

基于大数据和AI的5G定制网 智能运维方案研究与实现

Research and Implementation of Intelligent Operation and Maintenance
Solution for 5G Customized Network Based on Big Data and AI

李 伟¹,王旭峰¹,王 刚¹,邓子仪¹,田 琪¹,马宝泽²(1. 中国电信贵州分公司,贵州 贵阳 550001;2. 重庆邮电大学光电工程学院,重庆 400065)

Li Wei¹,Wang Xufeng¹,Wang Gang¹,Deng Ziyi¹,Tian Qi¹,Ma Baoze²(1. China Telecom Guizhou Branch,Guiyang 550001,China; 2. School of Optoelectronic Engineering,Chongqing University of Posts and Telecommunications,Chongqing 400065,China)

摘 要:

为提升运维效率,设计了一种基于大数据和AI技术的5G定制网智能运维方案,提出了5G定制网长短期记忆(LSTM-5G CN)故障预测模型。该方案利用网络运行过程中产生的核心数据,通过数据处理、LSTM模型训练和隐患预测、数据可视化等关键技术,实现对网络设备的监控、巡检、故障定位和故障预判。仿真结果表明,在多时间粒度网络性能指标条件下,所提出的LSTM-5G CN模型预测精度保持在98.99%以上。

Abstract:

To improve operation and maintenance efficiency, a 5G customized network intelligent operation and maintenance solution based on big data and AI technology is designed, and a long short term memory of 5G customized network (LSTM-5G CN) fault prediction model is proposed. This scheme utilizes core data generated during network operation, and achieves monitoring, inspection, fault localization, and fault prediction functions for network equipment through key technologies such as data processing, LSTM model training, hidden danger prediction, and data visualization. The simulation results show that under the condition of multi time granularity network performance indicators, the proposed LSTM-5G CN model maintains a prediction accuracy of over 98.99%.

Keywords:

5G customized network; Big data; AI; LSTM; Hidden danger prediction

引用格式:李伟,王旭峰,王刚,等. 基于大数据和AI的5G定制网智能运维方案研究与实现[J]. 邮电设计技术,2026(1):48-54.

关键词:

5G定制网;大数据;AI;LSTM;隐患预测

doi:10.12045/j.issn.1007-3043.2026.01.011

文章编号:1007-3043(2026)01-0048-07

中图分类号:TN915

文献标识码:A

开放科学(资源服务)标识码(OSID):



0 引言

截至2022年底,中国电信已累计打造5G定制网项目超3 800个,5G DICT项目超1.3万个^[1],相关项目涉及工业、采矿、医疗、交通物流、教育等15个行业,与数百个大型企业开展了5G定制网合作。当前,5G定制网行业用户规模已进入持续增长的稳健阶段,如何利用大数据和AI技术确保复杂网络的安全稳定性及

故障处理的时效性,是提升2B行业用户感知的关键。

1 现状和解决思路

1.1 现状

5G定制网是一种通过按需部署5G资源并提供定制化服务的端到端网络,具备低时延、大带宽、高安全性等优势。目前,5G定制网有致远、比邻、如翼3种模式,行业客户主要以比邻模式为主,即用户面UPF(user plane function)网元下沉到用户园区实现数据流量分流,通过N4接口实现专网和公网共享5G核心网控制面功能。比邻模式5G定制网网络架构如图1所

基金项目:国家自然科学基金项目(62201113)

收稿日期:2025-11-28

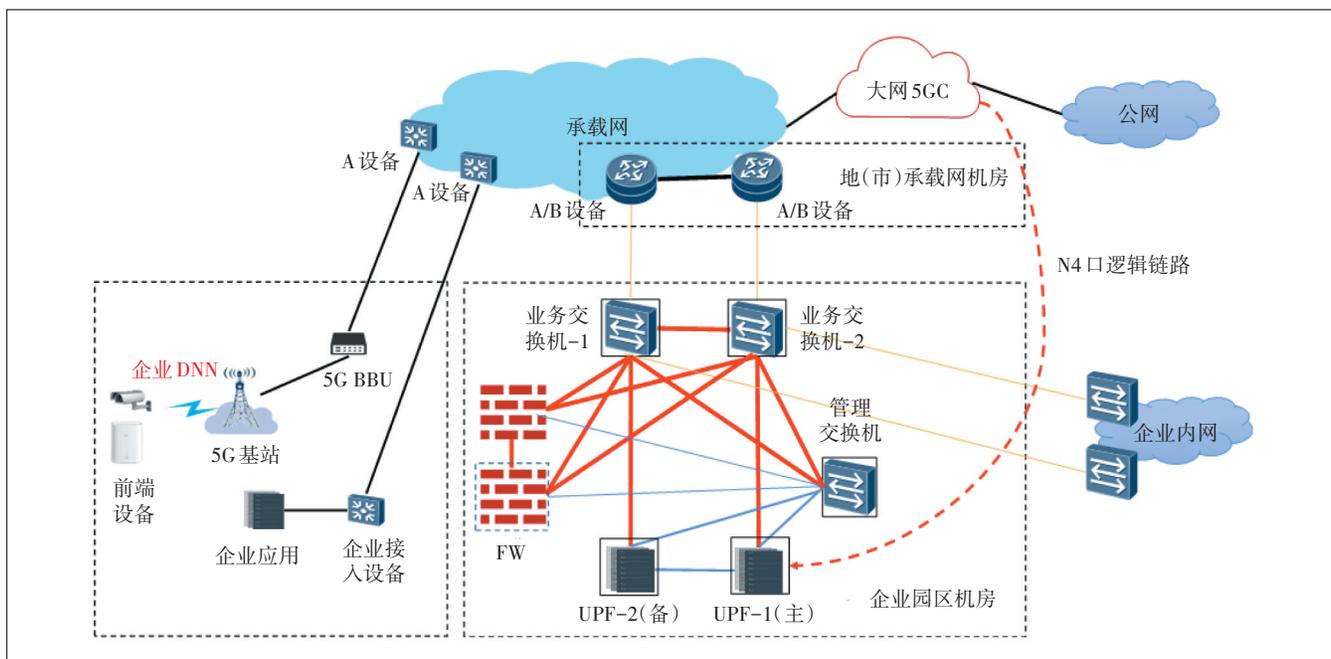


图1 比邻模式5G定制网网络架构

示,园区设备使用定制化DNN接入5G定制网网络,产生的数据流量通过下沉的UPF网元实现分流,园区内部流量可以直接流向企业内网或MEC服务器,公网流量通过N9接口流向公网UPF网元访问公网资源。由于5G定制网端到端网络的复杂性及设备终端的多样性,行业用户对网络稳定性的要求与日俱增,基于信令跟踪、指标观察、日志分析等传统运维巡检手段已无法满足行业客户需求,且传统网络维护手段对网络隐患的发现和定位存在滞后性。因此,本文设计了一种基于大数据和AI技术的5G定制网通用智能运维方案,利用5G定制网运行过程中产生的核心重要数据实现对网络设备的故障定位、隐患预测等功能,并对园区端到端5G定制网网络进行7×24 h的监控、业务巡检、故障诊断等。基于大数据和AI技术的5G定制网智能运维方案是一种通用运维解决方案,适用于致远、比邻、如翼3种5G定制网模式。由于如翼模式属于区域专属网络,它通过下沉5G核心网控制面网元满足高隔离、高安全的敏感型政企客户需求,为满足如翼模式的高安全性,该5G定制网智能运维方案需部署在客户园区内。

1.2 解决思路

针对5G定制网端到端的运行维护及故障处置,通过系统化梳理5G定制网重要设备清单和关键网络性能指标清单,自研了基于大数据和AI技术的5G定制

网智能运维方案(见图2)。

基于大数据和AI技术的5G定制网智能运维方案实现了系统巡检、定制网监控、故障预判、用户侧故障定位等功能。整个方案基于Diango架构进行开发,采用SSH、SNMP等协议,利用了Pandas、Scikit-learn、Bootstap5.0等开源软件开发了定制网监控模块、系统巡检模块、用户故障定位模块、故障预判模块四大基础功能组件,实现了对5G定制网网络的数据采集、系统巡检、网络监控以及故障定位;通过引入故障预测算法模型,实现对用户侧故障和网络设备故障的预判。

目前,故障预测算法模型主要分为机器学习类故障预测算法模型^[2]和时序分析类故障预测算法模型^[3]。机器学习类故障预测算法模型主要利用机器学习算法对异常点故障和数据的空间分布进行分析,预测是否存在故障,相关算法模型主要包括孤立森林算法^[4-5]、单类支持向量机算法^[6-7]等。时序分析类故障预测算法模型主要基于时间序列的性能指标数据,通过分析拟合误差的分布和置信区间阈值来预测是否存在故障,相关深度学习算法主要包括长短期记忆(Long Short-Term Memory, LSTM)模型及其变体^[8-9]、多模态动态卷积核(Multimodal Dynamic Convolutional Network, MTNet)^[10-11]等。由于LSTM模型将时间序列引入到网络结构中,其在复杂时序网络数据分析和预

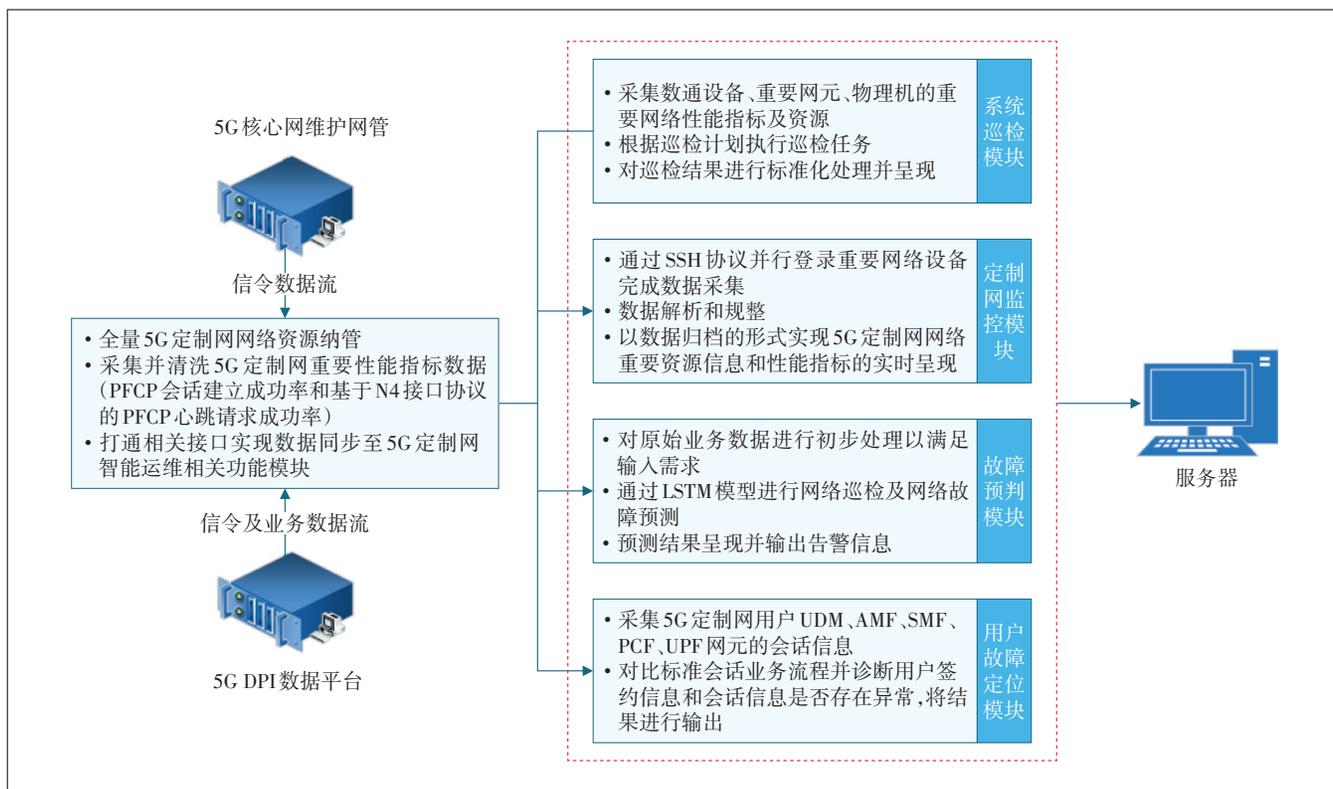


图2 5G定制网智能运维框架

测方面具有明显的优势^[12-14]。因此,在5G定制网故障预测中引入LSTM模型,以网络运行过程中产生的重要性能指标数据作为输入,通过算法训练,实现对5G定制网的故障预判。

2 方案实现与验证

2.1 定制网监控模块

定制网监控模块调用已纳管至数据库的网络设备全量资源信息,并通过SSH协议并行登录重要网络设备执行数据采集操作指令,利用Python的TextFSM库函数将采集的指令结果解析为基于Pandas的结构化数据,从而实现5G定制网全量网络设备数据的自采集。采集的重要数据会被写入消息队列并通过接口方式转发给监控模块,监控模块以数据归档的形式实现5G定制网网络设备重要资源信息和性能指标的实时呈现。定制网监控模块支持多维度的监控看板展示,运维人员可以通过生成临时看板,在短时间内查询任一定制网项目下某项或多项指标的历史记录。

2.2 系统巡检模块

系统巡检模块通过设计合理的巡检规划实现高效的目标巡检,整个巡检模块由4个部分组成。

a) 巡检指标。在数通设备及配套设施方面,巡检指标主要包括防火墙CPU和内存负荷、数通交换机设备的IP端口状态及端口流量、设备温度等信息;在重要核心网元方面,巡检指标主要包括重要网元License和地址池利用率;在物理机方面,巡检指标主要包括物理机网卡性能、物理机内存利用率和CPU、磁盘利用率等;在虚拟机方面,巡检指标主要包括虚拟机内存利用率和CPU、虚拟机网卡性能、磁盘利用率等。

b) 巡检计划。根据实际生产维护需要,制定立即执行任务和定时执行功能选项。

c) 巡检方案。针对巡检的网元和网络设备,构建算法进行数据采集并输出巡检结果。针对数量较少的网元和网络设备(如防火墙),串行与并行巡检方案均可。针对数量较多的设备(如配套业务交换机),为实现高效巡检,主要通过并行方案进行采集巡检。在巡检数据采集方面,本模块利用SSH协议到相关网元和设备执行指令,通过Python的TextFSM库函数将采集指令结果进行规整并输出Pandas结构数据。

d) 巡检结果。由于5G定制网巡检网元和设备众多且字段复杂,可以利用相关数据结构算法将巡检结果异常标准汇总,然后将巡检的各项指标进行全量呈

现。

2.3 用户故障定位模块

用户故障定位模块主要利用定制网用户的MSISDN或IMSI进行用户故障快速定位,并通过调用核心网维护网管和DPI接口来查看用户的全量日志。在用户数据会话信令流程中,该模块主要采集用户UDM、AMF、SMF、PCF、UPF网元的会话信息,诊断用户签约信息和会话信息是否存在异常,并通过PING功能测试网络是否正常。在用户会话数据采集方面,利用SSH协议并行到同类网元进行用户会话数据采集和数据清洗,对采集的信息进行二次判断,判断其是否存在用户信息,如果存在用户信息,则转换成字典类型数据进行输出呈现,从而实现用户故障定位。

2.4 故障预判模块

故障预判模块为5G定制网智能运维方案的核心模块。该模块以采集的全省5G定制网运行过程中产生的基于用户指标的PCFP会话建立成功率和基于N4接口协议的PCFP心跳请求成功率为原始数据,利用数据处理方法对5G定制网业务维度关键指标进行提炼与汇总,以汇总后的5G定制网业务维度关键指标作为输入,通过LSTM模型对输出的5G定制网业务质量

进行预测,并根据设定的告警门限值输出故障预警信息。根据PCFP会话建立成功率和PCFP心跳请求成功率的数据样本特征以及LSTM模型的设计原则,构建的LSTM故障预判模型如图3所示。

整个故障预判模块由5个部分组成,分别为输入层、隐藏层、输出层、网络训练和网络预测。其中,输入层负责对原始业务数据进行初步处理以满足输入需求;输出层提供预测结果;网络训练采用文献[15]提出的基于梯度的适应性动量估计(Adaptive moment estimation, Adam)算法;网络预测采用迭代的方法进行逐个指标的预测;核心的隐藏层采用基于LSTM模型搭建的前向单层循环神经网络。LSTM模型涉及输入门 i_t 、遗忘门 f_t 、单元状态 c_t 、输出门 o_t ,相关门函数如式(1)~式(5)所示。

$$i_t = \sigma(W_{xi}x_t + W_{hi}h_{t-1} + W_{ci}c_{t-1} + b_i) \quad (1)$$

$$f_t = \sigma(W_{xf}x_t + W_{hf}h_{t-1} + W_{cf}c_{t-1} + b_f) \quad (2)$$

$$c_t = f_t c_{t-1} + i_t \tanh(W_{hc}h_{t-1} + W_{xc}x_t + b_c) \quad (3)$$

$$o_t = \sigma(W_{co}c_t + W_{xo}x_t + W_{ho}h_{t-1} + b_o) \quad (4)$$

$$h_t = o_t \tanh(c_t) \quad (5)$$

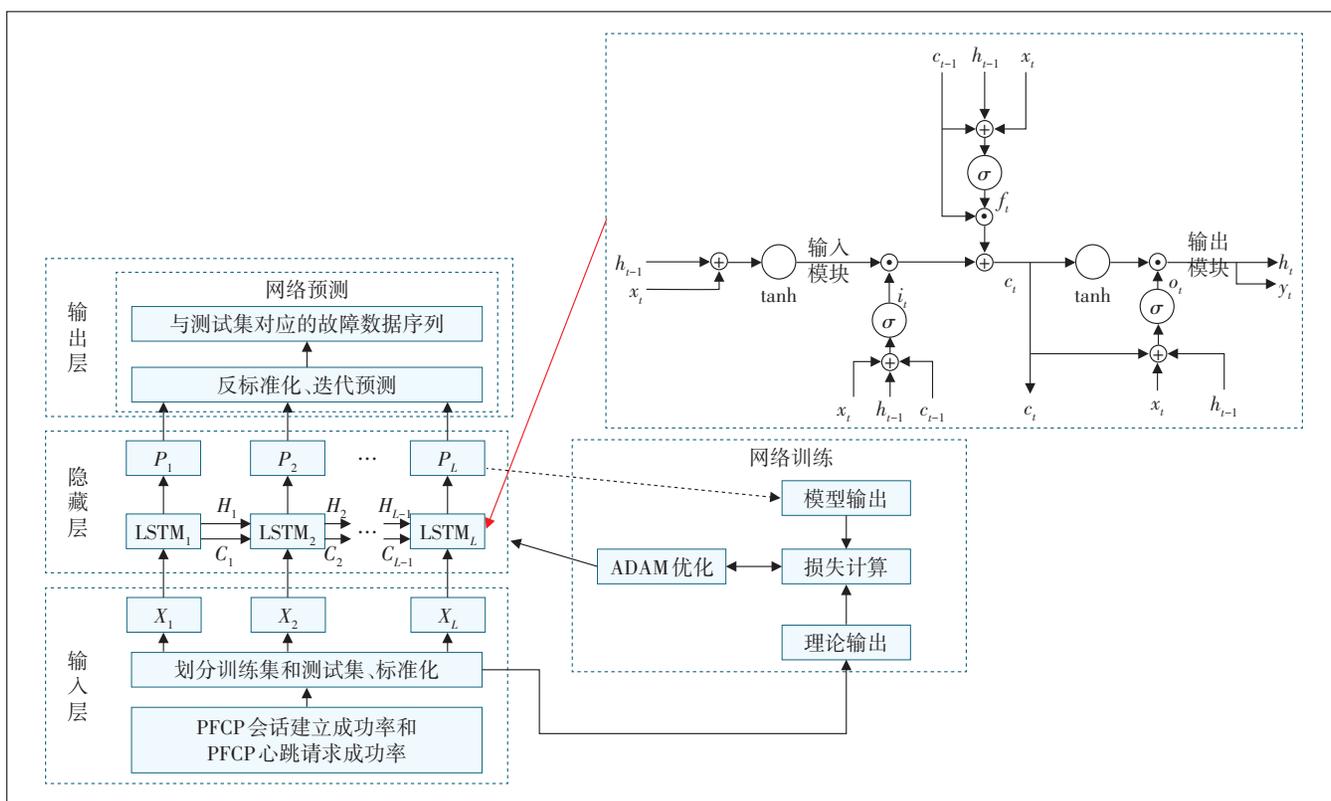


图3 LSTM故障预判模型

其中, σ 为 sigmoid 函数, W 为权重系数矩阵, b 为隐藏层的偏置向量, x_t 为输入向量, h_t 为输出向量, $\tanh(\cdot)$ 为双曲正切激活函数, 下标 t 表示时刻。

2.4.1 网络训练

网络训练主要在 LSTM 故障预判模型的隐藏层中进行。在输入层中, 定义原始的输入数据为 $F_{\text{Data}} = \{f_1, \dots, f_n\}$, 并将原始的网络性能指标数据划分为训练集 F_{Train} 和测试集 F_{Test} , 训练集与测试集分别表示为 $F_{\text{Train}} = \{f_1, \dots, f_m\}$ 和 $F_{\text{Test}} = \{f_{m+1}, \dots, f_n\}$, 其中, $m < n$ 且 $m, n \in N^*$ 。为消除输入训练数据的差异, 利用数据统计方法对训练数据 F_{Train} 进行标准化, 并定义标准化的训练数据为 F'_{Train} 。 F'_{Train} 和标准化方法如式(6)和式(7)所示。

$$F'_{\text{Train}} = \{f'_1, \dots, f'_m\} \quad (6)$$

$$f'_i = \frac{\left(f_i - \frac{1}{n} \sum_{i=1}^n f_i\right)}{\sqrt{\frac{1}{n} \sum_{i=1}^n \left(f_i - \frac{1}{n} \sum_{i=1}^n f_i\right)^2}}, 1 \leq i \leq m \quad (7)$$

将标准化后的训练集划分为与 LSTM 单元数量相等的 L 个模型输入, 相关模型输入和理论输出如式(8)和式(9)所示。

$$X = \{X_1, \dots, X_L\} \quad (8)$$

$$Y = \{Y_1, \dots, Y_L\} \quad (9)$$

其中, X_j, Y_j 分别表示为 $X_j = \{f'_j, \dots, f'_{m-L+j-1}\}$ 和 $Y_j = \{f'_{j+1}, \dots, f'_{m-L+j}\}$, $j \in N^*, 1 \leq j \leq L$ 。

根据图3所示的 LSTM 故障预判模型, X 通过隐藏层 LSTM 各单元计算后, 输出的 P 如式(10)所示。

$$P = \{P_1, \dots, P_L\} \quad (10)$$

其中, P_j 可以表示为 $P_j = \text{LSTM}(X_j, C_{j-1}, H_{j-1})$, C_{j-1} 和 H_{j-1} 分别表示隐藏层中第 $(j-1)$ 个 LSTM 单元的状态和输出, $\text{LSTM}(\cdot)$ 表示 LSTM 模型计算方法[式(1)~式(4)]。本文假设 LSTM 计算单元的状态向量大小为 S , 则 C_{j-1} 和 H_{j-1} 的向量大小也为 S 。根据对训练数据的建模方法, 可以得出 X, Y, P 均为 $(L, m-L)$ 的二维数组。因此, 可以定义训练过程的损失函数为式(11)。

$$\text{Loss} = \sum_{i=1}^{L \times (m-L)} (p_i - y_i)^2 / [L \times (m-L)] \quad (11)$$

为进一步得到隐藏层网络模型, 可以设定初始化随机种子数、训练步数、学习率, 并将损失函数 Loss 定

义为最小优化目标函数, 根据文献[15]中提到的 Adam 算法逐步更新隐藏层网络权重以得到最终的网络训练模型 $\text{LSTM}^*(\cdot)$ 。

2.4.2 网络预测

网络预测主要通过迭代的方法, 利用训练的 $\text{LSTM}^*(\cdot)$ 模型进行预测。根据图3所示的 LSTM 故障预判模型, 将 $(L, m-L)$ 的二维理论输出 $Y = \{Y_1, \dots, Y_L\}$ 中的最后一项元素 Y_L 输入 $\text{LSTM}^*(\cdot)$ 网络模型, 可以得到输出结果 $P_{Y_L}^* = \{P_{m-L+2}^*, \dots, P_{m+1}^*\}$, 则元素 P_{m+1}^* 对应为测试数据 $F_{\text{Test}} = \{f_{m+1}, \dots, f_n\}$ 中 f_{m+1} 的预测结果。为进一步得到所有测试数据的预测结果, 将 Y_L 中的后 $(L-1)$ 项元素与 P_{m+1}^* 合并为新的输入数据 $Y_{L+1} = \{f'_{L+2}, \dots, f'_m, P_{m+1}^*\}$, 将 Y_{L+1} 输入 $\text{LSTM}^*(\cdot)$ 可以得到对应 f_{m+2} 的预测结果 P_{m+2}^* 。同理, 通过迭代可以依次得到测试数据 F_{Test} 对应的预测结果 $P_{F_{\text{Test}}}^* = \{P_{m+1}^*, \dots, P_n^*\}$ 。为得到标准的预测结果, 根据2.4.1节的网络训练数据标准化方法, 对 $P_{F_{\text{Test}}}^*$ 进行反标准化, 即可得到标准的预测结果 $\hat{P}_{F_{\text{Test}}}^*, \hat{P}_{F_{\text{Test}}}$ 和反标准化计算方法如式(12)和式(13)所示。

$$\hat{P}_{F_{\text{Test}}}^* = \{\hat{P}_{m+1}^*, \dots, \hat{P}_n^*\} \quad (12)$$

$$\hat{P}_k^* = P_k^* \times \sqrt{\frac{1}{n} \sum_{i=1}^n \left(f_i - \frac{1}{n} \sum_{i=1}^n f_i\right)^2} + \frac{1}{n} \sum_{i=1}^n f_i \quad (13)$$

$m+1 \leq k \leq n, k \in N^*$

此外, 通过均方根误差 (Root Mean Square Error, RMSE) 方法计算训练的 $\text{LSTM}^*(\cdot)$ 模型的预测精度, RMSE 计算如式(14)所示。

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{n=1}^N (y_n - f_n)^2} \quad (14)$$

其中, y_n 为测试数据, f_n 为预测结果, N 为数据个数。

2.4.3 实验分析与部署实现

故障预判功能模块开发使用的服务端处理器为 13th Gen Intel(R) Core(TM) i7-13700H, 频率为 2.40 GHz; 内存为 32.00 GB; 集成开发环境为 PyCharm Community。程序开发涉及的 LSTM 模型和数据处理算法主要由 Python 相关的程序包实现 (Pandas, TextFSM, Tensorflow 等)。客户端通过调用 5G 维护网管和 5G DPI 平台相关数据完成预处理, 将处理后的数据输入故障预判功能模块生成故障预判结果, 最后通过 Http

接口协议完成应用层故障预判结果的呈现。

在5G定制网网络性能指标数据集选择方面,分别以5 min、15 min、30 min为时间粒度,提取5G定制网PFCP会话建立成功率和N4接口协议的PFCP心跳请求成功率作为故障时间序列数据集。根据2.4.1节提出的数据预处理方法对数据集进行标准化划分,并通过训练集优化LSTM预测模型。根据模型训练经验设定初始化模型参数,数据划分长度取最小值2,状态向量取故障次数12,随机种子数取值为1,训练步数取值为700。在训练过程中,通过设定不同的学习率 η 判断LSTM模型的损失变化, η 的值分别选取0.05、0.1及0.5,其变化关系如图4所示。从图4可以看出,当 $\eta=0.1$ 时损失较小。因此,LSTM训练模型关键参数设置如表1所示。

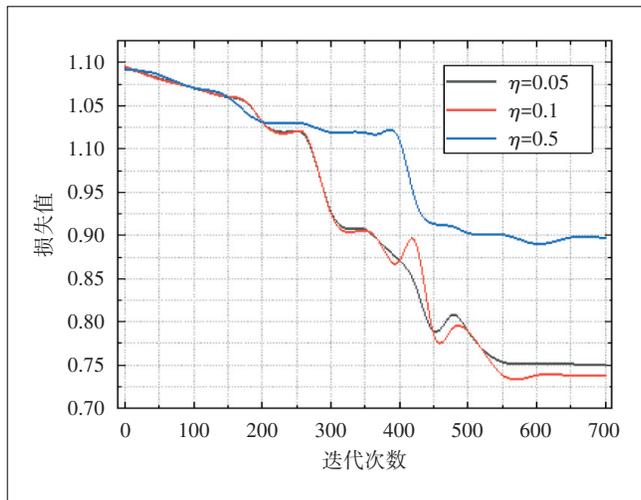


图4 损失变化关系

表1 LSTM模型关键参数

参数名称	取值
数据划分长度	2
状态向量	12
随机种子数	1
训练步数	700
学习率	0.1

根据2.4.2节的网络预测方案,分别以5 min、15 min、30 min为粒度对5G定制网性能指标数据进行预测,并选取均方根误差RMSE对5 min、15 min、30 min为粒度的期望指标的预测精度进行评估。从表2可以看出,在以5 min为粒度作为输入指标数据进行预测时,PFCP会话建立成功率及PFCP心跳请求成功率预

表2 RMSE均方误差评估

粒度	预测PFCP会话建立成功率期望/%	实际PFCP会话建立成功率期望/%	PFCP会话建立成功率RMSE	预测PFCP心跳请求成功率期望/%	实际PFCP心跳请求成功率期望/%	N4接口PFCP心跳请求成功率RMSE
5 min	99.995 1	99.997 5	0.003 2	99.998 9	99.998 6	0.035 3
15 min	99.994 9	99.997 6		99.997 6	99.998 2	
30 min	99.994 9	99.998 1		99.998 1	99.997 2	

测准确率较高,其主要原因是5 min粒度输入的数据量较大。此外,以5 min、15 min、30 min为粒度的PFCP会话建立成功率期望和N4接口PFCP心跳请求成功率期望的均方根误差RMSE分别为0.003 2、0.035 3,预测精度和预测结果满足评估要求。图5~图7分别展示了以5 min、15 min、30 min为粒度的5G定制网性能指标数据的全量预测效果,其PFCP会话建立成功率和N4接口PFCP心跳请求成功率预测指标分别在99.25%~100%、99.26%~100%、98.99%~100%之间波动。

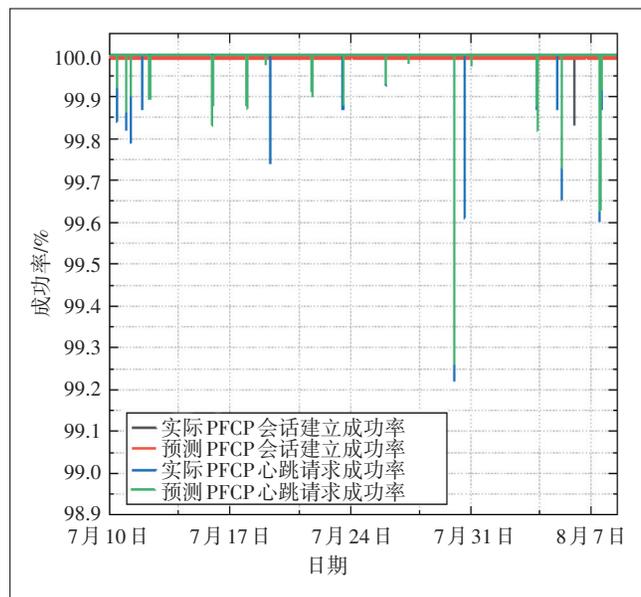


图5 会话建立成功率和心跳请求成功率变化关系预测效果 (5 min)

3 结束语

本文聚焦于5G定制网网络维护场景,通过参考智能化运维理念,引入大数据和AI算法为传统5G定制网网络维护赋能。该方案极大提升了5G定制网网络运维过程中的安全性和灵活性,具有很好的推广意义。同时,利用后端计算和前端呈现实现自动化运

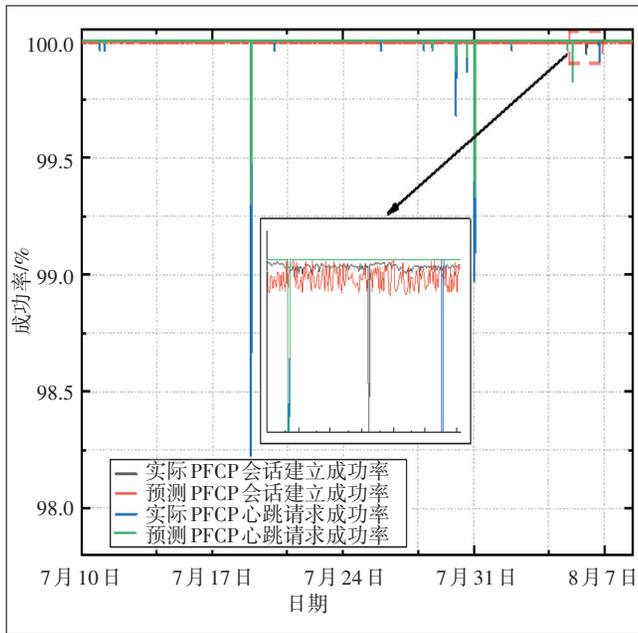


图6 会话建立成功率和心跳请求成功率变化关系预测效果 (15 min)

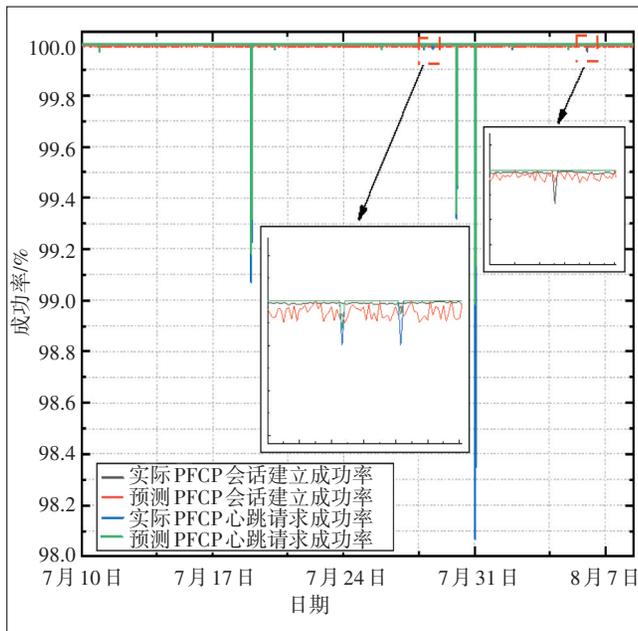


图7 会话建立成功率和心跳请求成功率变化关系预测效果 (30 min)

维,通过引入LSTM模型算法实现了故障预判并获得了理想的预测效果。本文提出的5G定制网智能运维方案同样适用于5G大网,但5G大网涉及更多的核心网网元以及更大的数据量。因此,为保障大网的稳定性和安全性,所提出的LSTM故障预判模型需考虑更多的数据指标因素,算法的有效性也需要进一步验证。

证。后续工作将聚焦大模型领域,利用大模型进一步提高故障预判性能,为5G行业客户设备互联、数据共享、数字孪生、自动化和智能化运维提供有力的基础条件。

参考文献:

- [1] 杜峰. 中国电信5G定制网“千案千面”助力行业个性转型[N]. 通信信息报,2022-11-30(3).
- [2] 余萍,曹洁. 深度学习在故障诊断与预测中的应用[J]. 计算机工程与应用,2020,56(3):1-18.
- [3] 张勉知,刘惜吾,叶晓斌,等. AI智能运维在5G SA网络中的应用研究[J]. 邮电设计技术,2020(10):47-50.
- [4] 宫晓婷,董培信. 基于改进孤立森林算法的交通流异常数据检测模型[J]. 重庆交通大学学报(自然科学版),2024,43(5):61-69,90.
- [5] 牛罡,王蕾,谭磊,等. 基于孤立森林算法的电力数据异常检测方法设计[J]. 自动化与仪器仪表,2023(8):52-55.
- [6] 武玉坤,李伟,倪敏雅,等. 单类支持向量机融合深度自编码器的异常检测模型[J]. 计算机科学,2022,49(3):144-151.
- [7] 边策. 基于单类支持向量机的拷贝数变异检测方法[D]. 西安:西安电子科技大学,2022.
- [8] 郑大庆,林陈威,王昺杰. 基于注意力-长短期记忆模型的偶发性交通流异常检测[J]. 同济大学学报(自然科学版),2023,51(6):923-931.
- [9] 李忠媛,鲜果,龚晓峰,等. 基于GAIN-LSTM网络的雷达PRI序列还原及识别方法[J]. 电讯技术,2024,64(3):343-349.
- [10] 张浩然. 基于多模态的深度假检检测算法研究[D]. 福州:福建师范大学,2023.
- [11] 黄小虎. 基于多模态特征的步态识别方法研究[D]. 武汉:华中科技大学,2023.
- [12] 王鑫,吴际,刘超,等. 基于LSTM循环神经网络的故障时间序列预测[J]. 北京航空航天大学学报,2018,44(4):772-784.
- [13] 班雪飞,倪峰,周玮,等. 基于LSTM神经网络算法互联网电视EPG业务隐患预测的应用研究[J]. 现代信息科技,2024,8(1):99-103.
- [14] 冯强,赵建光,杨草,等. 时间序列中非平稳性和波动性的建模及预测[JOL]. 计算机科学与探索,2024:1-10 [2025-08-12].<http://kns.cnki.net/kcms/detail/11.5602.TP.20240801.1607.006.html>.
- [15] KINGMA D P, BA J. Adam: a method for stochastic optimization [EB/OL]. [2025-08-12]. <https://arxiv.org/abs/1412.6980>.

作者简介:

李伟,工程师,硕士,主要从事5G核心网架构、IMS网络、大数据和AI技术的研究工作;王旭峰,工程师,学士,主要从事5G核心网2B应用、5G数字化运维的开发工作;王刚,工程师,学士,主要从事4G/5G无线通信技术管理规划工作;邓子仪,工程师,学士,主要从事4G/5G核心网维护工作;田琪,工程师,学士,主要从事4G/5G核心网维护工作;马宝泽,讲师,博士,主要从事智能信息处理、数据流分析等理论研究工作。