

引用本文: 王星宇, 张慧, 蔡安亮, 等. 基于图神经网络的 SDON 性能预测模型[J]. 光通信技术, 2024, 48(3): 38–44.

基于图神经网络的 SDON 性能预测模型

王星宇¹, 张慧², 蔡安亮¹, 沈建华^{1*}

(1. 南京邮电大学 通信与信息工程学院, 南京 210003; 2. 深圳赛柏特通信技术有限公司, 广东 深圳 518000)

摘要: 网络性能预测是实现软件定义光网络(SDON)高效网络管理的关键, 但目前亟需一种能够以较低成本准确预测关键指标的网络性能预测模型。提出一种基于图神经网络的 SDON 性能预测模型, 该模型将 BiGRU 和 Self-Attention 机制相结合, 能够学习网络拓扑、路由和流量矩阵之间的复杂关系, 从而准确地估计网络中源/目的地的分组延迟、抖动以及丢包率, 并且可以应用于训练中未遇到的网络。实验结果表明, 在不同流量模型测试中, 所提模型相较于基线模型的平均绝对百分比误差(MAPE)性能有明显提升。

关键词: 图神经网络; 网络性能预测; 软件定义光网络; 自注意力机制; 光通信

中图分类号: TN256 **文献标志码:** A **文章编号:** 1002-5561(2024)03-0038-07

DOI: 10.13921/j.cnki.issn1002-5561.2024.03.008

开放科学(资源服务)标识码(OSID):



SDON performance prediction model based on graph neural network

WANG Xingyu¹, ZHANG Hui², CAI Anliang¹, SHEN Jianhua^{1*}

(1. School of Communication and Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing 210003, China;

2. CypressTel SHENZHEN Communication Technology Company, Shenzhen Guangdong 518000, China)

Abstract: Network performance prediction is the key to achieving efficient network management of software defined optical networks(SDON), but there is an urgent need for a network performance prediction model that can accurately predict key indicators at limited cost. A graph neural network-based SDON performance prediction model is proposed, which combines BiGRU and Self-Attention mechanisms to learn the complex relationships between network topology, routing, and traffic matrices, accurately estimating the packet delay, jitter, and packet loss rate of the source/destination in the network. This model can be applied to networks that have not been encountered during training. The experimental results show that in different traffic model tests, the proposed model has a significant improvement in average absolute percentage error (MAPE) performance compared to the baseline model.

Key words: graph neural networks, network performance prediction, software-defined optical network, Self-Attention mechanisms, optical communication

0 引言

随着数字化时代的到来, 光通信和光网络在现代

通信和网络领域中的地位日益凸显^[1-2]。在这一背景下, 软件定义光网络(SDON)作为一种创新性的技术, 正逐渐成为引领光领域发展的关键驱动力之一^[3]。SDON 通过分离网络的控制平面和数据转发平面^[4], 使控制平面在光网络资源分配、路由配置以及服务质量(QoS)保障方面发挥核心作用。这种分离实现了网络的灵活编程和智能化管理, 从而大幅提升了网络的性能和可靠性。然而, SDON 在数据中心、算力网络等领域, 面临着巨大的挑战^[5], 包括对低延迟和高吞吐量的需求、网络规模的复杂性以及流量模式的快速变化。若要解决这些问题, 只有在 SDON 控制平面中准确地对网络建

收稿日期: 2024-01-27。

基金项目: 国家自然科学基金青年基金项目(62301284)资助; 南京邮电大学企业委托研发重点课题(KH0020322072)资助。

作者简介: 王星宇(1999—), 男, 硕士研究生, 现就读于南京邮电大学通信与信息工程学院电子信息专业, 主要研究方向为光网络、软件定义网络, 参与了南京邮电大学横向科研课题“新一代智能 SD-WAN 体系架构及其关键技术”研发。

*** 通信作者:** 沈建华(1972—), 男, 工学博士, 教授, 硕士生导师, 主要研究方向为高速率大容量光纤通信系统、智能光网络、软件定义网络等。



模,精确预测 SDON 环境下的网络性能,才能进行光网络的资源分配和网络优化(如路由优化、链路负载均衡等)^[9]。

学术界关注到机器学习(ML)和深度学习(DL)等人工智能技术在 SDON 网络性能预测方面的潜力。DL 作为一种更先进的方法,能够构建精确的数据驱动网络模型,无需对物理网络进行假设,从而确保网络模型的准确性。SDON 控制器监视网络流量、拓扑结构和性能指标^[7],并提供实时的网络状态信息和分析报告。将经过训练的 DL 模型集成到 SDON 控制器中,能够实现对网络性能的高效且精确预测,从而为资源分配和网络优化提供强有力的支持。WANG M 等人^[8]探讨了基于真实网络数据学习的数据驱动模型。文献[9-10]也尝试使用了全连接神经网络,即多层感知器(MLP)。然而,这些基于当前技术的早期网络尝试无法泛化到训练集之外的其它网络环境,不仅未在真实的流量模型中得到充分测试,而且也未考虑到 QoS 感知的队列调度配置。随着研究的深入,出现了更为复杂的神经网络模型,如变分自编码器^[11]和卷积神经网络(ConvNN)^[12],但这些模型也与 MLP 模型有相似的限制。文献[13-14]虽然使用了图神经网络(GNN),但仅仅采用了基本的 GNN 架构,并简化了网络模型,没有充分考虑流量模型、排队策略或泛化到更大网络的关键特性。

为了准确地预测 SDON 环境下的网络性能,在消息传递神经网络(MPNN)的基础上,本文结合双向门控循环单元(BiGRU)和自注意力(Self-Attention)机制的网络性能,提出一种基于 GNN 的 SDON 性能预测模型——BiGRU-AT 模型。

1 模型构建

BiGRU-AT 模型可以根据给定的网络样本准确地预测 SDON 中的各项网络性能指标。网络样本的信息包括:1)当前网络的拓扑结构信息;2)端到端的路由方案(流级别);3)节点的排队配置(接口级别);4)一组由一些参数表征的流量(流量矩阵)。该模型能够预测流级别的网络性能值,这些值包括但不限于延迟、抖动和丢包率等关键指标。

模型的设计主要基于 2 个原则:1)寻找模型支持的网络组件的有效表示方法(如流量模型、路由方案、节点的队列调度策略等);2)利用与网络规模无关的特征,可以确保模型在训练时对新遇到的更大规模网络保持准确的适用性。

BiGRU-AT 模型的整体框架如图 1 所示。该模型需要先对网络样本进行预处理,并且通过自定义的消息传递算法来对网络性能预测模型的关键元素进行信息交换和更新,然后通过 Self-Attention 层强化主要特征,最后通过 MLP 机读出网络性能的预测值。

1.1 网络样本预处理

1.1.1 网络元素表示

在 SDON 中,网络的定义^[15]包括:1 组源-目的地流集合 $F = \{f_i: i \in (1, \dots, n_f)\}$ 、1 组链路集合 $L = \{l_j: j \in (1, \dots, n_l)\}$ 以及 1 组节点(交换机或路由器)的队列集合 $Q = \{q_k: k \in (1, \dots, n_q)\}$,其中 i, j, k 分别表示为流集合 F 、链路集合 L 和节点队列集合 Q 的大小,根据网络样本中的路由配置,流集合 F 中的每一条流从源节点出发,途径多个链路和节点的队列,最后抵

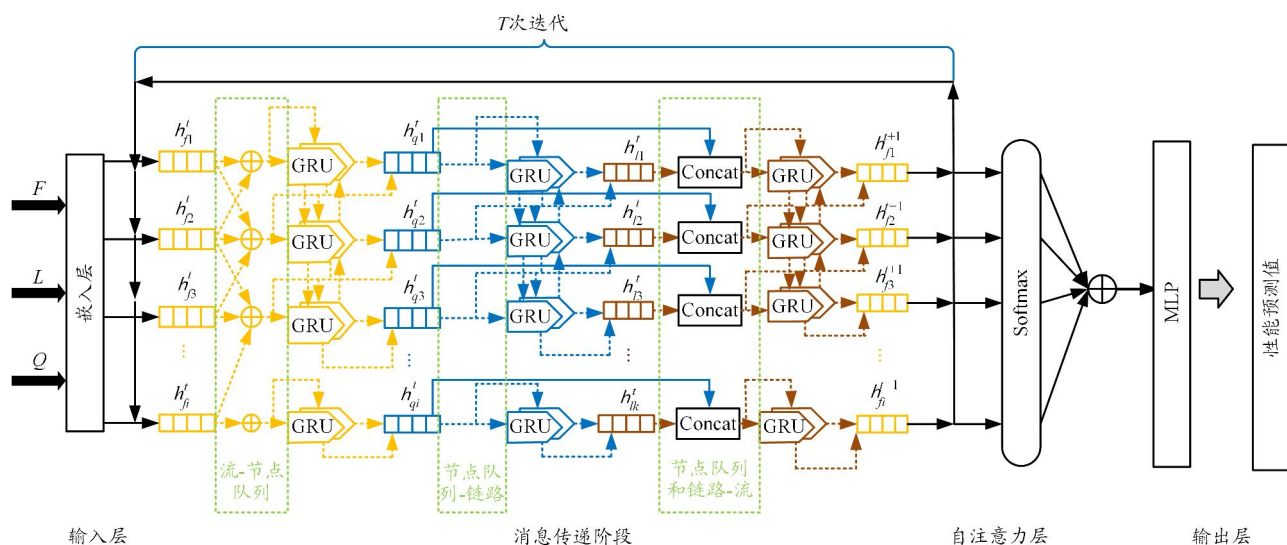


图 1 BiGRU-AT 模型整体框架

王星宇,张慧,蔡安亮,等: 基于图神经网络的 SDON 性能预测模型

达目的节点。因此,将流集合 F 中的流定义为元组序列 $f_i = \{(q^{i,1}, l^{i,1}), \dots, (q^{i,M}, l^{i,M})\}$, M 是流 f_i 的路径长度。

根据上述的网络组件定义,可以看出 BiGRU-AT 模型的输入包含 3 个主要的网络组件:网络流集合 F 、形成网络拓扑的链路集合 L 以及节点的输出端口的队列集合 Q 。基于这些网络组件的定义,可以得出以下 3 个基本原则:1)流的状态(如延迟、吞吐量、丢失),取决于其遍历的队列及链路的状态(如队列/链路利用率)的影响。2)队列的状态(如占用),取决于通过队列自身的流的状态(如流量、突发性)。3)链路的状态(如利用率),取决于可能将流量注入链路队列的状态和应用于这些队列的队列调度策略(如严格优先级、加权公平队列)。

1.1.2 网络元素特征嵌入

虽然网络性能预测模型中的关键元素是流集合 F 、链路集合 L 以及节点队列集合 Q ,但是模型的原始输入的形式并不适用于神经网络进行处理,需要将其转换为神经网络能够更好处理的形式。本文采用 $\text{dropout}=0.1$ 的 3 层 MLP 与 Self-Attention 层结合的预处理层,用于特征嵌入。嵌入层神经网络模型如图 2 所示。

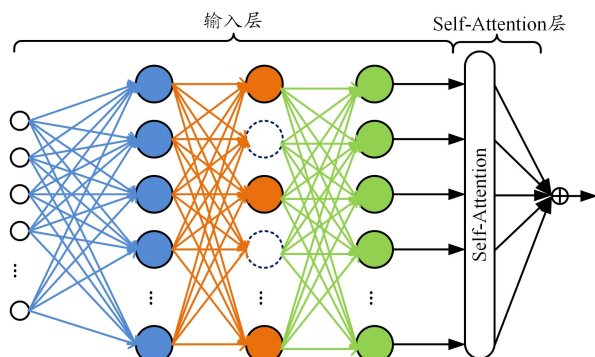


图 2 嵌入层神经网络模型

该神经网络模型是对初始特征进行预处理的模型。即该模型会对初始特征向量 x_f 、 x_l 和 x_q 进行处理,将其转化为隐藏状态向量。其中,初始特征向量 x_f 包括了流集合 F 的平均传输流量和对应流量模型的一些参数配置,如 On-Off 流量模型中的参数 t_{on} 和 t_{off} (t_{on} 表示流量处于活跃状态的持续时间, t_{off} 表示流量处于非活跃状态的持续时间)或者指数流量模型中的流量发生率 α 和流量持续时间 β 等信息;初始特征向量 x_l 则包括了链路 l_j 的链路负载和与输出流量到该链路的节点端口的队列调度策略等信息;初始特征向量 x_q 包

括了队列 q_k 的队列缓冲区大小和策略的权重等信息。

将特征嵌入到固定大小的向量后,通过自注意力层对这些向量进行处理,以进一步强化输入的主要特征。Self-Attention 网络中进行的计算如下:

$$Q = W^Q X \quad (1)$$

$$K = W^K X \quad (2)$$

$$V = W^V X \quad (3)$$

$$\text{Attention}(Q, K, V) = \text{Softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (4)$$

其中, X 表示全连接层输出数据, Q 、 K 、 V 分别表示查询向量、键向量、值向量, W^Q 、 W^K 、 W^V 分别表示 Q 的查询矩阵、 K 的键矩阵、 V 的值矩阵, d_k 表示特征维度, K^T 表示键向量 K 的转置。

将初始特征进行转换,得到嵌入后的隐藏状态向量,便可以进行 3 种关键元素之间的消息传递和状态更新。

1.2 BiGRU-AT 模型构建

1.2.1 BiGRU 神经网络

递归神经网络(RNN)在处理时间序列数据时虽然具有一定的优势,但是当处理特别长的序列时,容易出现梯度消失或梯度爆炸的问题,这导致它难以在长序列上有效地传递梯度,从而影响了其长期学习的能力。为了克服这一缺陷,HOCHREITER S 等人^[16]提出了长短期记忆(LSTM)网络。LSTM 网络是对 RNN 的一种改进,其增加了遗忘门、输入门和输出门 3 个关键门结构,这些结构使得 LSTM 网络保留重要信息同时遗忘不重要的信息,从而能够处理长期依赖问题。但是,这也使得 LSTM 网络的结构变得更为复杂,增加了训练的难度。基于提高效率和节约计算时间的考虑,CHUNG J 等人^[17]在 LSTM 的基础上提出了门控循环单元(GRU),GRU 的结构与 LSTM 相似,但其内部计算更简化。

本文使用 BiGRU 来进行流、链路以及节点 3 种关键元素之间的消息传递,其模型结构如图 3 所示。该模型同时受到先序信息所带来的正向反馈和后序信息所带来的逆向反馈,并且通过增加控制单元门的使用,有效地减少过拟合现象。同时,结合更多的信息也有助于提高消息传递的准确度。

图 3 中, $\omega_1 \sim \omega_6$ 表示不同的权重值, x_i 表示输入, h_i 表示输出,其中每个 GRU 单元计算公式如下:

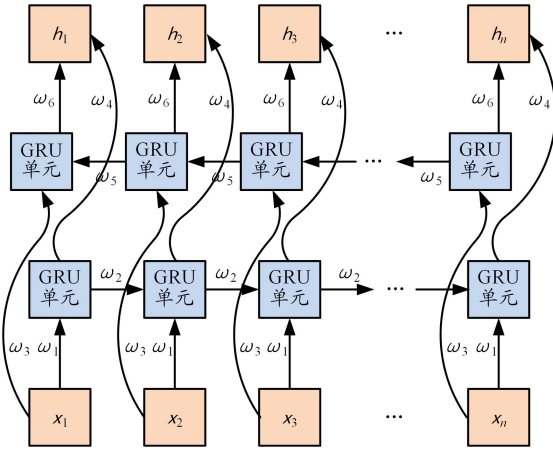


图3 BiGRU 模型结构

$$\begin{cases}
 z_t = \sigma(W_z[h_{t-1}, x_t]) \\
 r_t = \sigma(W_r[h_{t-1}, x_t]) \\
 h'_t = \tanh(\omega[r_t h_{t-1}, x_t]) \\
 h_t = (1 - z_t) h_{t-1} + z_t h'_t
 \end{cases} \quad (5)$$

其中, h_{t-1} 表示上个节点输出的隐状态, W_z 和 W_r 表示更新门和重置门的权重矩阵, z_t 和 r_t 分别表示更新门和重置门的输出, h'_t 表示经过重置门后的记忆内容, h_t 表示经过了门控单元的遗忘和记忆更新后的输出。

1.2.2 模型整体架构

BiGRU-AT 模型的整体方案如算法 1 所示。本文遵循 MPNN 范式, 该模型分为 3 个步骤:

1) 初始特征嵌入(如算法 1 中的 1)–3)所示): 根据 1.1.2 节的描述进行网络元素的初始特征嵌入, 将初始特征向量 x_f, x_l, x_q 预处理为 h_f^0, h_l^0, h_q^0 。

2) 消息传递阶段(如算法 1 中的 4)–19)所示): 进行 T 次消息传递, 每个消息传递迭代可分为流级别、链路级别以及节点队列级别的消息交换和更新。具体而言, 4)–10)是利用 CONCAT 函数聚合了与网络流相关的链路和节点队列隐藏状态信息即 h_q^t 和 h_l^t , 并通过 BiGRU 即 $\psi([h_q^t, h_l^t])$ 从正、反 2 个方向进一步提取特征 $h_{f,l}^t$, 并保存为 $s_{f,q}^{t+1}$ 。之后, 将 BiGRU 中的最终内部状态作为流的新的特征嵌入 s_q^{t+1} , 完成更新。11)–14)是对网络中的节点队列进行消息交换和更新, 聚合与该节点队列相关的流的信息 S_q^{t+1} , 然后通过 BiGRU 来提取特征 h_q^{t+1} 作为节点队列的新的特征嵌入, 并保存为 s_q^{t+1} 。15)–19)与上述提取特征操作类似, 通过聚合

与该链路相连接的所有节点队列的信息 s_q^{t+1} , 经过 BiGRU 提取特征 h_l^t , 并更新链路的特征嵌入 h_l^{t+1} 。

3) 读取阶段(如算法 1 中的 20)–26)所示): 消息传递结束后, 得到最终的流特征嵌入。首先, 通过一个 Self-Attention 层来强化主要的特征; 然后, 通过一个使用 Relu 激活函数的 3 层 MLP 来读取流经过队列时的延迟 d_q ; 最后, 与传播延迟 d 和传输延迟 d_p 相加即可得到预测的分组延迟, 即流级别的延迟 $y_{f,l}$ 。其中, x_l 是链路容量, $x_{f,l}$ 为数据包平均大小, $x_{l_{\text{loop}}}$ 为初始向量中传播延迟的赋值。

算法 1 BiGRU-AT 模型整体方案

输入: F, x_f, L, x_l, Q, x_q

输出: $y_{f,l}$

- 1) For each $f \in F$ do Embedding(x_f) $\rightarrow h_f^0$
- 2) For each $l \in L$ do Embedding(x_l) $\rightarrow h_l^0$
- 3) For each $q \in Q$ do Embedding(x_q) $\rightarrow h_q^0$
- 4) For $t=1$ to T do
- 5) For each $f \in F$ do
- 6) $\psi([\cdot, \cdot]) \leftarrow \text{BiGRU}(h_f^t, [\cdot, \cdot])$
- 7) For each $(q, l) \in f$ do
- 8) $h_{f,l}^t \leftarrow \psi([h_q^t, h_l^t])$
- 9) $s_{f,q}^{t+1} \leftarrow h_{f,l}^t$
- 10) $h_f^{t+1} \leftarrow h_{f,l}^t$
- 11) For each $q \in Q$ do
- 12) $S_q^{t+1} \leftarrow \sum_{f \in Q_f(q)} s_{f,q}^{t+1}$
- 13) $h_q^{t+1} \leftarrow \text{BiGRU}(h_q^t, S_q^{t+1})$
- 14) $s_q^{t+1} \leftarrow h_q^{t+1}$
- 15) For each $l \in L$ do
- 16) $\psi(\cdot) \leftarrow \text{BiGRU}(h_l^t, \cdot)$
- 17) For each $q \in L_q(l)$ do
- 18) $h_l^t \leftarrow \psi(s_q^{t+1})$
- 19) $h_l^{t+1} \leftarrow h_l^t$
- 20) For each $f \in F$ do
- 21) $y_{f,l} = 0$
- 22) For each $(q, l) \in f$ do

王星宇,张慧,蔡安亮,等: 基于神经网络的 SDON 性能预测模型

- 23) $d_q \leftarrow \text{MLP}(\text{Attention}(h_{t,l}^T))/x_{l_c}$
- 24) $d_t \leftarrow x_{f_c}/x_{l_c}$
- 25) $d_p \leftarrow x_{l_{\text{imp}}}$
- 26) $y_{l_i} \leftarrow d_q + d_t + d_p$

2 实验和数据分析

2.1 数据集

本文使用 OMNet++ 网络模拟器模拟各种网络场景来生成数据集。每一个数据集中的网络样本对应一次模拟,记录了网络中所有流的平均延迟、抖动、丢包数以及节点队列级别的信息,如平均占用率、平均分组大小和平均分组丢失等。通过对输入变量进行随机抽样,并选择不同的场景来进行模拟,具体的模拟输入变量包括网络拓扑、流量模型、流量强度以及节点的队列配置等。其中,网络拓扑包括 NSFNET、GEANT2 和 GBN;流量模型包括 On-Off、恒定比特率(CBR)、自相关指数(AE)、调制指数(ME)以及混合流量模型;流量强度为丢包率在 0~3% 的随机流量强度;节点的队列配置包括:每个节点包含 1~3 个队列,每个队列大小为 8、16、32 或 64 kb/s 以及调度策略。调度策略包括先进先出(FIFO)、严格优先级(SP)、加权公平队列(WFQ)以及递减式轮询调度(DRR)策略。

流量模型中的 AE 和 ME 能够准确地反映互联网流量的真实情况。流量强度表示网络的拥塞程度。在本文的数据集中,流量强度为 400~2 000 b/s,其中 400 b/s 是拥塞程度最低的网络(丢包率为 0%),2 000 b/s 是高度拥塞的网络(平均丢包率约为 3%)。

2.2 训练和结果

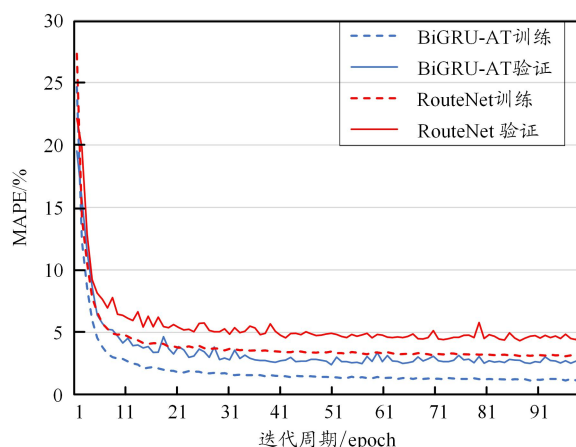
本文通过实验训练来验证 BiGRU-AT 模型在各种流量模型中的准确性。具体实验方式为:首先,通过系统地调整其一阶(方差)和二阶(自相关)统计量,逐步提升每个流量模型的内在复杂性。然后,将 NSFNET 数据集(14 节点)和 GEANT2(24 节点)数据集作为训练集,先进行 100 个 epoch 训练,再将 GBN 数据集(17 节点)作为验证集。每个 epoch 中训练步数为 2 000 步,验证步数为 200 步。同时,在尝试多种超参数组合进行实验后,得到 BiGRU-AT 模型的超参数值如表 1 所示。最后,从简单的流量模型(如 CBR)开始,逐步推进到测试几种更复杂的混合模型。这些混合模型的设计更加贴近真实互联网链路中的流量特性。实验使用的对比基线模型是 RouteNet 模型^[15],该模型是目前较为先进的网络性能预测模型,能够预测网络中的

表 1 BiGRU-AT 模型的超参数

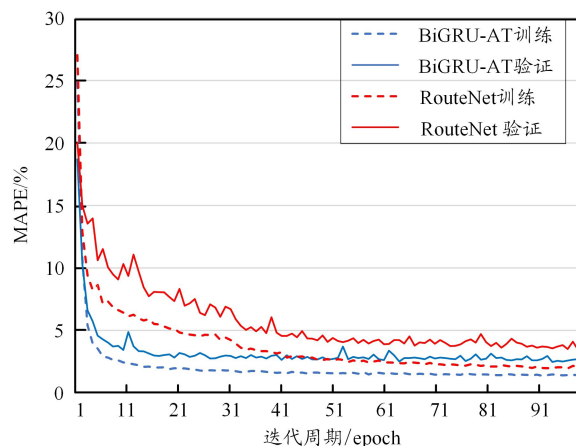
超参数	BiGRU-AT 模型
初始特征向量 x_t	32
初始特征向量 x_q	32
初始特征向量 x_g	32
消息传递阶段迭代次数/次	6
学习率	0.001
训练周期/epoch	100

延迟。

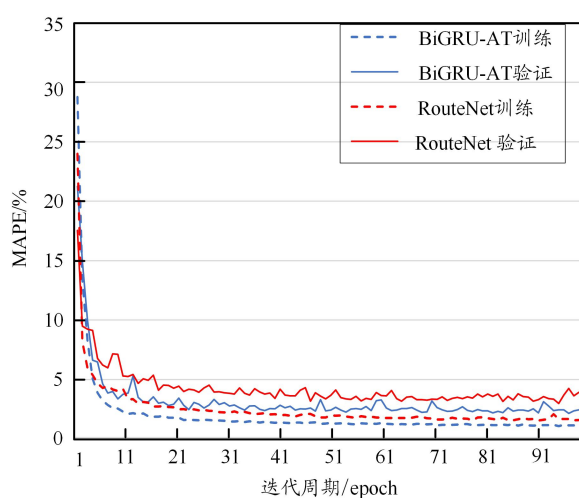
BiGRU-AT 模型与 RouteNet 模型在 5 种流量模型下的训练和验证结果如图 4 所示。可以看出,在 5 种流量模型中,BiGRU-AT 模型在训练集中的性能表现明显优于 RouteNet 模型,即在整体迭代周期中,BiGRU-AT 模型的 MAPE 值明显小于 RouteNet 模型;而在验证集中,BiGRU-AT 模型在整个迭代周期中的性能表现也要优于 RouteNet 模型。具体而言,在最简单



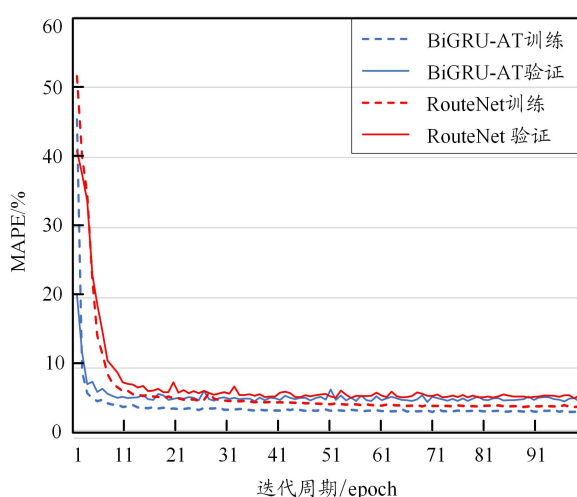
(a) CBR 流量模型



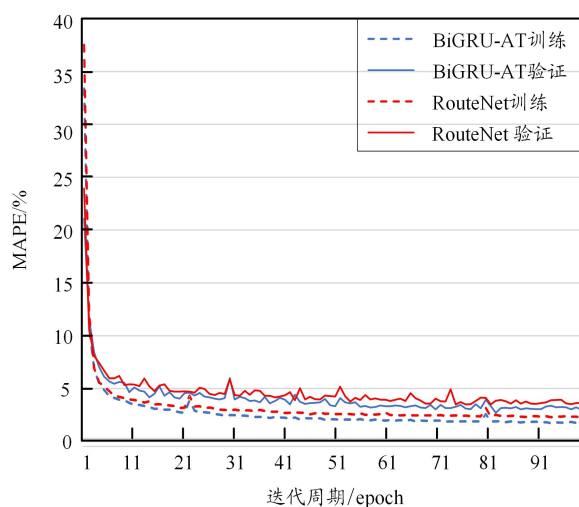
(b) On-Off 流量模型



(c) AE 流量模型



(d) ME 流量模型



(e) 混合流量模型

图4 BiGRU-AT模型和RouteNet模型
在5种流量模型下的训练与验证结果

的流量模型(CBR 流量模型)和非马尔可夫流量模型(On-Off 流量模型)中,相比 RouteNet 模型,BiGRU-AT 模型的最佳 MAPE 分别提升了 1.92% 和 1.13%,达到 2.40% 和 2.42%,从而能够提供更为准确的估计。在较复杂的具有强自相关(AE 流量模型)和重尾分布(ME 流量模型)的流量模型中,BiGRU-AT 模型相较于 RouteNet 模型性能有所提升,其最佳 MAPE 分别降低了 0.83% 和 0.49%,达到 2.16% 和 5.20%。由于这些模型近似于传输控制协议(TCP)生成的实际流量,因此 BiGRU-AT 模型能准确预测这些流量模型的延迟,这具有很强的实际意义。

在混合流量模型中,每个源-目的地从 On-Off、CBR、AE 或者 ME 这几种流量模型中随机选择其中一种,并使用相应模型的参数来生成流量。即在这种混合流量模型下,BiGRU-AT 模型仍然表现出良好的性能,其最佳 MAPE 能达到 2.71%,相比较于 RouteNet 模型,BiGRU-AT 模型的 MAPE 降低了 0.79%。

3 结束语

本文提出了一种基于 GNN 的 SDON 性能预测模型。该模型遵循 MPNN 范式,通过 Self-Attention 机制强化网络中的主要特征,并通过 BiGRU 结合双向处理和门控机制的优点,克服了基于 DL 模型的局限性,达到与排队论相同的目标,即在给定的 SDON 网络场景中准确地估计网络性能;此外,该模型支持与路由、队列调度和流量模型相关的各种配置参数,以及路由、队列调度和流量模型的任意组合。理论分析和实验结果表明:BiGRU-AT 模型在所有流量模型中对延迟的预测准确率均优于基线 RouteNet 模型,尤其是在具有复杂流量模型的场景中,仍然能表现出良好的 MAPE 性能。同时,该模型还可以结合 SDON 控制平面,在较短时间内对整体光网络的网络性能进行实时预测,为资源分配和网络优化提供有效的支持。

参考文献:

- [1] KAUSHAL H, KADDOUM G. Optical communication in space: challenges and mitigation techniques[J]. IEEE Communications Surveys & Tutorials, 2016, 19(1): 57-96.
- [2] KHALIGHI M A, UYSAL M. Survey on free space optical communication: a communication theory perspective [J]. IEEE Communications Surveys & Tutorials, 2014, 16(4): 2231-2258.
- [3] ZHAO Y, WANG Y, WANG W, et al. Software-defined optical networking (SDON): principles and applications[M]. Rijeka: InTech, 2017.
- [4] JI P N. Software defined optical network[C]//The 2012 11th International

王星宇, 张慧, 蔡安亮, 等: 基于图神经网络的 SDON 性能预测模型

al Conference on Optical Communications and Networks(ICOCN), November 28–30, 2012, Pattaya, Thailand. Pattaya: IEEE, 2012: 1–4.

[5] NSAFOA Y K, TCHAO E T, YEBOAH-AKOWUAH B, et al. Software-defined networks for optical networks using flexible orchestration: advances, challenges, and opportunities [J]. Journal of Computer Networks and Communications, 2022(1): 1–40.

[6] LIU P, JI W, LIU Q, et al. Failure prediction based on LSTM and SVM under SDON architecture [C]//2021 7th International Conference on Computer and Communications(ICCC), December 10–13, 2021, Chengdu, China. Chengdu: IEEE, 2021: 398–402.

[7] JHA R K, LLAH B N M. Software defined optical networks (SDON): proposed architecture and comparative analysis[J]. Journal of the European Optical Society-Rapid Publications, 2019, 15(1): 1–15.

[8] WANG M, CUI Y, WANG X, et al. Machine learning for networking: Workflow, advances and opportunities[J]. IEEE Network, 2017, 32(2): 92–99.

[9] MESTRES A, ALARCÓN E, JI Y, et al. Understanding the modeling of computer network delays using neural networks[C]//2018 Workshop on Big Data Analytics and Machine Learning for Data Communication Networks, August 20, 2018, Budapest, Hungary. New York: Association for Computing Machinery, 2018: 46–52.

[10] VALADARSKY A, SCHAPIRA M, SHAHAF D, et al. Learning to route [C]//16th ACM workshop on hot topics in networks, November 30–December 1, 2017, Palo Alto, USA. New York: Association for Computing Machinery, 2017: 185–191.

[11] XIAO S, HE D, GONG Z. Deep-q: traffic-driven qos inference using

deep generative network[C]//2018 Workshop on Network Meets AI & ML, August 24, 2018, Budapest, Hungary. New York: Association for Computing Machinery, 2018: 67–73.

[12] CHEN X, GUO J, ZHU Z, et al. Deep-RMSA: a deep-reinforcement-learning routing, modulation and spectrum assignment agent for elastic optical networks[C]//Optical Fiber Communication Conference, March 11–15, 2018, San Diego, USA. New York: Optica Publishing Group, 2018: 1–3.

[13] GEYER F, BONDORF S. DeepTMA: predicting effective contention models for network calculus using graph neural networks [C]//IEEE INFOCOM 2019-IEEE Conference on Computer Communications, April 29–May 2, 2019, Paris, France. New York: IEEE, 2019: 1009–1017.

[14] RUSEK K, SUÁREZ-VARELA J, MESTRES A, et al. Unveiling the potential of graph neural networks for network modeling and optimization in SDN [C]//2019 ACM Symposium on SDN Research, April 3–4, 2019, San Jose, USA: Association for Computing Machinery, 2019: 140–151.

[15] RUSEK K, SUÁREZ-VARELA J, ALMASAN P, et al. Routenet: leveraging graph neural networks for network modeling and optimization in SDN [J]. IEEE Journal on Selected Areas in Communications, 2020, 38(10): 2260–2270.

[16] HOCHREITER S, SCHMIDHUBER J. Long short-term memory[J]. Neural Computation, 2010, 9(8): 1735–1780.

[17] CHUNG J, GULCEHRE C, CHO K H, et al. Gated feedback recurrent neural networks [C]//32nd International Conference on Machine Learning, July 6–11, 2015, Lille, France. Cambridge: JMLR.org, 2015, 37: 2067–2075.