

利用人工智能对光传输链路劣化隐患的预测研究

Research on Prediction of Optical Transmission Link Deterioration Using Artificial Intelligence

王 瑜,朱 宏,周 莹,沙升升(中国联合网络通信集团有限公司,北京 100033)

Wang Yu,Zhu Hong,Zhou Ying,Sha Shengsheng(China United Network Communications Group Co.,Ltd.,Beijing 100033,China)

摘 要:

随着网络运维要求的提高,传统的被动故障处理和人工巡检已无法满足需求,因此提出了基于AI的主动预防方法。首先分析了传统巡检手段的局限性,随后提出一个AI预测光传输链路劣化隐患的设计思路,该思路包括数据采集、任务管理、AI预测和风险呈现及上报等模块。实际应用效果表明,AI回归算法预测KPI的准确率可达96%,而高风险隐患预测准确率为75%,显示了AI技术在网络运维中的潜力和提升空间。研究结果表明,AI技术能够有效识别传统巡检难以发现的缓慢劣化类链路风险隐患,提升网络运维的主动性和效率。

关键词:

传输隐患;智能分析;人工智能;大数据;网络运维
doi:10.12045/j.issn.1007-3043.2024.12.014
文章编号:1007-3043(2024)12-0083-06
中图分类号:TN913
文献标识码:A
开放科学(资源服务)标识码(OSID):



Abstract:

As network O&M requirements increase, traditional passive fault handling and manual inspection cannot meet the requirements. Therefore, AI-based proactive prevention methods are proposed. It firstly analyzes the limitations of traditional inspection methods. Then, a design idea of AI prediction of optical transmission link deterioration is proposed, which includes data collection, task management, AI prediction and risk presentation and reporting modules. The actual application results show that the accuracy of the AI regression algorithm in predicting KPIs reaches 96%, and the accuracy of predicting high risks is 75%, which shows the potential and improvement space of the AI technology in network O&M. The research results indicate that the AI technology can effectively identify slow-deterioration link risks that are difficult to detect in traditional inspection, improving the initiative and efficiency of network O&M.

Keywords:

Transmission hidden trouble; Intelligent analysis; Artificial intelligence; Big data; Network O&M

引用格式:王瑜,朱宏,周莹,等. 利用人工智能对光传输链路劣化隐患的预测研究[J]. 邮电设计技术,2024(12):83-88.

1 概述

在日常网络维护中,绝大多数网络故障都是事后处理,即在发生故障后再根据故障产生的告警来界定定位故障的网元、故障位置,分析故障的原因^[1]。但随着新时代的到来,对网络运维的要求越来越高,网络

质量对客户投诉有着最直接的影响,所以网络运维不能再停留在被动的故障处理中,而要对主动性预防提出新的要求。传统的主动预防主要是人工巡检,即通过人工对大量性能KPI数据进行分析检测,来确定所有传输链路是否存在风险和隐患^[2]。但由于人工的效率较低,而设备的日常KPI数据量过于庞大,导致能够分析出的风险和隐患较少,很多轻微劣化、缓慢劣化的故障并不能通过人工或者简单的识别规则及时发

收稿日期:2024-11-18

现,而且通过人工巡检识别风险隐患的成本较高^[3]。那么,是否能利用AI算法在大数据处理、自适应学习等方法的优点,提升光传输链路的劣化与隐患预测预防能力,本文将就此进行重点研究^[4]。

2 传统隐患巡检手段的弊端

传统的隐患巡检手段一般为简单地查询传输段两端的单波光功率,计算获得标称光功率以后,将实际光功率和标称光功率进行对比,如果差异比较大,则报告隐患风险。如表1所示,一般采用标称值分析,即分析OMS路径上光放板的光功率和合波标称光功率的偏差,来判断是否存在隐患。

表1 标称值分析

评估项	评估方法	标准
光放输出总功率	a) 通过单波标称输入输出功率计算光放输出合波功率 (a) 标称单波输出功率目标值:根据光放单板类型查表获取 (b) 标称输出总功率目标值 $P_{总}(\text{dBm}) = P_{单波}(\text{dBm}) + 10\lg N(\text{dB})$, N 为波数 b) 比较输出总功率标称值和实际值的偏差	正常:偏差 ≤ 3 dB; 异常:偏差 > 3 dB

如果单纯按上述巡检规则,实际生成的风险隐患出现偏差的概率会比较大,主要原因如下。

a) 用计算得到的标称光功率的准确率比较低。标称光功率依赖单波的标称值来计算,实际的单波光功率和标称值不一定相同。

b) 不一定每台OTN设备都按照最大波数进行传输,实际波数存在缺失的可能,实际标称光功率可能比计算的标称光功率要低。

c) 在很多场景下,实际光功率可能大部分时间和标称光功率存在3 dB以上的偏差,但不存在风险。因为如果光功率长期属于稳态,则不会影响业务。

3 AI预测光传输链路劣化隐患设计思路

为了解决上述传统巡检手段的弊端,引入人工智能技术作为一个研究方向。人工智能的核心是机器可以自学习,那么在巡检隐患、预测风险领域,AI应该如何使用,起到什么作用呢^[5]?当系统收集到所有传输段链路的长期性能指标数据时,发现风险隐患的本质是要识别性能数据的特征并分析性能数据的趋势,以便发现存在的风险隐患。这就需要性能数据的模型抽象出来,找到风险隐患的传输段性能数据模型特征,以这些特征为输入,以特定算法模型为基础,由

机器输出性能模型的AI拟合曲线来预测未来性能数据的数值,并根据训练的模型评估得出有劣化风险的传输段链路,最后根据专家反馈的风险隐患结果反向优化算法模型^[6]。

光传输风险AI预测预防方案的整体架构如图1所示,主要模块说明如表2所示。

方案总体思路分2步。

a) 根据链路的历史性能数据,通过AI回归算法拟合出链路的性能变化曲线,并识别出有缓慢劣化趋势的风险链路。

b) 根据AI输出的劣化趋势曲线预测未来性能值(见图2),并通过模型来确定其中有劣化趋势的风险隐患链路,为链路设置风险等级。

4 核心算法逻辑

核心算法的整体框架如图3所示,其逻辑主要包括以下部分。

a) 风险过滤。使用三次样条插值来拟合时间序列,计算一阶导等特征值,并结合Mann-Kandel假设检验,对链路进行初筛,过滤掉非风险链路,将存在风险的链路输入到后续模型中^[7]。

b) 长周期预测。长周期预测的基本原理包括承认事物发展的延续性,应用历史数据,推测事物的发展趋势;考虑到事物发展的随机性,任何事物发展都可能受偶然因素影响^[8]。

c) 风险二次确认及风险等级划分。利用对筛选出的风险链路的劣化幅度预测结果,结合链路本身的VOA,能够自动为劣化幅度划分风险等级。

在以前的方案中,风险过滤时使用强规则进行判定,算法效果对配置设置存在强依赖,因此不同局点需设置不同的参数;并且时间序列需人工设置特征提取方式(如插值导数等),无法完全体现时序的整体特征,准确率低。

本文提出的新方案通过改进风险过滤阶段的算法以改进整体方案的准确性与灵活度,且完全兼容了原OTN链路检测中风险过滤层的输入和输出,对上下游无影响。新方案采用了Tsfresh特征选取器,替换了原先的人工特征选取方案,使特征选取更全面,对链路指标的表征更精确^[9]。其次,新方案增加了XGBoost模型,替换原方案中的规则过滤,使用实时的数据训练模型,使模型的推理结果可根据训练数据的变化而变化,无需人工修改规则,更加灵活^[10]。新方案的风

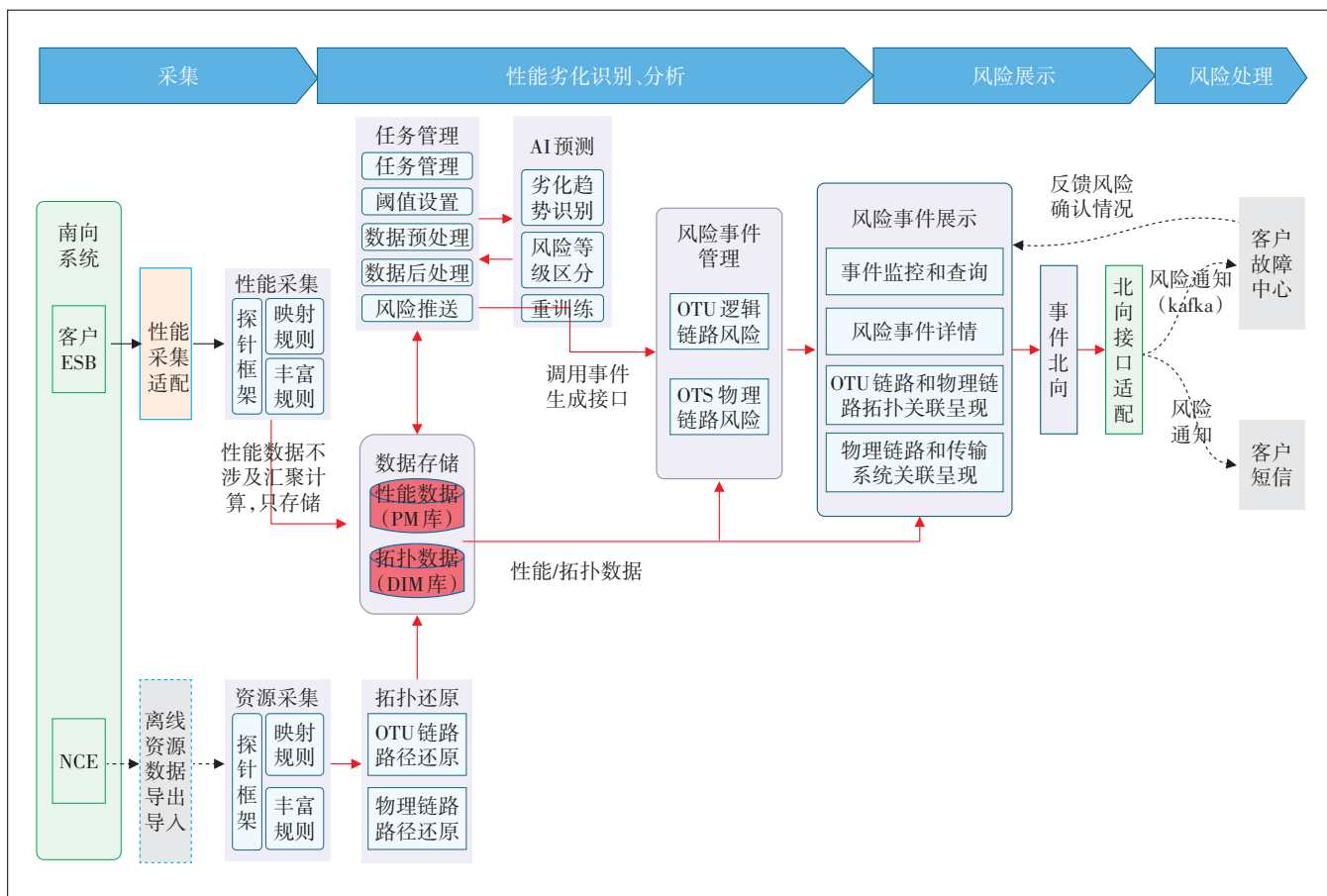


图1 光传输风险 AI 预测预防架构

表2 光传输风险 AI 预测预防功能

功能模块	主要功能点	功能简述
数据采集	性能数据采集	从设备EMS网管中导出的性能数据文件,通过SFTP下载到服务器,将其转换成NCE导出标准文件格式后,由PM模块进行采集入库
	资源数据采集	目前,资源数据从NCE网管的文件中直接导出并导入到实验室环境,再从实验室环境直接对接NCE的北向XML接口获取
任务管理	支持用户新建预测任务	用户可以新建预测任务,设定任务自动执行的开始时间、执行周期
	支持用户停止任务、编辑任务、手动启动任务、自动执行任务	任务执行过程中,用户可以停止任务;任务未执行的时候,用户可以编辑任务选项,并手动启动任务
AI预测	通过AI对性能数据进行分析,选出有劣化趋势的风险链路	通过对性能数据进行打点及线性拟合,找出具备缓慢劣化趋势的风险链路,并进行上报
风险呈现及上报	将AI预测出来的风险结果,在界面上进行呈现,并对接kafka接口和短信接口进行上报	①界面上能呈现出预测的风险链路,能查看相关风险链路的性能指标曲线图;②通过kafka接口和短信接口将相关的风险链路信息上报给网络运维管理人员

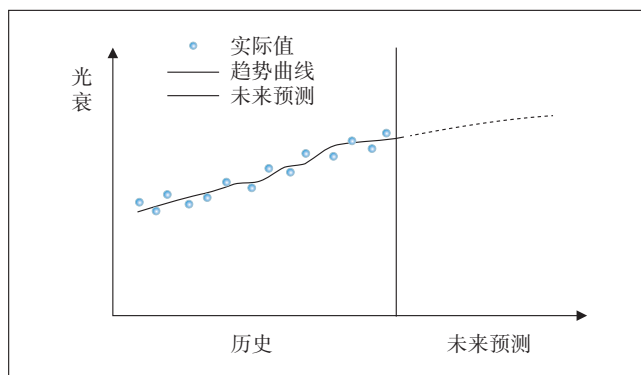


图2 劣化趋势曲线预测示意

险过滤模块如图4所示。

4.1 Tsfresh层设计

4.1.1 Tsfresh 特征提取器介绍^[11]

Tsfresh是一个用于自动提取时间序列特征的Python库,特别适用于机器学习和数据分析领域。Tsfresh能够自动检测并计算一组预定义的时间序列特征,无需手动选择或编写这些特征的计算代码,可以

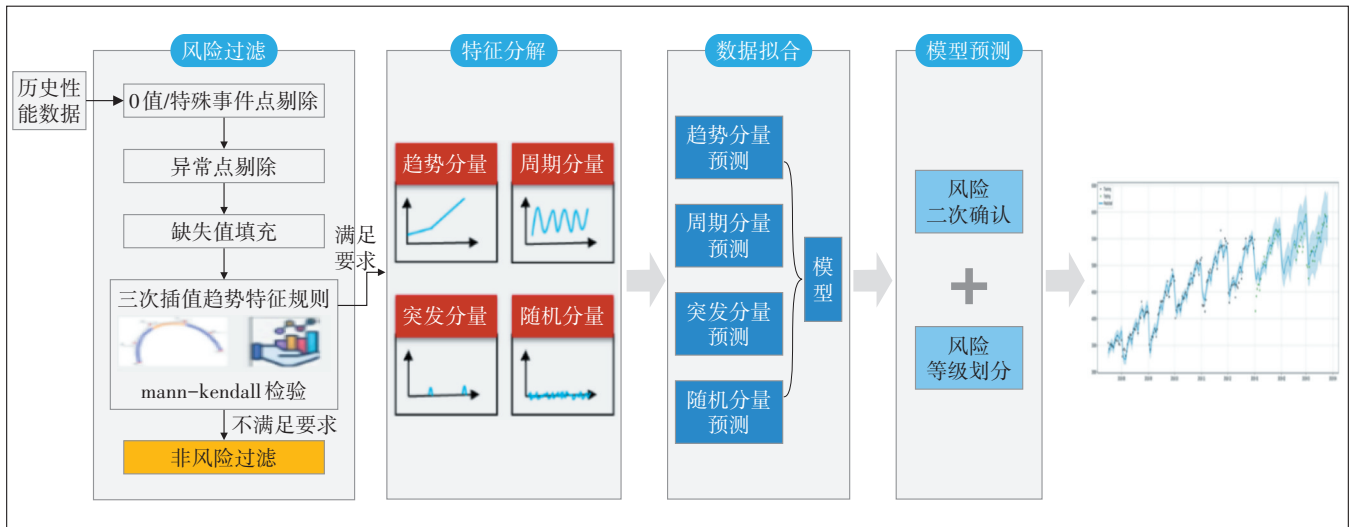


图3 AI逻辑算法示意

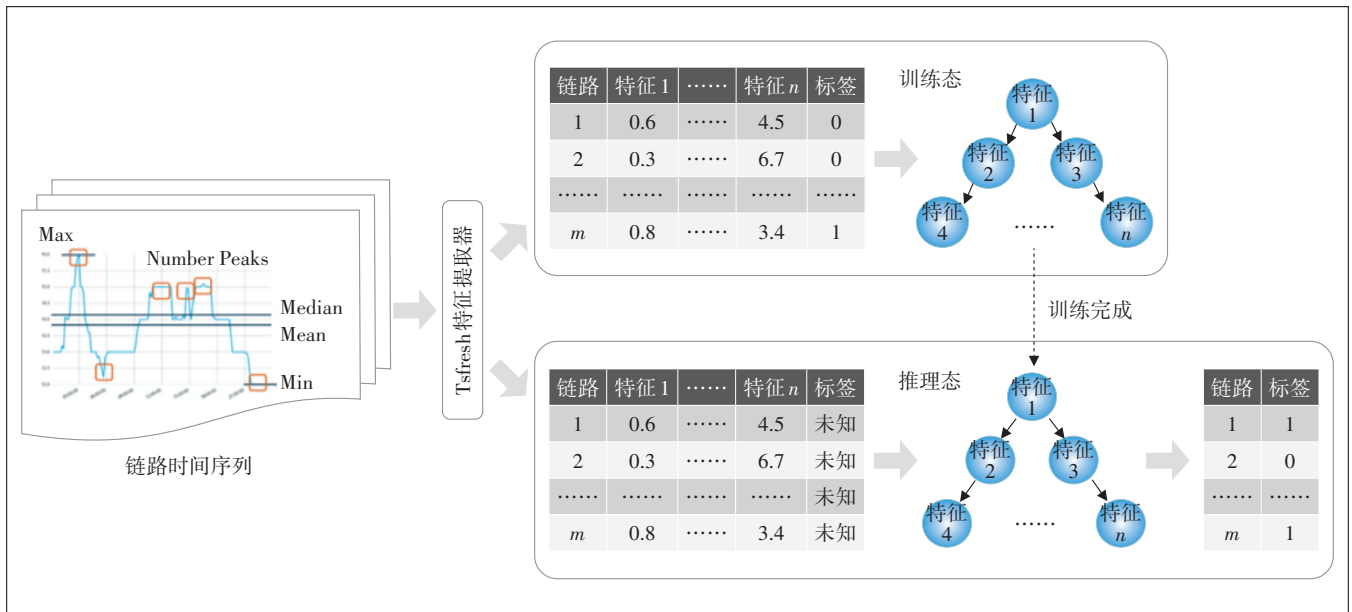


图4 风险过滤示意

显著减少数据预处理的时间。

Tsfresh的主要功能如下。

a) 自动特征提取。Tsfresh可以自动识别时间序列中的模式,并从中提取有意义的统计量,如最大值、最小值、均值、标准差、趋势、周期性等。

b) 特征选择。Tsfresh提供了一种基于假设检验的方法来筛选与目标变量最相关的特征,从而减少模型训练的复杂度。

c) 高效处理大规模数据。通过优化的算法设计,即使是在资源有限的环境下,Tsfresh也能够有效处理

大量时间序列数据。

4.1.2 数据准备

数据准备过程需要将时间序列数据加载到Pandas DataFrame中,包含时间序列值列、时间序列标号列以及时间列。

4.1.3 特征提取与筛选^[12]

使用Tsfresh包中的extract_features方法进行特征提取:

```
Features=extract_features (timeseries_container=
data,column_value=self.data_col,column_id=self.id_col,
```


column_sort=self. time_col, default_fc_parameters=cal_FCParameters)

其中, timeseries_containers 参数为时间序列数据的 DataFrame, column_value 为时间序列值列, column_id 为时间序列标号列, column_sort 为时间列, default_fc_parameters 为选取的特征集。特征集有 3 种不同的配置策略。

a) ComprehensiveFCParameters。该策略包括所有没有参数的特征和所有有参数的特征,每个特征都有不同的参数组合。如果不提交 default_fc_parameters,则采用 extract_features 的默认值。

b) MinimalFCParameters。仅包含少数功能,可用于快速测试。这里的特征“最小”。

c) EfficientFCParameters。其与 ComprehensiveFCParameters 中的大部分功能相同,但没有计算资源消耗较高的特征。如果运行时主要考虑性能,则可以使用此功能。

在链路时序数据中,通过 ComprehensiveFCParameters 可提取到 783 个特征,这些特征囊括了如最大值、最小值、均值、标准差、趋势、周期性等描述时间序列的特征。但是在亚健康检测任务中,并非每一个特征都能起到作用,还可能会造成计算资源的浪费。因此,对 783 个特征进行特征筛选:

selected_features = select_features(features, labels, ml_task='classification', fdr_level=1e-20)

其中, features 为提取的特征, labels 为每个链路是否劣化的标签, ml_task 为具体任务(分类或回归), fdr_level 参数是用来控制假发现率(False Discovery Rate, FDR)的阈值。FDR 是统计假设检验中的一项重要指标,它是指在所有被认定为“显著”的假设中,实际上为真(即不应被拒绝)的假设所占的比例的期望值。换句话说,这个参数的设置越低,那么选择得到的特征与任务的相关度就越高。

在实际的训练和推理之前,笔者做了一次完整的特征提取与筛选工作,使用 fdr_level=1e-20 选取了 173 个特征,并以 json 格式进行保存。在后续的任务中,可直接在特征提取时指定该批特征,无需进行多余的特征提取与特征筛选工作。

4.2 XGBoost 树分类器介绍^[13]

树模型是一种监督学习算法,用于分类和回归任务,它通过将特征空间递归地划分成多个区域来构建决策树。以判断某个人是否爱玩电动游戏的分类为

例,树模型用到了年龄和性别的特征,叶子结点处的值越高则代表越倾向于爱玩电脑游戏(见图 5)。

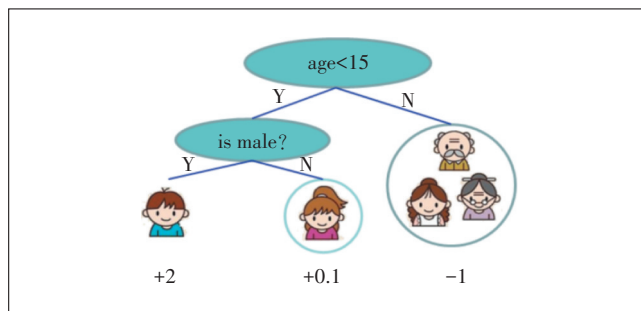


图 5 树模型示意

Boosting 是一种集成学习技术,它通过逐步添加弱预测模型(如树模型)来最小化损失函数。Boosting 算法的关键在于,每个新模型都在前一个模型的残差上进行训练,从而提高整体模型的预测能力。单纯只有一棵树,只用年龄和性别进行判断的话,在数据量更大的时候难免存在偏差,因此考虑使用 Boosting 技术增强树模型的拟合能力(见图 6)。

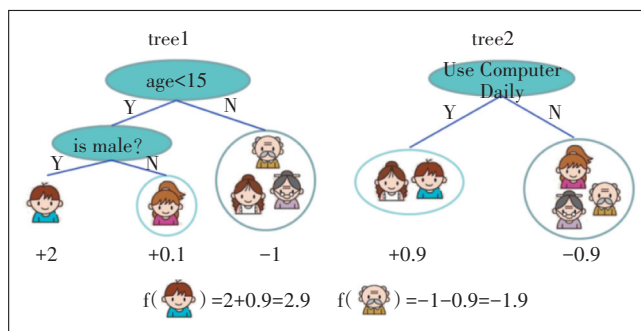


图 6 Boosting 技术增强树模型示意

XGBoost 算法正是基于上述概念产生的,在训练过程中,其损失函数如下。

$$L(\phi) = \sum_i l(\hat{y}_i; y_i) + \sum_k \Omega(f_k)$$

$$\text{where } \Omega(f) = \gamma T + \frac{1}{2} \lambda \|\omega\|^2$$

其中,第 1 项为每一棵树下,真实值与预测值的偏差之和;第 2 项为正则项,其值为所有数的复杂度。除此之外, XGBoost 算法还具备许多其他特点,总结如下。

a) 使用梯度下降算法来优化模型的损失函数,通过逐步添加新的树来最小化损失。

b) 在损失函数中加入了正则化项,包括 L1 和 L2 正则化,防止模型过拟合。

c) 采用深度优先的树生长策略, 并通过剪枝来控制树的复杂度。

d) 支持多线程和分布式计算, 可以有效地利用计算资源。

e) 支持多种数据类型和缺失值处理, 以及自定义的损失函数和评估标准。

5 AI 预测预防风险隐患的应用效果

如前所述, 预测预防风险隐患的实际应用效果主要依赖模型判断劣化链路的精准度以及 AI 回归曲线预测未来 7 天 KPI 值的准确度, 为此, 设计了 2 个指标来测试 AI 预测预防风险隐患的准确度。

a) AI 回归算法预测 KPI 准确率。主要是为了评估 AI 算法预测未来 7 天的具体数值和实际值的相对差值的误差, 以此来计算预测值的准确率。该值主要用来评估 AI 算法的能力。

$$\text{AI 预测 KPI 准确率} = 1 -$$

$$\left| \frac{\text{实际光衰平均值} - \text{AI 预测光衰值平均值}}{\text{实际光衰平均值}} \right| \times 100\%$$

b) 高风险隐患预测准确率。根据预测出来的有较大劣化风险隐患的数目, 和实际最后的劣化值做对比来评估预测的准确率。该值主要用来评估实际业务上的预测效果。

$$\text{高风险隐患预测准确率} = \frac{\text{实际劣化成 3 dB 的高风险链路数量}}{\text{预测高风险数量}} \times 100\%$$

经验证, 国内某运营商一干网络预测出的告警统计实例如表 3 所示。

表 3 AI 预测实例告警统计

总共风险隐患个数	预测高风险个数	实际高风险个数	AI 预测 KPI 准确率/%	高风险隐患预测准确率/%
57	33	24	96	73

从表 3 可以看出, AI 回归算法预测的未来 KPI 的准确率和实际的误差值在 5% 以内。但由于实际链路的光衰劣化, 1 dB 的劣化也许已经有质的区别, 所以目前实际导致高风险的比例达到了 73%, 还有较大的上升空间。未来将根据模型的不反馈, 重训练模型, 提高预测准确率。

6 结束语

通过对人工智能在传输预测预防领域的研究, 发

现传统巡检很难发现缓慢劣化类的链路风险隐患^[14]。本文创新性地将人工智能技术引入相关分析领域, 同时叠加了 CT 领域的专业知识作为第 3 维关键特征, 极大地提升了 AI 智能分析预测结果的准确性^[15]。同时, 还可以根据 AI 模型对劣化类链路的性能 KPI 曲线进行建模分析, 通过对模型的更新, 能更加准确地评估劣化风险隐患链路的特征, 准确识别劣化类风险隐患链路^[16]。

参考文献:

- [1] 伏玉笋, 杨根科. 人工智能在移动通信中的应用: 挑战与实践[J]. 通信学报, 2020, 41(9): 190-201.
- [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] 中国人工智能产业发展联盟. 电信网络人工智能应用白皮书[R/OL]. [2024-03-04]. <http://221.179.172.81/images/20210322/68111616377652538.pdf>.
- [4] 李开复, 王咏刚. 人工智能[M]. 北京: 文化发展出版社, 2017.
- [5] 王亮. 人工智能与数字出版的创新应用探讨[J]. 创新创业理论与实践, 2018, 1(9): 107-108.
- [6] 罗晓慧. 人工智能背后的机器学习[J]. 电子世界, 2019(14): 103.
- [7] 张新钰. 人工智能: 智能驾驶[M]. 北京: 电子工业出版社, 2020.
- [8] 杨爱喜, 卜向红, 严家祥. 人工智能时代: 未来已来[M]. 北京: 人民邮电出版社, 2018.
- [9] 野边继男. 深入理解 ICT 与自动驾驶[M]. 陈慧, 张诚, 陈恭羽, 译. 北京: 机械工业出版社, 2018.
- [10] 钱兵. 智能运维之道: 基于 AI 技术的应用实践[M]. 北京: 机械工业出版社, 2022.
- [11] 彭冬, 朱伟, 刘俊. 智能运维: 从 0 搭建大规模分布式 AIOps 系统[M]. 北京: 电子工业出版社, 2018.
- [12] 戴昆, 约翰逊, 默克. Splunk 智能运维实战[M]. 宫鑫, 康宁, 刘法宗, 译. 北京: 机械工业出版社, 2015.
- [13] 韦尔登. 下一代网络[M]. 宋飞, 徐格, 钱蔓黎, 等, 译. 北京: 人民邮电出版社, 2019.
- [14] 饶琛琳. 网站运维技术与实践[M]. 北京: 电子工业出版社, 2014.
- [15] 余洪春. DevOps 和自动化运维实践[M]. 北京: 机械工业出版社, 2018.
- [16] 周志华. 机器学习[M]. 北京: 清华大学出版社, 2016.

作者简介:

王瑜, 毕业于南京邮电大学, 高级工程师, 主要从事智能运维产品研发工作; 朱宏, 毕业于宁波大学, 工程师, 硕士, 主要从事智能运维产品研发工作; 周莹, 毕业于南京邮电大学, 工程师, 硕士, 主要从事智能运维产品设计工作; 沙升升, 毕业于江苏科技大学, 工程师, 学士, 主要从事智能运维产品架构管理工作。