

## ◎热点与综述◎

## 基于深度学习的网络流量预测研究综述

康梦轩<sup>1,2</sup>, 宋俊平<sup>1</sup>, 范鹏飞<sup>1</sup>, 高博文<sup>3</sup>, 周旭<sup>1</sup>, 李琢<sup>1,2</sup>

1. 中国科学院 计算机网络信息中心, 北京 100190

2. 中国科学院大学, 北京 100049

3. 中国联合网络通信有限公司 北京市分公司, 北京 100038

**摘要:**精准地预判网络流量变化趋势可以帮助运营商准确预估网络的使用情况,合理分配并高效利用网络资源,以满足日益增长且多样化的用户需求。以深度学习算法在网络流量预测领域的进展为线索,阐述了网络流量预测的评价指标和目前公开的网络流量数据集及应用,具体分析了网络流量预测中常用的深度信念网络、卷积神经网络、循环神经网络和长短时记忆网络共四种深度学习方法,并重点介绍了近年来针对不同问题所提出的改进神经网络模型,总结了各模型特点及应用场景。最后对网络流量预测未来发展进行了展望。

**关键词:**深度学习;网络流量预测;深度信念网络;卷积神经网络;长短时记忆网络

**文献标志码:**A **中图分类号:**TP393 **doi:**10.3778/j.issn.1002-8331.2101-0402

## Survey of Network Traffic Forecast Based on Deep Learning

KANG Mengxuan<sup>1,2</sup>, SONG Junping<sup>1</sup>, FAN Pengfei<sup>1</sup>, GAO Bowen<sup>3</sup>, ZHOU Xu<sup>1</sup>, LI Zhuo<sup>1,2</sup>

1. Computer Network Information Center, Chinese Academy of Sciences, Beijing 100190, China

2. University of Chinese Academy of Sciences, Beijing 100049, China

3. China United Network Communications Co., Ltd., Beijing Branch, Beijing 100038, China

**Abstract:** Precisely predicting the trend of network traffic changes can help operators accurately predict network usage, correctly allocate and efficiently use network resources to meet the growing and diverse user needs. Taking the progress of deep learning algorithms in the field of network traffic prediction as a clue, this paper firstly elaborates the evaluation indicators of network traffic prediction and the current public network traffic data sets. Secondly, this paper specifically analyzes four deep learning methods commonly used in network traffic prediction: deep belief networks, convolutional neural network, recurrent neural network, and long short term memory network, and focuses on the integrated neural network models used in recent years for different problems. The characteristics and application scenarios of each model are summarized. Finally, the future development of network traffic forecast is prospected.

**Key words:** deep learning; network traffic prediction; deep belief networks; convolutional neural network; long short term memory network

根据2020年4月中国互联网信息中心发布的《第45次中国互联网络发展状况统计报告》显示,截至2020年3月份,我国共有网民9.04亿,与2018年底相比,增长7508万人,普及率增至64.5%。不仅如此,我国网民平均每周每人上网30.8小时,与2018年底相比,增长3.2小时<sup>[1-2]</sup>。网民数量的增长及上网时间的延长带来网络流量的激增。

随着5G、边缘计算、NFV等技术的发展,对网络进行精细化、自动化、智能化运维及管理将成为新的挑战。为了应对这一挑战,需要对边缘网络、城域网、骨干网等多个层级的应用级网络流量进行精准感知。而网络流量预测能力则是核心技术之一。精准的网络流量预测技术能够实现如下功能:(1)帮助改善通信网络管理。在分配网络资源的过程中,传统方法仅仅依靠网络

**基金项目:**国家自然科学基金(U1909204)。

**作者简介:**康梦轩(1994—),男,硕士研究生,研究领域为网络流量预测、深度学习、智能网络运营分析技术等,E-mail:kangmengxuan@cnic.cn;周旭(1976—),男,博士,研究员,研究领域为未来网络、5G移动网络、大数据技术等。

**收稿日期:**2021-01-22 **修回日期:**2021-03-05 **文章编号:**1002-8331(2021)10-0001-09

当前流量使用状态进行资源分配,缺乏对未来状态的预判,很容易导致网络不断拥堵或资源浪费。网络流量的准确预测可以帮助运营商及早应对即将出现的拥堵,提前进行网络扩容、调整和优化。(2)提升通信网络效能。伴随着5G基站的大规模部署,未来将建设更多的微基站,需要更多的电力和回程光纤。流量的准确预测,使得根据实际业务需求弹性分配资源成为可能。在闲时,网络需求量低,可以使低负载的部分基站进入休眠节能状态,并动态调整基站覆盖范围,待忙时,再激活这部分基站,进而减少基站能耗<sup>[3]</sup>。(3)定制扩展网络增值服务。目前,运营商资费和商业模式都较为单一,对业务和用户的个性化服务不足。对业务流量分别进行预测,有助于针对不同业务提供差异化网络服务。例如对视频流媒体、P2P业务等占用大量带宽的业务提前智能调度,对移动办公业务、网络游戏等实时性要求高且高价值的业务提前部署缓存等。

为了使得网络流量能够被精准预测,模型的选择和设计至关重要。目前预测模型分为线性模型和非线性模型两类。自回归模型(Autoregressive Model, AR)<sup>[4]</sup>、移动平均模型(Moving Average, MA)<sup>[5]</sup>、结合自回归和移动平均特点的自回归移动平均模型(Autoregressive Moving Average Model, ARMA)<sup>[6]</sup>和自回归合成移动平均模型(Autoregressive Integrated Moving Average Model, ARIMA)<sup>[7]</sup>等都是线性模型,其特点是需要人工凭借经验设置多种参数来拟合数据,仅适用于短期流量预测<sup>[8]</sup>。而实际网络流量具有非常多的特性,如非线性、周期性、自相似性、突发性等,仅用线性模型很难对其准确拟合预测<sup>[9]</sup>。因此本领域的研究者也研究了很多非线性模型。传统的非线性模型有小波模型<sup>[10]</sup>、自回归分数整合滑动平均模型(Fractional Autoregressive Integration Moving Average, FARIMA)<sup>[11]</sup>、门限自回归模型和灰色模型以及机器学习领域的支持向量机模型(Support Vector Machine, SVM)<sup>[12]</sup>、人工神经网络模型(Artificial Neural Network, ANN)等。

目前网络流量预测有以下难点:一方面网络流量在时间和空间上分布复杂;另一方面网络流量在骨干网、数据中心、边缘网等不同场景流量特征不同。仅用传统模型很难充分捕获其特征差异性,从而影响预测模型的设计、训练和性能提升。近年来研究者们通过应用深度学习的方法,在以上问题的处理中取得了较大进展。

因此,本文以深度学习算法在网络流量预测领域的应用为线索,阐述了网络流量的预测评价指标,介绍了目前公开的网络流量数据集,分析了基于深度学习的预测方法和改进的神经网络,并总结了各模型特点及应用场景,对网络流量预测未来发展进行了展望。

## 1 网络流量预测评价指标

网络流量预测模型的性能多用误差评价指标来衡

量。误差越大,则预测准确率越低,所建立的网络流量预测模型性能越差。整理研究人员常用的评价指标如下:

(1)平均绝对误差<sup>[13]</sup>(Mean Absolute Error, MAE)是绝对误差的平均值,它可以很好地反映预测值误差的实际情况,其取值范围是 $[0, +\infty]$ ,当预测值与真实值完全吻合时等于0,即完美模型,误差越大,该值越大。

$$MAE = \frac{1}{K} \sum_{i=1}^K |x_i - x'_i| \quad (1)$$

(2)均方根误差<sup>[14]</sup>(Root Mean Square Error, RMSE)也称标准误差,其取值范围是 $[0, +\infty]$ ,当预测值与真实值完全吻合时等于0,即完美模型,误差越大,该值越大。

$$RMSE = \sqrt{\frac{1}{K} \sum_{i=1}^K |x_i - x'_i|^2} \quad (2)$$

(3)R-squared<sup>[15]</sup>将预测结果转换为准确度,其取值范围是 $[0, 1]$ ,针对不同问题的预测准确度,可以比较并判断所提模型更适合预测哪一类问题。 $R^2$ 值越大越好,当所提模型无任何误差时 $R^2 = 1$ ;当所提模型等于基准模型时 $R^2 = 0$ 。

$$R^2 = 1 - \frac{\sum_{i=1}^K (x_i - x'_i)^2}{\sum_{i=1}^K (x_i - \bar{x}_i)^2} \quad (3)$$

在以上评价指标中, RMSE与MAE的量纲相同,但计算结果一般RMSE比MAE要大。这是因为RMSE是先对误差进行平方的累加后再开方,这实际上是放大了误差之间的差距。而MAE反映的就是真实误差。因此在衡量预测模型性能时RMSE的值越小其意义越大,因为它的值更能反映出模型的最大误差。上述指标都存在的一个问题就是,没有明确的上下限,来判定当前的模型是否足够好。而R-squared最大的好处是将预测的评价归一到 $[0, 1]$ 的区间,使得对模型评价有了相对统一的标准<sup>[15]</sup>。

## 2 网络流量数据集及应用

深度学习是一种监督学习方法,只有取得大量可信的网络流量数据,根据网络流量的重要特征建立合适的预测模型,并使其充分训练,掌握网络流量所蕴含的复杂特性,才能准确地预测未来网络流量,发挥应用价值。受用户隐私、流量信息敏感等因素影响,可用于研究的网络流量数据并不多,整理研究人员常用的数据集如下:

(1)意大利电信大数据挑战赛

该数据集是米兰市和特伦托自治省这两个区域的呼叫详细记录数据。数据采集者将米兰市划分为100行、100列的 $100 \times 100$ 个区域,每个区域大小约为 $235 \text{ m} \times 235 \text{ m}$ ,流量数据统计以区域为单位,每隔10 min统计各区域内用户收发的短信服务(SMS)、呼叫服务(Call)和移动用户上网服务(Internet)等流量总量。时间跨度为2013年11月1日零点到2014年1月1日零点。适用于

对城市中各小区流量间的复杂时空依赖关系进行流量预测实验<sup>[16]</sup>。

张传亭博士团队搭建卷积神经网络模型对该数据集中的短信业务和呼叫业务进行流量预测,目的是验证在时间和空间上无线业务流量数据是否存在复杂的依赖关系<sup>[16]</sup>。

(2)GEANT

该数据集是由萨里大学研究团队收集并整理的来自GEANT网络的流量信息。该流量信息由内部网关协议和边界网关协议路由信息组成,每隔15 min记录一次,以15×15的矩阵形式呈现。时间跨度为5个多月共167天。可用该数据集预测链路负载。

Reis等人使用该数据集来揭示链接流量和起点至目的地的流量关系,实现对链接路由性能的评估<sup>[17]</sup>。

(3)流量文库数据集

该数据集记录了网络流量文库(<http://newsfeed.ntcu.net/~news/2006/>)主节点路由器Incoming articles从2016年7月21日至2016年9月30日的访问流量,以小时为单位。共采集1 680条记录,以时间序列的方式呈现。该数据集可用于对网页负载量的预测研究<sup>[18]</sup>。

曹建华等人分别使用灰色模型、神经网络模型、未加补偿器的灰色神经网络模型和添加补偿器的灰色神经网络模型对该数据集进行流量预测实验,以实现网页流量负载预测<sup>[19]</sup>。

张晗等提出一种基于小波分解的网络流量时间序列的分析和预测方法。将非平稳的网络流量时间序列通过小波分解成为多个平稳分量,采用自回归滑动平均方法分别对各平稳分量进行建模,将所有分量的模型进行组合,得到原始非平稳网络流量时间序列的预测模型。实验利用网络流量文库的时间序列数据建立预测模型,并对其进行独立测试检验。结果表明,该预测方法提高了网络流量时间序列的预测准确率<sup>[20]</sup>。

(4)Wikipedia页面流量数据集

该数据集记录了Wikipedia网站上的文章访问情况,包括文章名称以及访问该文章所用的流量类型(有线网络和无线网络等),时间跨度为2015年7月1日至2016年12月31日,以天为单位,包含大约145 000条时

间序列数据。数据集可用于预测未来时刻的网页负载量,其缺点是数据源无法过滤零流量值和缺失值(缺失值意味着流量为零或当天没有可用数据)。

Mettu等人利用该数据集来进行网络流量预测实验。为了建立访问者与页面阅读数间的内在联系,评估网页受欢迎的程度,实验中使用Boosting的集成技术、长短时记忆网络(Long Short-Term Memory,LSTM)模型和时间序列建模技术ARIMA来预测Wikipedia网络流量<sup>[21]</sup>。

Petluri等人利用该数据集来预测Wikipedia未来流量。在实验中建立一个带有序列到序列方法的循环神经网络模型,将模型的结果与传统模型结果进行对比,表明提出的模型在预测维基百科文章未来流量方面具有有效性<sup>[22]</sup>。

Paun等人比较了两个现有模型,即小波神经网络和带有序列到序列方法的递归神经网络模型,利用该数据集来预测维基百科文章的未來流量,然后提出了一种序列到序列方法的长时短期记忆模型<sup>[23]</sup>,使用该模型可以预测维基百科文章的未來流量。

(5)Monash大学网络流量数据集

该数据集记录了Monash大学网站的流量信息,数据集采集者将该网站的流量信息划分为上行流量和下行流量,并通过后台网站服务器来分别记录传输流量。时间跨度为2003年3月9日至2003年4月13日,以小时为单位。该数据集可用于预测局域网网络流量的负载情况。

李丽采用该数据集开展网络流量未来负载研究,实验结果能够为局域网网络负载预测研究提供参考依据<sup>[13]</sup>。

(6)CAIDA(Center for Applied Internet Data Analysis)

CAIDA是一个对全球范围Internet结构及数据进行研究的国际合作机构。CAIDA在各种不同的链路和交换中心收集了不同种类的网络数据,目前拥有的数据量已经达到89 TB,并在尽可能保证数据提供者隐私的前提下,向研究界共享这些数据。根据收集方法的不同,CAIDA的数据可以分为被动测量数据、主动测量数据、外部数据和IMDC<sup>[24]</sup>。

数据集的总结如表1所示。

表1 网络流量数据集整理

数据集	数据集属性	时间跨度	大小	应用场景
意大利电信数据	呼叫(Call)、短信(SMS)、流量(Internet)和地理空间分布信息	2013.11.1—2014.1.1,共62天	约为19 GB	用于预测时空序列的网络流量
GEANT	时间序列的流量信息	2008.10.5—2009.3.21,共167天	约为17 MB	用于预测时间序列的网络流量
流量文库数据集	页面访问流量信息	2006.7.21—2006.9.30,共71天	约为41 GB	用于预测时间序列下页面访问量
Wikipedia页面流量数据集	页面访问流量信息	2015.7.1—2016.12.31,共549天	约为580 MB	用于预测时间序列下页面访问量
Monash大学网络流量数据集	大学服务器采集的信息	2003.3.9—2003.4.13,共35天	约为1.3 GB	用于预测局域网中时间序列的网络流量
CAIDA	骨干网、数据中心、边缘网流量数据等	实时更新	—	对应不同种类项目,可研究点多



### 3 基于深度学习的网络流量预测方法

近年来,深度学习因在网络流量预测方面具有显著效果而受到研究者的广泛关注。在深度神经网络被使用前,研究人员通常使用BP神经网络<sup>[25]</sup>、小波神经网络<sup>[26]</sup>等浅层神经网络进行网络流量预测,相较于浅层神经网络,深度神经网络可以从训练样本中提取出更高维、更抽象的特征,从而学习到浅层网络遗漏的潜在信息,提升网络流量预测的准确率。下面分别介绍用于流量预测的四种典型深度学习模型。

#### 3.1 深度信念网络

深度信念网络(Deep Belief Networks, DBN)<sup>[27]</sup>是由多个受限玻尔兹曼机(Restricted Boltzmann Machines, RBM)为学习模块的组合<sup>[28]</sup>。RBM是两层无向图模型,由一个可见层和一个隐藏层组成,且网络层间全连接<sup>[29]</sup>,每层中的每个单元都通过无向边与另一层的所有单元连接,同一层中的单元彼此断开。DBN由若干RBM堆叠而成,顶部两层是无向图结构的RBM,其余层自上而下构成有向图的概率生成模型<sup>[27]</sup>。一个三层深度信念网络如图1所示。

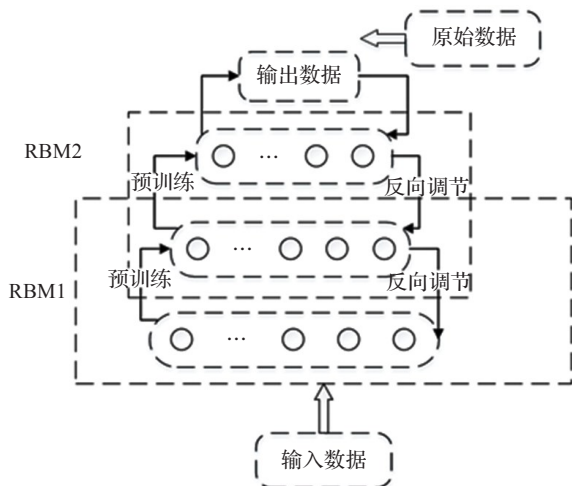


图1 深度信念网络示意图

文献[30]利用深度信念网络的非线性函数拟合能力,提出一种基于深度信念网络的网络流量预测模型。该预测模型由RBN组成,采用逐层无监督贪心算法训练参数,然后利用反向传播算法微调整个网络参数。最后基于该预测模型对收集到的真实网络流量进行预测和分析,实现了利用过去时刻数据对未来时刻的网络流量进行预测的功能。

深度信念网络可进行无监督学习,也可进行有监督学习。它通过“逐层初始化”的方式进行训练,有效地克服了深层网络结构在训练时遇到的梯度不稳定等问题<sup>[31]</sup>。与传统的神经网络相比,深度信念网络有训练时间短、不易陷入局部最优、处理数据快等优势。但深度信念网络的超参数需要人为设置,实验效率低。深度信念网络适合样本数量较小的非线性时间序列预测问题。

#### 3.2 卷积神经网络

在深度学习领域中,卷积神经网络(Convolutional Neural Networks, CNN)是重要且常见的一种模型,起初是为识别图片而设计,因而在处理图片类二维数据中具有较大优势。不仅如此,卷积神经网络利用权值共享减少了需要训练的权值个数,降低了网络的计算复杂度,同时通过池化操作减少了特征网络,提升了卷积神经网络的泛化能力<sup>[32]</sup>。卷积神经网络可以自动学习特征,能够避免手工提取特征造成的人为误差。典型的卷积神经网络主要由输入层、卷积层、池化层、全连接层及输出层组成<sup>[33]</sup>,具体结构如图2所示。

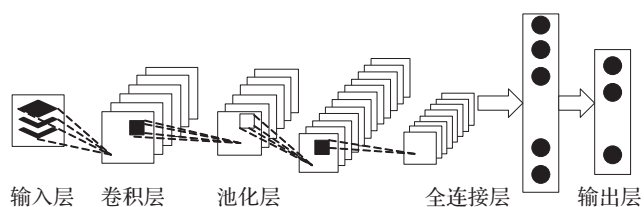


图2 卷积神经网络示意图

针对现有预测方法不能同时捕捉无线业务流量的复杂时空依赖关系问题,文献[16]提出了一种基于深度卷积神经网络的城市尺度无线业务流量预测模型。模型通过逐层卷积操作将小区间的局部依赖和全局依赖进行高效捕捉,通过两个共享参数的深度网络分别对时间维度的近邻依赖性和周期依赖性进行建模。此外,为了强化卷积层所学特征在不同网络层的传播,充分进行特征重用,模型采用全连接的模式对特征进行学习。实验结果表明,该模型能够大幅提升预测准确性。

卷积神经网络可以通过卷积操作来捕获数据的局部空间特征,但需要大量样本,计算复杂度较高,冗余性高<sup>[34]</sup>。卷积神经网络适合样本数量较大,对预测精度要求较高且时延不敏感的非线性时间序列预测,多用于预测具备空间特性的网络流量数据。

#### 3.3 循环神经网络和长短期记忆网络

循环神经网络(Recurrent Neural Networks, RNN)是一种反馈神经网络。Hopfield在1982年提出的Hopfield神经网络被公认为循环神经网络的始祖。循环神经网络在训练时很容易出现梯度消失和梯度爆炸的问题。梯度消失时,网络很难学习到远距离数据之间的依赖关系。梯度爆炸时,权重更新值繁杂冗余,导致网络更新困难。其中梯度爆炸问题可以通过裁剪梯度值等方法来解决<sup>[35]</sup>,但是梯度消失问题很难解决,因此使得循环神经网络的发展受到了严重阻碍。

为了解决循环神经网络的梯度消失问题,Hochreiter和Schmidhuber在1997年提出了长短期记忆网络(LSTM)<sup>[36]</sup>,其示意图如图3所示。该网络模型通过对循环神经网络的内部结构进行改进来克服梯度消失问题,推动了循环神经网络的发展。

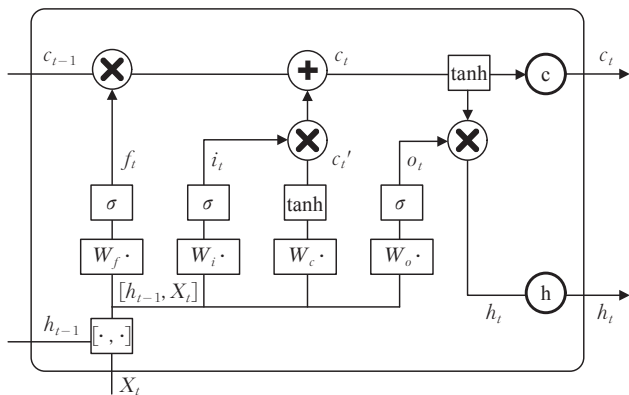


图3 长短时记忆网络示意图

文献[37]使用循环神经网络和长短时记忆网络等来对真实的骨干网数据进行预测,以均方根相对误差(RMSE)为评价指标,普通RNN的RMSE为0.067,而LSTM的RMSE为0.042,可见LSTM的预测能力要优于RNN。

为了建立一个响应时间短的网络流量预测模型,文献[38]使用长短时记忆网络进行全网络链路级流量预测。实验结果表明,LSTM高精度地预测链路吞吐量,以平均相对误差(Mean Absolute Percentage Error,MAPE)为评价指标,与自回归合成移动平均模型(ARIMA)模型相比,其性能提升约30%。

为了获取网络流量的长期相关性,张松等人分析了长短期记忆单元和门控循环单元的原理,提出了层叠式的双单向最小门控循环单元网络模型(SBU-M-GRUs)用于网络流量的预测。它的隐藏单元是最新的门控循环单元,只保留了更新门,同时控制历史信息的遗忘和新信息的添加,可以保持长期记忆。模型内部结构简单,训练速度快,并且在私有互联网服务提供商的网络流量数据集上进行验证,实验得到SBU-M-GRUs的MAE为0.034 9, RMSE为0.037 7,预测性能提升了20%<sup>[39]</sup>。

针对网络流量的安全性分析和校园网络流量的非线性和多维动态性引起的预测问题<sup>[40]</sup>,文献[41]设计了一种基于长短时记忆网络模型的网络流量预测系统,用于分析校园用户的网络行为。该预测系统使用西嘉教育日志分析工具收集和预处理来自各种网络应用的多源异构日志数据,并利用LSTM结构中的遗忘门和输入门来预测校园用户的网络流量。

长短时记忆网络很擅长对时间序列进行建模,能较好地处理大时间尺度的数据,但其在训练过程中收敛速度较慢,参数无法直接确定,易陷入局部最优。长短时记忆网络适合处理样本数量多且具有长期依赖性的时间序列预测问题。

3.4 改进型神经网络

近几年,伴随着我国不断地完善网络基础设施以及持续的社会数字化发展。网络流量数据呈现非结构化、

非线性化、实时性、突发性和时空性等特点,单一的神经网络结构难以满足实际网络流量预测需求<sup>[42]</sup>。因此,本领域的研究人员开始尝试对不同的神经网络结构进行修改来提高预测的精度。下面分别介绍几种典型的改进方案。

(1)实时性

对网络流量进行实时检测,及时获取流量变化趋势,分析各类网络资源的请求,可以帮助网络运营者进行动态智能化服务部署,提高资源利用率,同时可以有效检测网络故障,优化网络性能,确保网络安全。由于网络流量呈现出非结构性和非线性化,且大规模的网络流量数据量较大,很难实现实时预测。

文献[43]针对大规模网络中流量矩阵难以实时预测的问题,提出了基于卷积神经网络、深度信念网络和循环神经网络的流量预测模型,并从整体预测流量矩阵、预测每个起点—终点的流量和预测矫正过的流量矩阵等角度来预测流量。实验结果表明,循环神经网络在预测矫正过的流量矩阵时可以获得不错的实时预测效果。

文献[43]对于矫正过的网络流量有不错的预测效果,但是流量数据矫正需要预训练,难以满足下一代大规模动态网络的实时性需求。针对此问题,文献[44]在卷积神经网络的基础上提出了一种新的时空价值网络辅助深度学习的智能流量控制算法(ST-DeLTA)。首先将完整的网络流量数据按时间顺序划分为单独节点数据,每个节点数据包含流量负载量、剩余缓冲区数据量和链接通道的状态等;其次利用ST-DeLTA模型中的卷积操作从值矩阵中提取网络流量的时空特征;最后ST-DeLTA模型通过预测每个节点未来时刻的流量负载,做出自适应数据包转发决策,提升整条网络的负载能力。实验表明,ST-DeLTA模型在预测传输吞吐量和平均丢包率两方面上表现良好。

训练好的ST-DeLTA模型,在本网络结构中预测流量效果不错,且识别速度很快,满足网络流量实时性要求,但在未训练的网络结构中,泛化性不强。针对网络流量预测模型的泛化能力较差的问题<sup>[45]</sup>,李校林等人提出一种基于长短时记忆网络的网络流量预测模型。该模型的核心思想是首先通过距离比较和优化组合策略对粒子滤波算法的重采样过程进行优化改进,之后利用PE算法来替代原LSTM的训练过程,以提升模型训练速度,避免传统LSTM算法无法达到全局最优的缺陷。实验表明,与传统的LSTM算法相比较,PF-LSTM模型的预测精度及收敛速度都有所提升,能够更精确地实时描述网络流量的变化趋势<sup>[46]</sup>。

以上模型在同一数据空间进行预测,即网络流量数据结构相似。但在实际大型网络中业务量的时间和空间自相关,因此在灵活多变的光网络中进行准确的业务



量预测具有挑战性。时空图建模是分析系统流量的空间关系和时间趋势的有效方法,文献[47]提出了一种高效的基于图的神经网络,称为带有门控循环单元的图卷积网络(GCN-GRU)。基于现实世界的光网络流量数据集,GCN-GRU实现了98%的流量预测准确性,实现了光网络实施智能管理和实时自适应调整。

### (2) 突发性

由于网络流量具有突发性的明显特征,其时间序列呈现非线性,传统的线性方法很难准确地进行预测。针对此问题,文献[48]提出用小波变换的方法将原始流量分解为一个近似序列和几个细节序列。在此基础上,通过长短时记忆网络来预测流量变化趋势,并以多尺度提取突发信息,以完成对未来流量的预测。实验结果表明,相比于长短时记忆网络直接预测流量数据,长短时记忆网络针对小波变换后的流量数据的预测误差更低,性能更佳。

为应对网络流量的突发性和不确定性,文献[48]的思路是对流量数据进行处理后预测,而文献[49]的思路是不改变流量数据仅改变预测模型结构使其适用于预测突发流量,故其修改了传统长短时记忆网络的损失函数,提出了一种基于长短期记忆神经网络的网络流量预测模型。实验结果表明,与支持向量回归和反向传播神经网络构建的其他模型相比,该模型的预测精度更好,鲁棒性更强。

黄林等人继承了文献[49]的思路,通过修改网络结构模型来预测突发的网络流量,提出了一种并行的双LSTM预测器结构。其中一个作为主预测器,另一个用于突变时刻检测流量,两个预测器之间进行内部状态的交换,使主预测器可以利用突变检测器得到的信息进行多变量预测,从而改进预测的准确度。实验表明,相较于直接使用传统的单变量单LSTM对流量进行预测,该方法对突发业务流预测的准确率提高了30%~45%<sup>[50]</sup>。

黄林等人是将两个LSTM模型加以结合使用,而文献[51]提出的模型是将深度递归神经网络(RNN)和门控递归单元(Gated Recurrent Unit, GRU)神经网络相结合,应用于网络流量预测。数值计算和仿真验证结果表明,该模型的网络流量预测结果接近实际环境中网络流量的实际值,应用效果得到明显提升。

上述文献均为应对短期突发流量而设计,而文献[52]具体地将突发事件划分为高原、单发和双发三种类型,Plateau代表处于稳定流量下的网络,Single-Burst代表网络经历短期高峰流量后下降,趋于稳定,而Double-Burst代表网络经历短期高峰后又经历一个意外网络高峰后下降,并趋于稳定。针对这三种类型,文献[52]设计了一种非线性GCN-GAN网络模型,其主要思路是将图卷积神经网络和生成对抗网络加以结合,图神经网络

用于提取网络流量特征,生成对抗网络用于对数据结构建模,以实现突发流量的预测。通过将其模型与基本LSTM进行比较,证明了其预测实际光网络中突发事件的有效性。

### (3) 时空性

随着无线网络的飞速发展,基站的自我管理和主动调整能力变得至关重要。无线网络流量的准确预测是智能基站的重要前提。移动网络数据具有高度的非线性和复杂性,其特征存在于时间和空间相关性。现有的大多数预测方法在对流量数据进行建模的过程中都没有同时考虑时空情况。

文献[53]提出了一种时空卷积网络(LA-ResNet),该网络使用一种注意力机制来解决时空建模并预测无线网络流量。LA-ResNet由三部分组成:残差网络、递归神经网络和注意力机制。使用此方法,可以对无线网络流量数据的时空特征进行建模,并增强其相关功能,因此可以有效地捕获无线网络业务数据的时空相关性。残差网络可以捕获数据中的空间特征。递归神经网络和注意力机制的组合可以捕获数据的时间依赖性。最后,在真实数据集上的实验表明,LA-ResNet模型的预测效果得到验证。

文献[53]的预测效果明显提升,但是该模型结构比较复杂,计算复杂度较大。因为移动网络流量数据复杂,简单的模型难以更好地预测数据,对于长短时记忆网络而言,超参数越多模型越复杂。为寻找到模型复杂度与预测性能的平衡,文献[54]提出了一种基于遗传算法(Genetic Algorithm, GA)的LSTM方法,即GA-LSTM,用于预测网络流量。其利用LSTM提取时间流量特征,用GA为LSTM网络识别合适的超参数,最后建立GA-LSTM的网络流量预测模型。实验结果表明,与自回归综合移动平均值(ARIMA)和纯LSTM相比,提出的GA-LSTM模型复杂度最低,具有更高的预测精度和较小的预测误差,并且能够描述复杂变化的流量特征。

卷积神经网络适用于对空间数据的特征学习,故张杰等提出一种基于时空特征的移动网络流量预测模型STFM。STFM模型利用目标区域及周围区域的历史移动网络流量对目标区域的流量进行预测。其核心思想是,首先利用三维卷积网络(3DCNN)从流量中提取移动网络流量空间上的特征,再利用时间卷积网络(Temporal Convolutional Network, TCN)提取移动网络流量时间上的特征,最后全连接层对提取的特征与实际的流量值建立映射关系,产生预测的流量值。实验表明,STFM在移动网络流量预测上的RMSE相比TCN模型、CNN模型和CNN-LSTM模型分别减少了28.0%、21.7%和10.0%<sup>[55]</sup>。

表2总结了流量预测模型的特性及适用场景。

表2 预测模型特性整理

	方法	优点	缺点	适用场景	
线性预测模型	AR、MA、ARMA、ARIMA	模型简单,计算复杂度低,处理速度快	预测精度低,不能反映真实网络业务的突发性和长相关特性	适合具有短相关特性的业务预测	
	FARIMA	能够有效描述流量的长相关性,也能表示短相关流量	模型复杂,参数多,计算量大;实际应用局限性大	适合直接对数据曲线进行拟合,来进行变化趋势和特征的预测	
	SVM	泛化能力强,所需样本数少,不易陷入局部最优解	模型的选择需要凭借经验和大范围搜索,可靠性差	适合样本数量较小的非线性时间序列的预测	
神经网络	CNN	可以提取数据的局部空间特征	需要大样本量,计算复杂度较高,冗余性高,无记忆特性	适合样本数量较大,对预测精度要求较高且时延不敏感的非线性时间序列预测	
	RNN	可灵活捕获数据的时间依赖性	会出现梯度消失或梯度爆炸的问题	适合动态的短相关时间序列预测	
	LSTM	能较好地处理大时间尺度数据的影响	收敛速度较慢,参数无法直接确定,易陷入局部最优	适合样本数量多且具有长期依赖性的时间序列预测	
	DBN	训练速度更快,能够提取有用的数据特征,泛化能力强,可避免过拟合问题	计算复杂度较高	适合样本数量较小的非线性时间序列预测	
非线性预测模型					
	实时性	ST-DeLTA	可以预测动态变化的网络流量	模型复杂,参数多,计算量大	适用于大规模动态网络的数据包转发决策
		PE+LSTM	提高训练效率,克服收敛局部最优问题	训练流程相对复杂	适用于复杂突变的网络流量时序数据
		GCN-GRU	可以实现对灵活多变的光网络的实时自适应调整,响应快	模型复杂,参数多,计算量大	适用于大规模动态空间网络流量预测
	改进型突发性神经网络	小波+LSTM	模型简单,响应迅速,预测准确	经由小波变化后,有丢失关键信息的风险	适用于小型突发网络流量预测
		修改损失函数的LSTM	可以处理多源异构数据	针对特定用户的网络情况(如校园网络)	适用于多源异构数据的短期预测
		并行双LSTM	双LSTM对突发业务流预测的准确率较单LSTM强	计算复杂度较高	适合对突发业务流量预测
		GRU-RNN	预测流程简单,训练好的模型预测响应迅速	训练过程相对较慢	适合对大规模突发业务流量预测
		GCN-GAN	可以预测高原、单发和双发三种类型的流量	模型复杂,计算复杂度高	适合对大规模突发业务流量预测,且可以预测三种类型的流量
	时空关系	LA-ResNet	可以对高度复杂的非线性数据进行准确建模	响应较慢,对突发流量预测效果差	适用于非线性时空关系的流量预测
GA-LSTM		自动寻找最佳超参数	神经网络黑盒效应,人为无法控制超参数	适用于非线性网络流量时间序列数据	
STFM		适合对移动网络流量的时间和空间特征进行预测	只对短期流量有效	对短期移动网络流量进行预测	

4 未来研究方向和总结

在过去的十年中,研究人员设计了很多基于深度学习的网络流量预测算法,取得了较大的进展。本文研究了网络流量的评价指标,介绍了网络流量预测公开数据集及应用,并深入分析了基于深度学习的网络流量预测方法以及近年来出现的改进型神经网络预测方法。

就目前的研究成果来讲,网络流量预测从传统的方法到应用深度学习技术取得了较大的突破,但当前网络流量预测研究面临公开数据集不多,流量数据具有隐私性,公开受限等问题。研究人员将深度学习算法各自应用于不同的私有数据集上,评价指标不统一,难以客观衡量深度学习算法性能。

为了解决以上问题,未来需要建立统一的公开数据集和评价指标来评估不同预测模型的性能。同时,为了适应目前5G网络快速发展下应用需求和商业模式的变化,对网络流量预测未来研究点进行展望:

(1)高速网络下的流量预测。目前从工业界到民用带宽都在进行升级,随着网络带宽的增加,如何在高速网络环境下快速预测流量变化趋势将是互联网流量预测的一个重要研究方向。

(2)云边智能协同预测。随着未来5G网络的发展,小微基站、边缘节点的应用使得流量的管理与资源的配置可以更加精细化和个性化,结合边缘设备的云边协同智能化预测也是未来的一个研究方向,可服务于智能资

源调度、热点内容缓存、计算卸载等行业应用。

(3)应用级网络流量预测。在更靠近用户的网络边缘位置进行流量预测能够及时获取应用级的流量变化趋势,分析各类网络资源的请求,实现精细化资源编排和管理,进行动态智能化服务部署,能够提高资源利用率,节约能耗并为用户提供定制化服务。这将是未来网络流量预测的重要研究领域。

### 参考文献:

- [1] 韩莹,井元伟,金建宇,等.基于改进黑洞算法优化ESN的网络流量短期预测[J].东北大学学报(自然科学版),2018,39(3):311-315.
- [2] 姜明,吴春明,张旻,等.网络流量预测中的时间序列模型比较研究[J].电子学报,2009,37(11):2353-2358.
- [3] 吴雨聪.基于神经网络的网络流量预测研究[D].成都:电子科技大学,2020.
- [4] GAN M, PENG H. Stability analysis of RBF network-based state-dependent autoregressive model for nonlinear time series[J]. Applied Soft Computing, 2012, 12(1): 174-181.
- [5] XU S N, ZENG B Q. Network traffic prediction model based on auto-regressive moving average[J]. Journal of Networks, 2014, 9(3): 653-659.
- [6] 余德良.基于深度学习的移动网络流量态势预测[D].哈尔滨:哈尔滨工业大学,2019.
- [7] CELENK M, CONLEY T, GRAHAM J, et al. Anomaly prediction in network traffic using adaptive Wiener filtering and ARMA modeling[C]//2008 IEEE International Conference on Systems, Man and Cybernetics, 2008.
- [8] 刘豪.基于机器学习的校园网流量准确预测[D].北京:北京邮电大学,2019.
- [9] 王思远.深度学习在移动业务预测中的研究与应用[D].北京:北京邮电大学,2018.
- [10] 龙图景,孙政顺,李春文,等.一种新的网络业务流的多重分形小波模型[J].计算机学报,2004,27(8):1074-1082.
- [11] LEDERER C, ALTSTADT S, ANDRIAMONJE S, et al. Network traffic prediction based on FARIMA-GARCH model[C]//2013 International Conference on Insulating Materials, Material Application and Electrical Engineering, Changsha, 2013.
- [12] LIU X W, FANG X M, QIN Z H, et al. A short-term forecasting algorithm for network traffic based on chaos theory and SVM[J]. Systems and Network Management Journal, 2011, 19(4): 427-447.
- [13] 李丽.基于支持向量机的网络流量预测研究[D].南京:南京理工大学,2007.
- [14] 蒋品.基于机器学习的蜂窝网络基站流量分析与预测研究[D].北京:北京邮电大学,2019.
- [15] 黎佳玥,赵波,李想,等.基于深度学习的网络流量异常预测方法[J].计算机工程与应用,2020,56(6):39-50.
- [16] 张传亭.基于深度学习的城市尺度无线流量预测[D].济南:山东大学,2019.
- [17] REIS J, ROCHA M, PHAN T K, et al. Deep neural networks for network routing[C]//International Joint Conference on Neural Networks, 2019.
- [18] 殷荣网.基于FCM-LSSVM网络流量预测模型[J].计算机工程与应用,2016,52(1):105-109.
- [19] 曹建华.基于灰色神经网络模型的网络流量预测算法研究[D].江苏无锡:江南大学,2008.
- [20] 张晗,王霞.基于小波分解的网络流量时间序列建模与预测[J].计算机应用研究,2012,29(8):3134-3136.
- [21] METTU N R, SASIKALA T. Prediction analysis on web traffic data using time series modeling, RNN and ensemble techniques[C]//International Conference on Intelligent Data Communication Technologies and Internet of Things, 2018.
- [22] PETLURI N, AL-MASRI E. Web traffic prediction of Wikipedia pages[C]//2018 IEEE International Conference on Big Data, 2018: 5427-5429.
- [23] PAUN K P, MAKWANA C H. A survey on web traffic forecasting on time series data[J]. Journal of Applied Science and Computations, 2019, 6(4): 3588-3594.
- [24] ZSEBY T, VAZQUEZ F I, KING A, et al. Teaching network security with IP darkspace data[J]. IEEE Transactions on Education, 2016, 59(1): 1-7.
- [25] LAN X. Analysis and research of several network traffic prediction models[C]//2013 Chinese Automation Congress, Changsha, 2013.
- [26] 黎嘉诚.基于小波神经网络的校园网络流量预测研究[D].南昌:南昌大学,2019.
- [27] HINTON G E, OSINDERO S, TEH Y W. A fast learning algorithm for deep belief nets[J]. Neural Computation, 2006, 18(7): 1527-1554.
- [28] HUANG W, SONG G, HONG H. Deep architecture for traffic flow prediction: deep belief networks with multi-task learning[J]. IEEE Transactions on Intelligent Transportation Systems, 2014, 15(5): 2191-2201.
- [29] NIE L, JIANG D, YU S, et al. Network traffic prediction based on deep belief network in wireless mesh backbone networks[C]//Wireless Communications and Networking Conference, San Francisco, 2017.
- [30] 任玮.基于深度信念网络的网络流量预测模型[J].山西电子技术,2016,2(1):62-64.
- [31] NAREJO S, PASERO E. An application of Internet traffic prediction with deep neural network[M]//Multidisciplinary approaches to neural computing. Cham: Springer, 2017: 139-149.
- [32] 张庆辉,万晨霞.卷积神经网络综述[J].中原工学院学报,2017,28(3):82-86.
- [33] 曹钰.基于区域信息的深度卷积神经网络研究综述[J].电子世界,2017,41(6):32-36.



- [34] 杜爽,徐展琦,马涛,等.基于神经网络模型的网络流量预测综述[J].无线电通信技术,2020,46(2):216-222.
- [35] BENGIO Y, SIMARD P, FRASCONI P. Learning long-term dependencies with gradient descent is difficult[J]. IEEE Transactions on Neural Networks, 1994, 5(2): 157-166.
- [36] HOCHREITER S, SCHMIDHUBER J. Long short-term memory[J]. Neural Computation, 1997, 9(8): 1735-1780.
- [37] 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, Udipi, 2017.
- [38] LAZARIS A, PRASANNA V K. Deep learning models for aggregated network traffic prediction[C]//2019 15th International Conference on Network and Service Management, Halifax, 2019.
- [39] 张松. 基于改进门控循环神经网络的网络流量预测研究[D]. 长沙: 湖南大学, 2019.
- [40] 曹淑琳. 基于深度学习的无线网络流量预测[D]. 西安: 西安电子科技大学, 2019.
- [41] GENG Y, LI S A. LSTM based campus network traffic prediction system[C]//2019 IEEE 10th International Conference on Software Engineering and Service Science, Beijing, 2019.
- [42] 曹国昕. 网络流量分析与预测模型研究[D]. 长春: 吉林大学计算机科学与技术学院, 2009.
- [43] LIU Z, WANG Z, YIN X, et al. Traffic matrix prediction based on deep learning for dynamic traffic engineering[C]//2019 IEEE Symposium on Computers and Communications, Barcelona, 2019.
- [44] TANG F, MAO B, FADLULLAH Z M, et al. ST-DeLTA: an novel spatial-temporal value network aided deep learning based intelligent network traffic control system[J]. IEEE Transactions on Sustainable Computing, 2020, 5(4): 568-580.
- [45] 陈曦. 基于包容性检验和神经网络的网络流量预测[J]. 电视技术, 2014, 38(11): 130-133.
- [46] 李校林, 吴腾. 基于 PF-LSTM 网络的高效网络流量预测方法[J]. 计算机应用研究, 2019, 36(12): 3833-3836.
- [47] GUI Y, WANG D, GUAN L, et al. Optical network traffic prediction based on graph convolutional neural networks[C]//2020 Opto-Electronics and Communications Conference, Taipei, China, 2020.
- [48] LU H, YANG F. A network traffic prediction model based on wavelet transformation and LSTM network[C]//9th IEEE International Conference on Software Engineering and Service Science, Beijing, 2018.
- [49] LU H, YANG F. Research on network traffic prediction based on long short-term memory neural network[C]//2018 IEEE 4th International Conference on Computer and Communications, Chengdu, 2018.
- [50] 黄林, 王电钢, 刘萧, 等. 基于 LSTM 的网络流量预测方法[J]. 计算机应用研究, 2020, 37(A1): 264-265.
- [51] FAN J, MU D, LIU Y. Research on network traffic prediction model based on neural network[C]//2019 2nd International Conference on Information Systems and Computer Aided Education, Dalian, 2019.
- [52] VINCHOFF C, CHUNG N, GORDON T, et al. Traffic prediction in optical networks using graph convolutional generative adversarial networks[C]//2020 22nd International Conference on Transparent Optical Networks, Bari, 2020.
- [53] LI M, WANG Y, WANG Z, et al. A deep learning method based on an attention mechanism for wireless network traffic prediction[J]. Ad Hoc Networks, 2020, 107: 102258.
- [54] CHEN J, XING H, YANG H, et al. Network traffic prediction based on LSTM networks with genetic algorithm[C]//2019 International Conference on Signal and Information Processing, Networking and Computers, 2019.
- [55] 张杰, 白光伟, 沙鑫磊, 等. 基于时空特征的移动网络流量预测模型[J]. 计算机科学, 2019, 46(12): 108-113.