

风电业务流程剩余时间预测与系统应用

李建元¹, 田玉超², 尹昱妍², 田甜², 刘新锋²

(1. 中国移动通信集团山东有限公司; 2. 山东建筑大学 计算机科学与技术学院, 山东 济南 250101)

摘要: 针对工业场景下运维事件日志复杂以及数据量较少等问题, 提出一种基于时间卷积网络的剩余时间预测方法。首先, 使用因果卷积对流程实例数据进行建模; 其次, 对不同长度的轨迹前缀分别进行模型训练, 以提高模型的针对性; 最后, 将剩余时间预测模型用于实际的风电运维业务场景中, 开发风电运维业务系统, 实现工作票表单数据以及事件日志数据的可视化展示。与传统的剩余时间预测方法相比, 该方法的预测效果有显著提升, 平均绝对误差值降低了5%左右。开发的风电运维业务系统能够预测未完成工作票开票过程中流程的剩余时间, 实现业务流程超时告警。

关键词: 业务流程分析; 剩余时间预测; 时间卷积网络; 风电业务流程; 流程可视化

DOI: 10.11907/rjdk.231451

开放科学(资源服务)标识码(OSID):

中图分类号: TP181

文献标识码: A

文章编号: 1672-7800(2024)006-0018-07



Wind Power Business Process Remaining Time Prediction and System Application

LI Jianyuan¹, TIAN Yuchao², YIN Yuyan², TIAN Tian², LIU Xinfeng²

(1. China Mobile Communications Group Shandong Co., Ltd;

2. School of Computer Science and Technology, Shandong Jianzhu University, Ji'nan 250101, China)

Abstract: Aiming at the complexity of industrial scene event logs and the small amount of data in industrial scenarios, a remaining time prediction method based on time convolutional network is proposed. Firstly, causal convolution is used to model process instance data. Then, the model is trained on track prefixes of different lengths to improve the pertinence of the model. Finally, the remaining time prediction model is used in the actual wind power operation and maintenance business scenario to develop the wind power operation and maintenance business system to realize the visual display of work ticket form data and event log data. Compared with the traditional remaining time prediction method, the prediction effect of the proposed method is significantly improved, and the average absolute error is reduced by an average of 5%. The developed wind power wind power operation and maintenance business system can predict the remaining time in the process of invoicing unfinished work tickets, and realize business process overtime alarms.

Key Words: business process analysis; remaining time prediction; temporal convolution network; wind power business process; process visualization

0 引言

业务流程管理旨在提供规范化的技术和工具, 针对生产流程、行政申请流程等各类业务流程, 利用流程挖掘技

术进行分析、建模、监控和持续优化^[1,2]。清洁能源的运行业务场景所面临的随机性和不确定性对运维工作提出了更高要求, 因此运维工作需要规范化的流程管理^[3]。风电运维管理系统被视为一种流程性管理系统, 其数据信息对于运维人员的决策起着重要作用。在业务流程中, 时

收稿日期: 2023-04-29

基金项目: 国家自然科学基金项目(51975332); 山东省重点研发计划(重大科技创新工程)项目(2021CXGC011204)

作者简介: 李建元(1985-), 男, 硕士, 中国移动通信集团山东有限公司高级咨询师, 研究方向为5G+新一代信息技术; 田玉超(1997-), 女, 山东建筑大学计算机科学与技术学院硕士研究生, 研究方向为流程挖掘; 尹昱妍(1998-), 女, 山东建筑大学计算机科学与技术学院硕士研究生, 研究方向为流程挖掘; 田甜(1987-), 女, 博士, 山东建筑大学计算机科学与技术学院副教授, 研究方向为软件工程、算法优化; 刘新锋(1977-), 男, 博士, CCF会员, 山东建筑大学计算机科学与技术学院副教授, 研究方向为工业大数据分析、流程挖掘。本文通讯作者: 刘新锋。

间限制是非常重要的因素,因其直接影响着流程的成本、风险和效率等关键指标。通过预测未来执行过程中的流程剩余时间,可以帮助业务流程执行者及时发现潜在问题,并采取相应措施,以优化整个业务流程的执行过程。因此,剩余时间预测是流程挖掘技术中非常重要的一环^[4-5]。通过流程挖掘技术可以对历史数据进行分析 and 挖掘,从而构建出预测模型,用来预测未来执行过程中的流程剩余时间,从而为流程管理者提供必要的决策依据,以有效地进行资源调度,避免执行超时等问题。

本文提出一种基于时间卷积网络的剩余时间预测模型,通过提取事件日志数据中的属性特征进行模型训练和优化,利用事件日志数据进行工作票中当前执行流程的时间预测,以判断流程是否超时,并将该剩余时间预测模型应用于风电智慧运维中,以实现风电运维开票流程的剩余时间预测和实时告警,从而提高运维的效率和安全性^[6]。

1 相关工作

针对剩余时间预测与风电业务场景数据预测,已经有了许多研究成果。在传统的业务流程模型预测方法中,Folino等^[7]将不同执行场景的发现与日志事件的自动抽象相结合,对流程剩余时间预测任务中的特征设计进行研究,可更有效、更容易地使用日志数据进行过程挖掘预测。Claes等^[8]提出一种合并组织间流程不同部分的输入数据技术,将其作为流程挖掘算法的输入,同时为用户提供了一种用于合并数据的规则算法。Aalst等^[9]采用一种抽象机制学习历史事件日志信息,并构建带有标注的变迁系统,将到目前为止执行的活动序列通过状态表示函数映射到变迁系统中的一个状态,然后基于类似状态下较早过程实例的平均完成时间进行预测。除使用变迁系统外,后续还有一些研究者使用队列流程模型^[10]、设计序列编码方案^[11-12]、声明式流程模型^[13]等方式进行剩余时间预测。

随着人工智能技术的快速发展,利用机器学习与深度学习方法进行业务流程剩余时间预测的研究越来越多,应用机器学习与神经网络模型构建的剩余时间预测方法也取得了很好的预测效果。Polato等^[14]提出数据感知的变迁系统,采用 ϵ -SVR机器学习算法,通过业务过程的历史踪迹进行模型训练,并通过编码技术将部分踪迹的特征转换为适合 ϵ -SVR输入的格式,模型的输出值即为剩余时间预测值。Rogge-Solti等^[15]根据历史信息构建了分布式变迁随机Petri网,对正在执行的业务过程获取当前状态,然后收集模拟结果(完成时间),通过模拟样本均值,采用最大似然估计值获得剩余执行时间的预测值。Rogge-Solti等^[16]还通过历史踪迹计算每个变迁活动的持续时间,以此统计时间信息,对于正在执行中的流程,根据当前活动的执行时间从原始分布的截断分布中进行抽样,获取所有样本模拟的完成时间,并取平均值作为最后的预测结果。

Ceci等^[17]利用序列模式挖掘并识别部分过程模型,使用匹配部分过程模型的活动附加信息从事件日志中训练嵌套预测模型,用于预测新(运行)流程实例的下一个活动和完成时间。Bevacqua等^[18]使用聚类技术处理事件日志,然后利用回归技术构建剩余时间预测模型,在预测准确性和鲁棒性方面都取得了令人满意的结果。Senderovich等^[19]考虑了流程实例自身的属性、流程实例内部事件的属性以及业务流程实例的上下文信息,采用有监督的机器学习方法进行预测,验证了该方法的实用性和适用性。Verenich等^[20]采用白盒预测流程实例剩余执行时间的方法,使预测过程中的指标具有可解释性。首先使用聚类技术对历史流程实例进行处理划分,然后采用回归分析技术进行剩余时间预测。该方法获得了较准确的预测结果,缺点是未考虑流程实例活动的可重复性。

倪维健等^[21]针对业务流程剩余时间预测问题提出一个基于注意力双向循环神经网络的业务流程剩余事件预测方法,提取流程实例的轨迹前缀及其对应的剩余执行时间,利用双向结构对轨迹前缀进行建模,并利用注意力机制学习到的权重对其作最后的处理,最后采用一个全连接层进行剩余时间预测。该方法采用平均绝对误差进行效果评估,验证了所提出模型的优越性。因为不同长度的流程实例之间存在很大差异,会影响预测结果^[22]。刘彤等^[23]针对流程轨迹前缀长度的不同设计一种事件表示方法,并利用深度迁移学习进行剩余时间预测,考虑了轨迹前缀之间的关联性,克服了长轨迹前缀数据量少的问题,提高了剩余时间预测效率。Navarin等^[24]通过将活动信息转化为模型支持的输入格式,将未完成的踪迹信息输入模型中进行时间预测,不仅考虑了每个事件的时间特征,而且考虑了与事件相关的轨迹开始时间、最后一个事件开始时间等其他特征,但该方法的编码维数会随着离散特征取值的变大而增加。徐兴荣等^[25]设计一种准循环神经网络,充分利用输入序列的顺序信息用于业务流程剩余时间预测,解决了对序列数据跨时间步长并行处理的问题。

现有风电运维管理系统实现了对运维过程中数据的管理,例如康磊等^[26]针对现有风场管理模式造成的信息不能共享、专业人员分散等问题,提出建立层级管理模式,按照权限对每一级工作建立功能模块。Besnard^[27]主要针对海上风力发电运维提出风电系统,可高效地维护风电场设备,并且降低运维成本。陈雷^[28]采用.NET技术以及SQL Server数据库设计实现了风电运维管理系统等。针对风电运维过程中产生的数据信息,秦政等^[29]基于先进统计方法实现了对风电功率的预测。郭娜等^[30]提出一种基于优先级的特征自选取策略,以解决风机检修过程中存在的操作失误和工作延期等问题。

目前的业务流程剩余时间预测方法虽然取得了很好的预测效果,但是模型的可解释性较差,限制了业务流程管理人员对模型预测结果的理解和应用。同时,风电运维

过程中业务流程复杂性高、变更内容多、流程实例差异大,而时间卷积网络具有较好的可解释性以及较高的模型精度,并方便进行可视化,可以自适应地学习不同时间尺度上的特征,更好地处理复杂的业务流程数据,并实现实时的业务流程剩余时间预测,具有较高的实际应用价值^[31]。

2 基本定义

定义1 事件。业务流程中某个活动的一次执行实例。例如,风电开票流程中工作票的填写可以形式化表示为一个元组 $e = (a, id, start, end, p_1, \dots, p_m)$ 。其中, a 表示该事件执行的活动, id 表示该事件所对应流程实例的ID, $start, end$ 分别表示该事件执行的开始时间与结束时间, $p_1, \dots, p_m$ 表示该事件的属性。

定义2 流程轨迹。业务流程中每一个活动节点都可成为事件,该流程中所有事件集合组成的非空序列可称为流程轨迹,可用 $\sigma = \langle e_1, e_2, \dots, e_{|\sigma|} \rangle$ 表示。其中, e_i 表示流程轨迹中的一个事件, $|\sigma|$ 表示流程轨迹中的事件数量, $1 \leq i \leq |\sigma|$ 。

定义3 轨迹前缀。轨迹前缀是流程轨迹 σ 的前 k 个事件,可表示为 $\sigma^{(k)} = \langle e_1, e_2, \dots, e_k \rangle (1 \leq k \leq |\sigma|)$ 。

定义4 业务流程剩余时间。在一个流程轨迹中,从

目前正在执行的事件到最后一个事件结束之间的时间,即 $remain = e_{|\sigma|,end} - e_{k,end}$ 。

定义5 事件日志。所有流程轨迹的集合称为事件日志,记录了整个业务系统的历史执行情况,即 $L = \{\sigma_1, \sigma_2, \dots, \sigma_l\}$ 。

定义6 业务流程剩余时间预测。给定某业务系统,令该系统的事件空间为 E ,则流程实例及前缀空间可表示为 E^* 。剩余时间预测任务的目标是利用事件日志中历史流程实例数据构建映射函数 $f: E^* \rightarrow \mathcal{R}^+$,利用该函数可以预测给定流程的剩余时间。

3 基于时间卷积网络的剩余时间预测方法

基于时间卷积网络的剩余时间预测方法整体流程如图1所示。首先,指定一个长度范围 $[\min, \max]$,用于对事件日志中的流程轨迹进行截取,得到不同长度的实例前缀数据信息,并根据不同长度的训练样本将其划分为不同的训练数据集。然后基于划分的训练集训练剩余时间预测模型。此处,考虑到不同长度的训练样本单独训练剩余时间预测模型时无法有效利用各个轨迹前缀之间的关联性,不确定因素太多,因此将迁移学习的思想加入其中。在训练前缀长度为 l 的数据时,以前缀长度为 $l-1$ 的数据训练模型所得的参数作为模型初始参数。

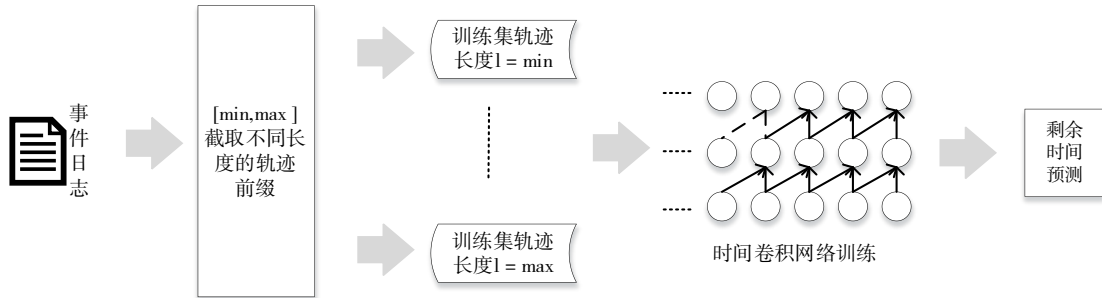


Fig. 1 Flow of remaining time prediction method

图1 剩余时间预测方法流程

3.1 数据预处理

在训练集划分阶段,需要先进行数据处理工作,将事件日志 .xes 格式转化为神经网络模型所接受的 .csv 数据格式,然后采用基于轨迹前缀长度的方式构建训练集。通过截取不同长度的轨迹前缀作为训练数据集中的训练样本,最后返回所有长度的训练数据集。

对于连续属性时间,需要先进行离散化处理,再采用 CBOW (Continuous Bag of Words) 编码方式进行数据编码,将数据按照 7:3 的比例划分为训练集和测试集。CBOW 模型结构如图2所示。

将事件日志数据利用 one-hot 编码后,得到当前事件的上下文事件向量,将其输入到 CBOW 模型中,通过与隐藏层的共享矩阵相乘,加权求平均,最后利用 softmax() 函数获取当前事件的概率分布,将概率值最大的向量表示为

当前事件。

3.2 时间卷积网络模型

传统的卷积神经网络对于时序数据的处理较为困难,不能很好地捕捉长期依赖信息,时间卷积网络 (Temporal Convolutional Network, TCN) 是将卷积神经网络思想引入时序数据处理的一种方法,其结构包括膨胀因果卷积层和残差结构,具有灵活的感受野和稳定的梯度,能够有效解决训练过程中网络性能减弱等问题,可以很好地处理时间序列,在多种任务上处理效果甚至超过通常用于时序数据处理的循环神经网络模型^[32-33]。

时间卷积网络包括因果卷积和膨胀卷积两种。因果卷积结构中,对于上一层 t 时刻的值,只依赖于下一层 t 时刻及其之前的值,是一种单向结构。由于网络层数的限制,因果卷积对历史特征信息的获取范围小,而膨胀因果

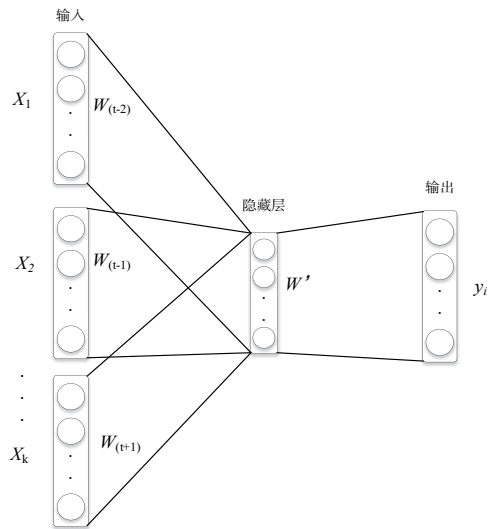


图 2 CBOW 模型结构

卷积通过调节膨胀系数能够灵活改变卷积操作的感受野。膨胀因果卷积包含一个膨胀因子 d , 允许卷积时输入存在间隔采样, 采样大小受膨胀因子影响。膨胀因果卷积结构如图 3 所示。

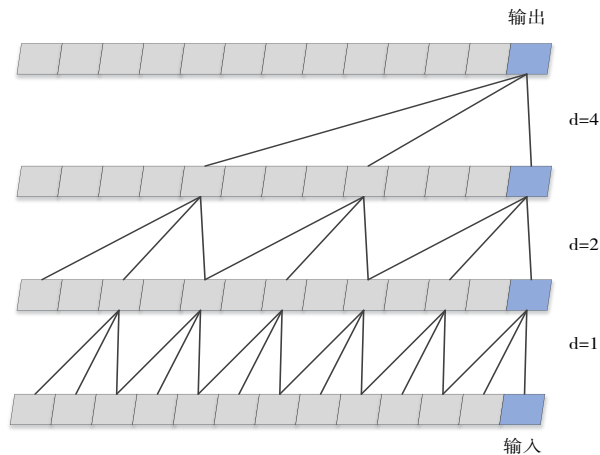


图 3 膨胀因果卷积结构

TCN 使用残差结构可以使输入数据 x 和输出数据 $F(x)$ 通过 add 操作实现线性叠加, 使网络实现跨层传递信息, 加快训练过程, 避免 TCN 层在叠加过程中产生梯度消失现象。残差结构如图 4 所示, 其中包含两层膨胀因果卷积, 每层中都包含参数重写、随机失活等操作, 用于进行正则化处理。

本文使用时间卷积网络对轨迹前缀数据进行建模, 设置膨胀因子 d 为 8, 卷积核大小为 2×2 , 将其与残差结构相结合, 提高模型的鲁棒性, 最后通过一个全连接层实现剩余时间预测。

3.3 实验结果与分析

本文使用 3 个公开事件日志数据进行对比实验。事件日志数据分别是 BPI_2012_A、BPI_2012_O、BPI_

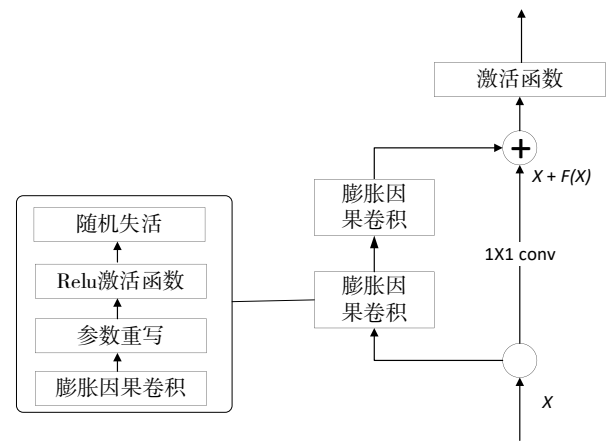


图 4 残差结构

2012_W, 均来自不同领域, 可在 4TU Center for Research Data(<https://data.4tu.nl>)上下载。其中, BPI_2012 类数据是某财政机构贷款审批的事件日志。数据统计信息如表 1 所示。

表 1 数据统计信息

数据集	轨迹数量	事件数量	活动数量	轨迹最大长度	轨迹最小长度
BPI_2012_A	13 087	73 022	10	10	3
BPI_2012_O	5 015	41 728	7	39	4
BPI_2012_W	9 658	147 450	6	153	1

进一步采用平均绝对误差 (Mean Absolute Error, MAE) 作为剩余时间预测模型的评估指标, MAE 值越小, 说明预测效果越好。计算公式如式 (1) 所示。

MAE = \sum_{(\sigma^{(k)}, remain(\sigma, k))} |f(\sigma^{(k)}) - remain(\sigma, k)| \tag{1}

其中, $f(\sigma^{(k)})$ 表示长度为 k 的轨迹前缀剩余时间预测值, $remain(\sigma, k)$ 表示该轨迹前缀剩余时间的真实值。

选择 5 个基准方法进行对比实验, 分别是集合 (TS-set)、序列 (TS-sequence) 基于 Petri 网的方法 (SPN)、长短记忆神经网络 (Long Short Term Memory, LSTM)、门控循环神经网络 (Gate Recurrent Unit, GRU)。实验结果如表 2 所示。

本文提出的剩余时间预测方法在各个数据集上均优于传统的剩余时间预测方法, 主要原因是事件日志具有复

表 2 业务流程剩余时间预测方法实验结果

对比方法	BPI_2012_A	BPI_2012_O	BPI_2012_W
TS-set ^[23]	7.505	7.392	8.429
TS-sequence ^[23]	7.488	9.612	8.619
SPN ^[23]	8.880	6.385	8.516
LSTM ^[23]	3.588	7.993	8.021
GRU ^[23]	3.895	7.324	6.153
本文方法(TCN)	3.498	5.783	4.986

杂且序列较长的特点。时间卷积网络可以接受任意长度的时间序列,并且具有灵活的感受野,对时序数据往往具有较好的预测效果。

本文方法的 MAE 值较对比方法具有明显下降,在各个数据集上均优于对比方法,MAE 值平均下降了 5%,验证了本文方法的优越性。

4 风电业务流程数据预测及应用

4.1 风电业务流程剩余时间预测

风电运维流程中可能涉及多种子流程,如开票流程、人员变更流程、延期流程、签发流程、终结流程等。其中,开票流程包括工作票填写、工作票签发、安全措施补充、工作时间批准、工作时间确认、工作许可、工作确认等,主要发生在工作票申请开始阶段;延期流程包括工作票延期申请、工作票延期批准,在工作票流转过程中,当实际运维工作出现延期时发生;人员变动流程涉及的活动包括工作组成员变动、工作组负责人变动审批,主要发生在运维过程中出现工作人员或负责人变更时;终结流程包括工作终结申请、工作终结许可、工作验收单开具、工作票终结、方案活动完善等,在工作票结束时发生。不同的业务场景会面临不同业务流程,造成风电流程的多样性与不确定性。因此,在流程可视化时需要考虑各个流程直接的关联性。

根据风电运维系统产生的日志数据筛选出工作票编码、工作 ID、工作名称、流程 ID、流程名称、活动 ID、活动名称、活动完成时间等数据。对于已执行完成的流程,将每个流程中最后一个活动的完成时间表示为该流程的完成时间。经过整合后,原始的日志数据中保留了工作单编码、工作 ID、工作名称、流程 ID、流程名称与流程完成时间。

进行流程剩余时间预测时,将实际业务流程作为事件前缀,参与轨迹编码、时间戳预处理等工作。当轨迹长度为 n 时,轨迹前缀为 $\sigma^n = \langle e_1, e_2, \dots, e_{n-1} \rangle$,其中 $n \geq 2$,即保证该轨迹的轨迹前缀长度至少为 1。

在模型训练前,首先要对时间属性进行离散化处理,具体操作方式为利用年月日与秒之间的数量关系计算时间属性,将其转化为秒数,用 `second()` 函数表示。对于流程轨迹的剩余时间表示,假设该轨迹中当前事件的完成时间为 t_n , s_n 是本流程轨迹的结束时间与当前流程结束时间的时间差,则 s_n 的计算如式(2)所示。

$$s_n = \text{second}(t_n) \quad (n \geq 1)$$
 (2)

基于时间卷积神经网络模型进行风电工作票流程的剩余时间预测,需要将训练数据中已完成工作票的流程完成时间转换为 s_n ,时间转换示例如表 3 所示。然后进行流程中的活动属性编码,并将其与流程 ID 进行对应。将处理后的事件日志数据按照 CBOW 编码方式进行编码作为模型的输入数据,实现剩余时间预测。将预测结果 d 与当

前流程的完成时间相加,即可得到总流程的结束时间预测值 $\bar{s}_n$ 。后期将预测结果以相反的方式转换为年月日格式(yyyy-mm-dd),即可得到工作票执行流程的预测完成时间。

Table 3 Example of process end time conversion for work ticket 102A20200001

表 3 工作票 102A20200001 流程结束时间转换示例

流程名称	流程结束时间 t_n	s_n	$\bar{s}_n$
电气一种票开票流程	2020-02-15 13:59:13	$s_0=158\ 174\ 635\ 3$	$\bar{s}_0 = s_0 + d$
电气一种票人员变动	2020-02-15 14:05:30	$s_1=158\ 174\ 673\ 0$	$\bar{s}_1 = s_1 + d$
电气一种票延期流程	2020-02-15 14:10:00	$s_2=158\ 174\ 700\ 0$	$\bar{s}_2 = s_2 + d$
电气一种票终结流程	2020-02-15 14:15:25	$s_3=158\ 174\ 732\ 5$	$\bar{s}_3 = s_3 + d$

所得部分预测结果如表 4 所示。将该预测结果与真实记录的流程结束时间进行对比,作为工作票流程超时告警功能实现的基础,帮助工作人员及时处理工作票开票流程超时产生的问题,以及预防因流程超时带来的资源冲突等问题。

Table 4 Forecast results of work ticket process end time

表 4 工作票流程结束时间预测结果

工作票编码	流程 ID	流程名称	预测结束时间
107A20200009	373	电气一种票开票流程	2020/10/21 9:16:50
201B20200019	377	电气二种票开票流程	2020/11/8 11:36:32
105C20200007	380	风机检修票开票流程	2020/1/20 23:43:20
102C20200004	380	风机检修票开票流程	2020/1/21 10:18:01
105A20200003	373	电气一种票开票流程	2020/11/11 14:20:23
102C20200036	380	风机检修票开票流程	2020/2/19 9:17:35
202B20200014	377	电气二种票开票流程	2020/5/19 17:50:34
103C20200005	380	风机检修票开票流程	2020/2/22 17:29:26
107A20200008	373	电气一种票开票流程	2020/10/21 9:06:52

4.2 系统功能架构

为了适应系统的业务流程,将工作票信息、工作票开票流程、工作活动信息整合在一起,包括工作票表单编码、流程名称、开始时间、结束时间、超时时间等,并展示开票过程中工作票流程相关信息,以实现对工作票流程的监控,便于工作人员进行查询。针对未完成的工作票,若正在执行的流程时间晚于预测时间,且该流程仍未完成,则该流程处于超时状态,同时显示超时时长,起到超时警告的作用;若该流程在预测时间之前结束,则该流程为正常结束状态。其中,业务单元包括工作票流程信息管理、工作票流程监控、工作流程执行时间预测。根据各级人员的职责分为工作人员、部门管理者、领导层。具体系统功能如图 5 所示。

4.3 流程可视化展示

本文在已有风电运维系统的基础上加入流程挖掘算法,实现一个基于人工智能算法进行数据分析、业务全面

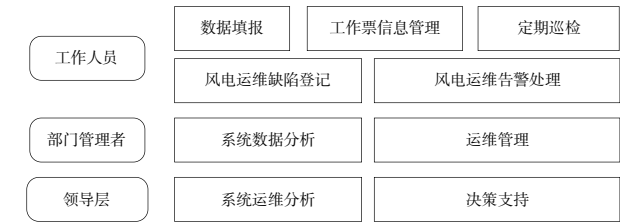


Fig. 5 System function
图 5 系统功能

的风电流程可视化系统,将利用事件日志预测风电运维流程的剩余时间预测信息通过系统进行展示,实现风电运维流程的实时监控,提高工作人员的工作效率。

基于时间卷积网络实现对工作票流程的剩余时间预测,以已完成的工作票流程为训练数据,利用训练结果实现对未完成工作票流程的剩余时间预测。将预测结果与真实结果进行对比,如表 5 所示,可以看出,基于时间卷积网络的剩余时间预测结果与流程实际结束时间误差较小,可将其应用于实际风电运维系统中对流程进行超时判断。

Table 5 Comparison of partial forecast results for wind power operation and maintenance data

表 5 风电运维数据部分预测结果对比

工作票编码	流程ID	真实结束时间	预测结束时间
102A20200001	373	2020-02-15 13:59:42	2020-02-15 14:10:42
102A20200001	374	2020-02-15 14:05:00	2020-02-15 14:05:29
102A20200001	375	2020-02-15 14:10:34	2020-02-15 14:10:34
102A20200001	376	2020-02-15 14:15:00	2020-02-15 14:44:46
102A20200001	377	2020-02-15 13:06:01	2020-02-15 13:06:01
102A20200001	378	2020-02-15 13:10:00	2020-02-15 13:11:55
102A20200001	379	2020-02-15 13:24:51	2020-02-15 13:34:51

图 6 给出了表 5 中预测结果与实际流程结束时间的拟合情况图,通过将预测结果、实际流程结束时间与流程开始时间的数值作差,以秒为单位,查看整个流程执行过程所用时间是否一致。从图 6 可以看出,预测结果与时间流程结束时间基本拟合,其他存在误差的情况也不超过 30min,说明本文提出的剩余时间预测模型取得了较好效果,具有一定的可行性。

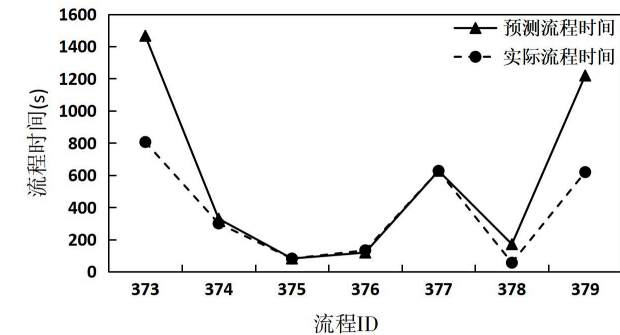


Fig. 6 Visualization of remaining time prediction
图 6 剩余时间预测可视化图

5 结语

本文基于时间卷积网络实现业务流程的剩余时间预测研究,并在公开事件日志上进行了实验验证。相比于已有方法,本文方法可有效降低剩余时间预测误差,具有良好的预测效果。将剩余时间预测方法应用于风电运维服务中,可以对风电工作票开票过程中的业务流程进行实时监控,辅助工作人员进行决策,提高风电运维效率。因此,本文提出的基于时间卷积网络的剩余时间预测方法具有重要的理论与实际应用价值。在未来的研究中,可进一步探索如何提高预测精度和预测的实时性,以满足更加复杂的业务场景需求。

参考文献:

[1] ZHANG Y C, LIU K P, LIAO X B, et al. Research on multi-time scale source-load coordination dispatching model of power system including large-scale wind power [J]. High Voltage Technology, 2019, 45 (2) : 600-608.
张亚超,刘开培,廖小兵,等. 含大规模风电的电力系统多时间尺度源荷协调调度模型研究[J]. 高电压技术,2019,45(2):600-608.

[2] AALST W M P. Process mining[J]. Communications of the ACM, 2012, 55(8) : 76-83.

[3] AALST V D. Using process mining to bridge the gap between BI and BPM [J]. Computer, 2011, 44(12) : 77-80.

[4] CUI L. Prediction of business process system based on machine learning [D]. Beijing: Beijing University of Posts and Telecommunications, 2019.
崔亮. 基于机器学习的业务流程系统的预测[D]. 北京:北京邮电大学,2019.

[5] TEINEMAA I, DUMAS M, ROSA M L, et al. Outcome-oriented predictive process monitoring: review and benchmark [J]. ACM Transactions on Knowledge Discovery from Data, 2019, 13(2) : 1-57.

[6] ZHAO Y, WANG H M, KANG L, et al. Short-term electric load forecasting based on time convolutional network [J]. Chinese Journal of Electrotechnical Society, 2022,37(5):1242-1251.
赵洋,王瀚墨,康丽,等. 基于时间卷积网络的短期电力负荷预测[J]. 电工技术学报,2022,37(5):1242-1251.

[7] FOLINO F, GUARASCIO M, PONTIERI L. Mining predictive process models out of low-level multidimensional logs [C]//International Conference on Advanced Information Systems Engineering, 2014: 533-547.

[8] CLAES J, POELS G. Merging event logs for process mining: a rule based merging method and rule suggestion algorithm [J]. Expert Systems with Applications, 2014, 41(16) : 7291-7306.

[9] AALST W M P V D, SCHONENBERG M H, SONG M. Time prediction based on process mining [J]. Information Systems, 2011, 36 (2) : 450-475.

[10] SENDEROVICH A, WEIDLICH M, GAL A, et al. Queue mining for delay prediction in multi-class service processes [J]. Information Systems, 2015, 53 : 278-295.

[11] LEONTJEVA A, CONFORTI R, FRANCESCO MARINO C D, et al. Complex symbolic sequence encodings for predictive monitoring of business processes [C]//International Conference on Business Process Management, 2016: 297-313.

[12] SURIADI S, ANDREWS R, TER H A H M, et al. Event log imperfec-

- tion patterns for process mining: towards a systematic approach to cleaning event logs[J]. *Information Systems*, 2017, 64: 132–150.
- [13] JIMENEZ-RAMIREZ A, BARBA I, FERNANDEZ-OLIVARES J, et al. Time prediction on multi-perspective declarative business processes[J]. *Knowledge and Information Systems*, 2018, 57(3): 655–684.
- [14] POLATO M, SPERDUTI A, BURATTIN A, et al. Time and activity sequence prediction of business process instances[J]. *Computing*, 2018, 100(9): 1005–1031.
- [15] ROGGE-SOLTI A, WESKE M. Prediction of business process durations using non-markovian stochastic petri nets[J]. *Information Systems*, 2015, 54: 1–14.
- [16] ROGGE-SOLTI A, WESKE M. Prediction of remaining service execution time using stochastic Petri nets with arbitrary firing delays[C]// *International Conference on Service-oriented Computing*, 2013: 389–403.
- [17] CECI M, LANOTTE P F, FUMAROLA F, et al. Completion time and next activity prediction of processes using sequential pattern mining[C]// *International Conference on Discovery Science*, 2014: 49–61.
- [18] BEVACQUA A, CARNUCCIO M, FOLINO F, et al. A data-driven prediction framework for analyzing and monitoring business process performances[C]// *International Conference on Enterprise Information Systems*, 2013: 100–117.
- [19] SENDEROVICH A, DI FRANCESCO MARINO C, GHIDINI C, et al. Intra and inter-case features in predictive process monitoring: a tale of two dimensions[C]// *International Conference on Business Process Management*, 2017: 306–323.
- [20] VERENICH I, NGUYEN H, LA R M, et al. White-box prediction of process performance indicators via flow analysis[C]// *Proceedings of the 2017 International Conference on Software and System Process*, 2017: 85–94.
- [21] NI W J, SUN Y J, LIU T, et al. Prediction method of remaining time of business process based on attention bidirectional recurrent neural network[J]. *Computer Integrated Manufacturing System*, 2020, 26(6): 1564–1572.
倪维健, 孙宇健, 刘彤, 等. 基于注意力双向循环神经网络的业务流程剩余时间预测方法[J]. *计算机集成制造系统*, 2020, 26(6): 1564–1572.
- [22] SENDEROVICH A, FRANCESCO MARINO C D, GHIDINI C, et al. Intra and inter-case features in predictive process monitoring: a tale of two dimensions[C]// *International Conference on Business Process Management*, 2017: 306–323.
- [23] LIU T, NI W J, SUN Y J, et al. Prediction method of remaining execution time of business process instance based on deep transfer learning[J]. *Data Analysis and Knowledge Discovery*, 2020, 4(2): 134–142.
刘彤, 倪维健, 孙宇健, 等. 基于深度迁移学习的业务流程实例剩余执行时间预测方法[J]. *数据分析与知识发现*, 2020, 4(2): 134–142.
- [24] NAVARIN N, VINCENZI B, POLATO M, et al. LSTM networks for data-aware remaining time prediction of business process instances[C]// *2017 IEEE Symposium Series on Computational Intelligence*, 2017: 1–7.
- [25] XU X R, LIU C, LI T, et al. Prediction method of remaining time of business process based on bidirectional quasi-recurrent neural network and attention mechanism[J]. *Journal of Electronics*, 2022, 50(8): 1975–1984.
徐兴荣, 刘聪, 李婷, 等. 基于双向准循环神经网络和注意力机制的业务流程剩余时间预测方法[J]. *电子学报*, 2022, 50(8): 1975–1984.
- [26] KANG L, WANG F W, MO H, et al. Research on wind farm management based on centralized control mode[J]. *Yunnan Hydropower*, 2018, 34(3): 159–162.
康磊, 王凡武, 莫浩, 等. 基于集控模式下的风电场管理研究[J]. *云南水力发电*, 2018, 34(3): 159–162.
- [27] BESNARD F. On optimal maintenance management for wind power systems[EB/OL]. <https://core.ac.uk/display/41185063>.
- [28] CHEN L. Design and implementation of Kangbao wind farm integrated management system[D]. Dalian: Dalian University of Technology, 2014.
陈雷. 康保风电场综合管理系统设计与实现[D]. 大连: 大连理工大学, 2014.
- [29] QIN Z, BAO D M, LAI X L, et al. Research on wind power prediction system of wind farm[J]. *Computer Technology and Development*, 2013, 23(7): 73–76.
秦政, 包德梅, 赖晓路, 等. 风电场功率预测系统研究[J]. *计算机技术与发展*, 2013, 23(7): 73–76.
- [30] GUO N, LIU C, LI C, et al. Predictive monitoring method of fan maintenance process based on feature selection[J/OL]. *Computer Integrated Manufacturing System*, 2023–03–24. <http://kns.cnki.net/kcms/detail/11.5946.tp.20230322.1753.044.html>.
郭娜, 刘聪, 李彩虹, 等. 基于特征选择的风机检修流程预测性监控方法[J/OL]. *计算机集成制造系统*, 2023–03–24. <http://kns.cnki.net/kcms/detail/11.5946.tp.20230322.1753.044.html>.
- [31] BAI S, KOLTER J Z, KOLTUN V. An empirical evaluation of generic convolutional and recurrent networks for sequence modeling[DB/OL]. <https://arxiv.org/abs/1803.01271>.
- [32] ZHOU D, LI Z, ZHU J, et al. State of health monitoring and remaining useful life prediction of lithium-ion batteries based on temporal convolutional network[J]. *IEEE Access*, 2020, 8: 53307–53320.
- [33] WU P, SUN J, CHANG X, et al. Data-driven reduced order model with temporal convolutional neural network[J]. *Computer Methods in Applied Mechanics and Engineering*, 2020, 360: 112766.

(责任编辑: 黄健)