0.034 16	0.824 30	0.949 60	0.065 16	0.044 44	0.762 78	0.906 72	0.067 81	0.046 15	0.753 12	0.898 75
	T-GCN	0.029 06	0.019 10	0.892 21	0.980 45	0.046 68	0.028 27	0.830 04	0.951 33	0.055 92	0.033 86	0.79640	0.930 18
	ST-Fusion	0.026 58	0.013 66	0.903 23	0.984 32	0.045 71	0.024 46	0.833 58	0.953 34	0.055 65	0.030 80	0.79740	0.930 84

注:表中黑体代表最优效果。

将表1中的预测结果按照数据集、预测时刻进行可视化,见图2。可以看到,在15 min 的流量预测 任务中,ST-Fusion 相较于 T-GCN 模型,在3个数据 集中,RMSE分别减少3.14%、4.26%、8.52%。同 样在30 min 的预测任务中,RMSE 减少约1.20%、 1.69%、2.08%。同样在60 min 的预测任务中, RMSE 减少约1.15%、1.07%,但是在CERNET 数 据集上相差0.49%。这表明ST-Fusion 模型较 T-GCN 更能够提取网络流量数据的空间特征,采用并 行特征提取可以有效减少特征损失。

为了验证 ST-Fusion 模型是否能从网络流量数 据中提取时间特征和空间特征来提高预测效果,本 文将 ST-Fusion 模型与 GRU 和 GCN 模型进行比较。 从图 2 中可以看到,基于时空融合的方法比单纯时 间特征预测或者空间特征预测具有更好的精度。这 说明 ST-Fusion 模型能够很好地从网络流量数据中 捕获时空特征。例如,在 3 个数据集 15 min 的预测 任务中,相较于 GCN 模型,RMSE 指标分别降低了 16.25%、18.24%、44.92%,在 30 min 预测任务中, RMSE 指标分别降低了 10.53%、24.81%、29.84%, 在 60 min 预测任务中,RMSE 指标分别降低了 9.22%、19.82%、17.93%,说明 ST-Fusion 能够很好 地捕捉空间依赖性。相较于 GRU 模型,在 15 min 预测任务中, RMSE 指标分别降低了 15.23%、 23.96%、28.81%,在 30 min 预测任务中, RMSE 指 标分别降低了 7.34%、14.53%、10.26%,在 60 min 预测任务中, RMSE 指标分别降低了 1.79%、 9.01%、6.05%,说明 ST-Fusion 能够很好地捕捉时 间依赖性。

从图 2 可以看出, ST-Fusion 模型在 3 种不同拓 扑下均取得最好的预测表现, 说明 ST-Fusion 具有良 好的泛化能力, 可以在不同拓扑、具有不同时空特征 倾向的网络流量数据预测任务中均取得不错的成绩。

图 3 为 ST-Fusion 模型在 GEANT、ABILENE、 CERNET 数据集下,在 15 min 任务中的预测结果。 由图中预测曲线可知,ST-Fusion 模型对网络流量数 据进行预测时,其预测曲线相较于其他模型更加接 近真实流量曲线,这表示 ST-Fusion 取得很好的预测 效果。同时 ST-Fusion 模型在 3 种网络数据中均较 好地拟合真实曲线,说明其针对不同的网络拓扑都 具有较好的适应性结果。在一些突发的网络流量和 流量波动的峰值和低谷时,ST-Fusion 模型相较于 T-GCN模型更加接近真实曲线,取得更小的预测 误差。



图 2 ST-Fusion 模型在 3 个网络流量数据集 MAE、RMSE、ACCURACY、VAR 指标对比









Fig. 3 Prediction results of ST-Fusion model on GEANT, ABILENE, and CERNET datasets

## 5 结 论

提出一种新的基于时空特征融合的神经网络, 称为 ST-Fusion。将网络拓扑抽象为非加权图。其 中节点表示路由节点,边表示两个路由节点之间的 网络连接,路由与其他路由的流量表示为节点的属 性。模型采用编码器-解码器结构,首先,双流编码 器使用图卷积神经网络提取网络流量的空间结构, 使用门限神经网络和自注意力机制相结合的方式来 提取节点属性的动态变化。然后采用双边门控注意 力机制进行融合最终获得时空依赖。最后使用 GRU 递归循环解码进行流量预测。使用 ST-Fusion 进行网络流量预测,在 GEANT、ABILENE、CERNET 3个不同真实网络流量数据集中进行评估,同时与 ARIMA、SVR、GCN、GRU、T-GCN 模型进行对比,ST-Fusion 在不同的评价指标中均取得更好的性能。综 上所述,ST-Fusion 模型能够较好地从网络流量数据 中捕捉时空特征,采用 ST-Fusion 网络流量预测模型 有利于提高 SDN 网络中路由调度的合理性和设备 的利用率。由于提出的 ST-Fusion 模型本身并没有 加入网络流量所独有的特征约束,所以该模型适用 于数据本身具有时空两种特征的预测任务中。

# 参考文献

- FENG Huifang, SHU Yantai. Study on network traffic prediction techniques [ C ]//2005 International Conference on Wireless Communications, Networking and Mobile Computing. Wuhan: IEEE, 2005: 1041. DOI: 10.1109/WCNM. 2005.1544219
- [2] LANER M, SVOBODA P, RUPP M. Parsimonious fitting of longrange dependent network traffic using ARMA models [J].
   IEEECommunications Letters, 2013, 17 (12): 2368. DOI: 10. 1109/LCOMM. 2013. 102613. 131853
- [3] HONG W C. Traffic flow forecasting by seasonal SVR with chaotic simulated annealing algorithm [J]. Neurocomputing, 2011, 74 (12/ 13): 2096. DOI:10.1016/j.neucom.2010.12.032

- [4] BERMOLEN P, ROSSI D. Support vector regression for link load prediction[J]. Computer Networks, 2009, 53(2): 191. DOI: 10. 1016/j. comnet. 2008. 09. 018
- [5] BARABAS M, BOANEA G, RUS A B. Evaluation of network traffic prediction based on neural networks with multi-task learning and multiresolution decomposition [C]//2011 IEEE 7th International Conference on Intelligent Computer Communication and Processing. Cluj-Napoca: IEEE, 2011: 95. DOI: 10. 1109/ICCP. 2011. 6047849
- [6] JIA Yuhan, WU Jianping, DU Yiman. Traffic speed prediction using deep learning method [C]//2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC). Rio de Janeiro: IEEE, 2016: 1217. DOI:10.1109/ITSC.2016.7795712
- [7] VINAYAKUMAR R, SOMAN K P, POORNACHANDRAN P. Applying deep learning approaches for network traffic prediction
   [C]//2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI). Udupi: IEEE, 2017: 2353. DOI: 10.1109/ICACCI.2017.8126198
- [8] LAZARIS A, PRASANNA V K. An LSTM framework for modeling network traffic[C]//2019 IFIP/IEEE Symposium on Integrated Network and Service Management (IM). Arlington: IEEE, 2019: 19
- [9] RAMAKRISHNAN N, SONI T. Network traffic prediction using recurrent neural networks [C]//2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA). Orlando: IEEE, 2018; 187. DOI: 10.1109/ICMLA.2018.00035
- [10] ZHOU Qinzheng, LI Qianmu, YAN Han. Long short-term memory neural network for network traffic prediction [C]//2017 12th International Conference on Intelligent Systems and Knowledge Engineering (ISKE). Nanjing: IEEE, 2017: 1. DOI: 10.1109/ ISKE. 2017. 8258815
- [11] ZHU Yu, ZHU Junxiong, HOU Jie. A brand-level ranking system with the customized attention-GRU model [C]//Proceedings of the 27th International Joint Conference on Artificial Intelligence (IJCAI'18). [S.1.]: AAAI, 2018: 3947
- [12] TIAN Jiyu, QIN Jing, CHEN Liming. A novel method for network traffic prediction using residual Mogrifier GRU [C]//2021 IEEE World Congress on Services (SERVICES). Chicago: IEEE, 2021; 47. DOI: 10.1109/SERVICES51467.2021.00035
- [13] CHEN Juan, XING Huanlai, YANG Hai. Network traffic prediction based on LSTM networks with genetic algorithm [C]//International Conference on Signal and Information Processing, Networking and

Computers. Singapore: Springer, 2018: 411. DOI: 10. 1007/ 978-981-13-7123-3\_48

- [14] FAN Jianyong, MU Dejun, LIU Yang. Research on network traffic prediction model based on neural network [C]//2019 2nd International Conference on Information Systems and Computer Aided Education (ICISCAE). Dalian: IEEE, 2019: 554. DOI: 10.1109/ICISCAE48440.2019.221694
- [15] LU Haipeng, YANG Fan. A network traffic prediction model based on wavelet transformation and lstm network [C]//2018 IEEE 9th International Conference on Software Engineering and Service Science (ICSESS). Beijing: IEEE, 2018; 1. DOI: 10.1109/ ICSESS.2018.8663884
- [16] DU Shuang, XU Zhanqi, LÜ Jianxin. An EMD-and GRU-based hybrid network traffic prediction model with data reconstruction
   [C]//2021 IEEE International Conference on Communications Workshops (ICC Workshops). Montreal: IEEE, 2021: 1. DOI: 10.1109/ICCWorkshops50388.2021.9473822
- [17] 王天保,刘昱,郭继昌,等.图卷积神经网络行人轨迹预测算法[J].哈尔滨工业大学报,2021,53(2):53
  WANG Tianbao, LIU Yu, GUO Jichang, et al. Pedestrian trajectory prediction algorithm based on graph convolutional network
  [J]. Journal of Harbin Institute of Technology, 2021, 53(2):53. DOI:10.11918 /202006051
- [18] YANG Li, GU Xiangxiang, SHI Huaifeng. A noval satellite network traffic prediction method based on GCN-GRU [C]//2020 International Conference on Wireless Communications and Signal Processing (WCSP). Nanjing: IEEE, 2020; 718. DOI: 10.1109/ WCSP49889.2020.9299774
- [19] ZHAO Ling, SONG Yujiao, ZHANG Chao. T-GCN: a temporal graph convolutional network for traffic prediction [J]. IEEE Transactions on Intelligent Transportation Systems, 2019, 21(9): 3848. DOI:10.1109/TITS.2019.2935152

- [20] GUI Yihan, WANG Danshi, GUAN Luyao. Optical network traffic prediction based on graph convolutional neural networks [C]//2020 Opto-Electronics and Communications Conference (OECC). Taipei: IEEE, 2020; 1. DOI:10.1109/OECC48412.2020.9273638
- [21] YAO Zhenjie, XU Qian, CHEN Yongrui. Internet traffic forecasting using temporal-topological graph convolutional networks [C]//2021 International Joint Conference on Neural Networks (IJCNN). [s.l.]: IEEE, 2021; 1. DOI; 10.1109/IJCNN52387.2021.9534096
- [22] HE Kaiwen, HUANG Yufen, CHEN Xu. Graph attention spatialtemporal network for deep learning based mobile traffic prediction
   [C]//2019 IEEE Global Communications Conference (GLOBECOM).
   Waikoloa: IEEE, 2019: 1. DOI: 10.1109/GLOBECOM38437.
   2019.9013136.
- [23] WU Yuankai, TAN Huanchun, QIN Lingqiao. A hybrid deep learning based traffic flow prediction method and its understanding
   [J]. Transportation Research Part C: Emerging Technologies, 2018, 90: 166. DOI: 10.1016/j. trc. 2018.03.001
- [24] CAO Xiaofeng, ZHONG Yuhua, ZHOU Yun. Interactive temporal recurrent convolution network for traffic prediction in data centers
   [J]. IEEE Access, 2018, 6: 5276. DOI: 10.1109/ACCESS. 2017.2787696
- [25] ANDREOLETTI D, TROIA S, MUSUMECI F. Network traffic prediction based on diffusion convolutional recurrent neural networks [ C ]//IEEE INFOCOM 2019-IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS). Paris: IEEE, 2019: 246. DOI: 10. 1109/INFCOMW. 2019. 8845132
- [26] GUO Shengnan, LIN Youfang, FENG Ning. Attention based spatialtemporal graph convolutional networks for traffic flow forecasting [C]. Proceedings of the AAAI Conference on Artificial Intelligence, 2019, 33(1): 922. DOI:10.1609/aaai.v33i01.3301922

(编辑 苗秀芝)