文章编号：2097-5899(2025)09-0046-06　**中图分类号**：U29　**文献标识码**：A
DOI：10.16669/j.cnki.issn.2097-5899.202410040001
引用格式：李瑞辰，尚麟宇，刘逸明，等. 技术站列车进路占用时段智能预测方法研究[J]. 铁路物流，2025，43(9)：46-51.
LI Ruichen，SHANG Linyu，LIU Yiming，et al. Intelligent Prediction Method for Train Route Occupation Period in Technical Stations[J]. Railway Logistics，2025，43(9)：46-51.

技术站列车进路占用时段智能预测方法研究

李瑞辰[1]，尚麟宇[1]，刘逸明[1]，杨雨佳[2]

(1. 中国铁道科学研究院集团有限公司 通信信号研究所，北京　100081；2. 北京华铁信息技术有限公司 安全质量管理部，北京　100081)

摘　要：准确地预测出技术站列车接发进路的占用时段，可以为调度员提供辅助决策，对于提升铁路运输行车指挥和运营控制水平，提高车站整体通过能力有着重要意义。首先，分析技术站接车作业流程与发车作业流程，梳理接发车进路占用时段的影响因素；其次，提出智能预测模型的总体方案、算法选择和数据预处理方法，并构建列车进路占用时段智能预测模型；再次，利用实际数据对智能预测模型进行验证，结果表明模型的训练效果较好；最后，结合SAM系统列车进路自动触发功能现状提出了整合优化方案。

关键词：技术站；列车接发；进路占用时段；智能预测；随机森林

0　引言

技术站作为铁路运输网络中的关键节点，承担着铁路货物列车到发、解编、机车换挂等工作[1]。在技术站的同一咽喉区域内，列车进路、调车进路以及本务机车进路相互交叉排列，进路间会在时间和空间上产生冲突，对彼此的作业效率及行车安全造成影响[2]。

在铁路技术站实际生产作业过程中，因车流不均衡、设备故障、施工作业、天气以及人为因素等原因，货物列车到发存在晚点的情况[3]。目前只能依靠经验丰富的行车指挥人员根据站场状态、作业情况、联锁关系等条件估算列车进路占用时间，并结合《车站行车工作细则》(以下简称《站细》)规定的停调时间通过人工干预完成进路的选排。这种依赖人工的方式效率低下且存在安全风险，制约着整个铁路货运枢纽的运输组织和周转效率。

近年来，编组站综合自动化系统(Synthetic Automation of Marshalling Yard，SAM)、编组站综合集成自动化系统(Computer Integrated Process System，CIPS)和调度集中系统(Centralized Traffic Control，CTC)等具备进路自动触发功能的系统广泛应用于技术站，技术站进路冲突干扰问题仍未有效解决，成为制约技术站进一步提升智能化水平的突出瓶颈[4-5]。目前，国内外学者提出了模型分析、仿真模拟和数据驱动3种方法预测未来到达时间，但缺少技术站列车进路占用时段预测方面的研究[6-8]。模型分析和仿真模拟分析方法均需要明确规则声明，而全面考虑实践中影响技术站接发车的错综复杂的因素以及这些动态因素在不同地点和不同时间的变化比较困难[9-10]。以SAM系统为基础，利用数据驱动方法研究列车进路占用时段智能预测方法。

1　接发列车作业及影响因素分析

1.1　接车作业流程分析

SAM系统自动接收调度所下达的列车接车阶段计划。列车接车阶段计划按照类型可分为“终到接车”“中转接车”“通过接车”3种类型。其中，“终到接车”类型列车需要在本站进行技术作业；“中转接车”类型列车在本站不进行技术作业，停留若干时间后继续前行；“通过接车”类型列车不在本站停留直接通过。是

收稿日期：2024-10-04　**修订日期**：2025-06-20
基金项目：中国铁道科学研究院集团有限公司通信信号研究所科研项目(2023TH04)

否在本站进行停车将影响接车进路占用时段的推算，所以可以将接车作业分为“停车”与“通过”2类。

技术站接车作业流程图如图1所示，本站在t_0时刻同意A站办理的发车手续，t_1时刻A站开放发车进路，t_2时刻列车越过A站出站信号，t_3时刻列车占用A站离去区段，t_4时刻本站接车进路开放，t_5时刻列车占用本站接近区段，t_6时刻列车完全进入股道，从t_0到t_6为技术站接车作业的过程及关键时刻。以A站办理发车手续t_0作为技术站列车进路占用时段智能预测推算的起始点，t_1—t_5作为关键调整点，依靠数据挖掘等智能化手段，结合日常积累的大量历史数据，就可以精准推算出各作业过程所需时长(T_1，T_2，T_3，T_4，T_5，T_6)，最终得到接车进路在本站的占用时段[t_4，t_6]。

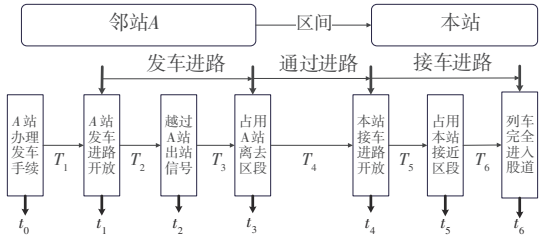


图1　技术站接车作业流程图

Fig.1　Train receiving operation diagram of technical stations

通过技术站现场调研及分析，梳理出接车作业影响因素如表1所示。

1.2　发车作业过程分析

SAM系统自动接收调度所下达的列车发车阶段计划，按照类型可分为“站编始发”“中转发车”“通过发车”3种类型。其中，“站编始发”类型列车在本站编组成列后发车；“中转发车”类型列车在本站不进行相关改编作业，停留若干时间后继续前行；“通过发车”类型列车不在本站停留直接通过。是否在本站进行停车将影响接车进路占用时段的推算，所以可以将接车作业分为“始发”与“通过”2类。

技术站发车作业流程如图2所示，本站在t_7时刻办理发车手续，然后t_8时刻本站开放发车进路，t_9时刻列车越过本站出站信号，t_{10}时刻列车占用本站离去区段，从t_7到t_{10}为技术站发车作业的过程及关键时刻。以办理发车手续t_7作为技术站列车进路占用时段智能预测推算的起始点，t_8—t_9作为关键调整点，依靠数据挖掘等智能化手段，结合日常积累的大量历史数据，就可以精准推算出各作业过程所需时长(T_7，T_8，T_9)，最终得到发车进路在本站的占用时段[t_8，t_{10}]。

通过技术站现场调研及分析，梳理出发车作业影响因素如表2所示。

表1　接车作业影响因素

Tab.1　Influencing factors of train receiving operation

自变量	影响因素	变量类型	说明
J_1	A站发车类型	分类特征	始发、通过
J_2	A站发车进路	分类特征	进路ID
J_3	列车类型	分类特征	A类：城际、动车 B类：直达、特快、快速、普通旅客列车 C类：路用车、单机、通勤车 D类：行包车、普通货物列车
J_4	列车总重	数值特性	单位：t
J_5	列车辆数	数值特性	单位：辆
J_6	限速条件	分类特征	限速、不限速
J_7	天气条件	分类特征	良好、一般、较差、恶劣
J_8	接车进路	分类特征	进路ID
J_9	接车进路自触	分类特征	自触、取消自触
J_{10}	接车进路类型	分类特征	停车、通过

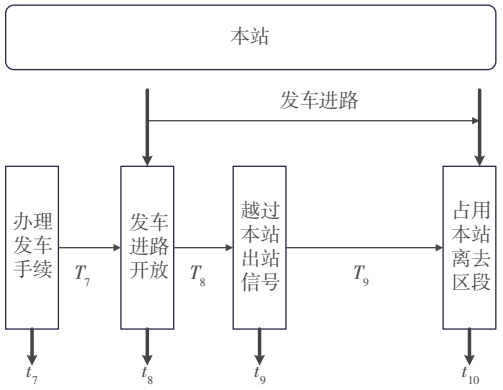


图2　技术站发车作业流程图

Fig.2　Train departure operation diagram of technical stations

2　智能预测模型设计

2.1　总体方案

铁路技术站相关系统在日常运用过程中存储了海量的历史运用数据，当系统中的一些问题无法构建简单合适的数学模型解决时，可以运用数据驱动技术，利用大规模的历史数据针对现实问题构建近似模型进行拟合，即通过数据价值和机器算力换取复杂问题的研究时间[11]。运用数据驱动技术可以从历史数据中挖掘出这些信息来解释列车时空轨迹的规律特征，并在

表2　发车作业影响因素

Tab.2　Influencing factors of train departure operation

自变量	影响因素	变量类型	说明
F_1	发车类型	分类特征	始发、通过
F_2	发车进路	分类特征	进路ID
F_3	列车类型	分类特征	A类：城际、动车 B类：直达、特快、快速、普通旅客列车 C类：路用车、单机、通勤车 D类：行包车、普通货物列车
F_4	列车总重	数值特性	单位：t
F_5	列车辆数	数值特性	单位：辆
F_6	限速条件	分类特征	限速、不限速
F_7	天气条件	分类特征	良好、一般、较差、恶劣

此基础上预测最终的列车进路占用时段。

在进行数据预测并优化决策的过程中，将数据驱动技术的7个步骤对数据驱动的研究进行总结，形成了3个关键过程：首先进行数据驱动相关算法的选择，然后对数据进行预处理，最后进行智能推算模型的构建。智能预测总体方案如图3所示。

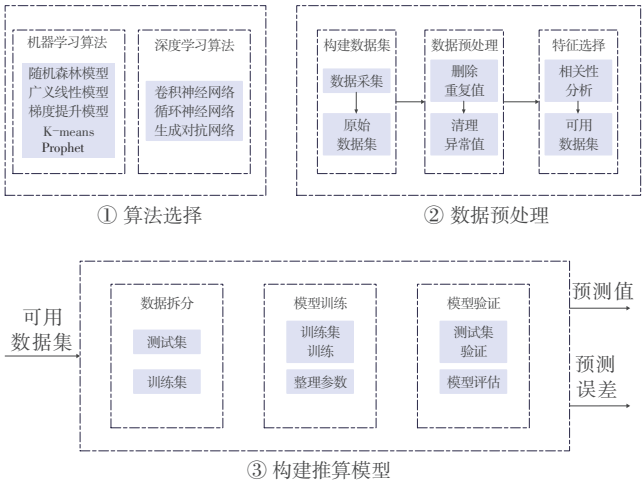


图3　智能预测总体方案

Fig.3　Overall scheme of intelligent prediction

2.2　算法选择

在预测模型领域中，机器学习算法与深度学习算法是两大主流算法，它们各自拥有独特的应用方向[12]。机器学习算法因其强大的大数据处理能力，在大数据模型预测方面得到了广泛应用；而深度学习算法则在处理文本、图形等复杂数据方面表现出色，尤其是在图像识别领域，卷积神经网络凭借其卓越的性能，能够精确识别各种物体。然而，在诸如技术站晚点预测等未来时空模型预测方面，机器学习算法，特别是随机森林算法，因其独特的优势而被广泛采用[11]。

选用随机森林算法，主要基于以下几个方面的考虑。首先，随机森林通过采用多棵树的运算方式，有效降低了单棵树的误差，从而提高了预测的准确性和稳定性；其次，该算法具有出色的抗过拟合能力，能够在处理大量数据的同时，避免模型过度拟合训练数据而导致的预测性能下降；最后，对于多变量预测问题，随机森林能够估计变量的重要性，这不仅有助于理解数据特征对预测结果的影响，还能在数据丢失时保持较高的预测准确性。

2.3　数据预处理

（1）构建数据集。SAM系统实时跟踪报点服务器可以通过采集站场表示信息，并结合列车车次窗信息搜索列车作业进路，采集进路内区段的占用、出清信息并形成包含进路的跟踪结果信息，同时对作业轨迹跟踪的结果数据和执行的作业计划进行关联，得到列车的相关信息，并将跟踪结果信息保存于数据库内，形成原始数据集。原始数据集需要进行数据清洗和检查，对缺失、异常和重复等信息进行删除和修正，得到可用的基础数据集。

（2）归一化处理。根据不同影响因素的特点，对数据集进行标准化、归一化处理。将影响因素划分为数值特征影响因素与分类特征影响因素。数值特征是特征取值为数字的一类特征，如列车总重、列车辆数等。需要对特征进行归一化，以消除特征间单位和尺度差异的影响，同时提高模型的性能、收敛速度和稳定性，并减少取值过大的特征对训练模型精度的影响程度。利用最小－最大归一化方法(Min-Max Scaling)对数据集进行处理，即将特征值缩放成[0，1]范围内，并在预测结束后进行反归一化处理，最小－最大归一化方法及逆过程的基本公式如下。

$$X_{\text{normalized}}=\frac{X-X_{\min}}{X_{\max}-X_{\min}} \tag{1}$$

$$X'=X_{\text{normalized}}\times(X_{\max}-X_{\min})+X_{\min} \tag{2}$$

式中：X为原始数据；$X_{\max}$和$X_{\min}$表示该特征因素的最大值和最小值；$X_{\text{normalized}}$为归一化后的数据；X'为逆过程处理后得到的原数据。

分类特征是特征取值为不同类别的一类特征。不同的分类特征，其取值通常为具有不同含义的字符，为使训练模型可以识别其含义并进行处理运算，需进行独热特征编码(One-Hot编码)。

（3）特征选择。将接车作业过程分析出的影响因素J_1—J_{10}以及上一阶段用时T_{n-1}作为接车作业特征变量，将发车作业过程分析出的影响因素F_1—F_7以及上一阶段用时T_{n-1}作为发车作业特征变量，分别对特征变量及预测目标值进行Spear-man相关性分析，选取与预测目标值相关性显著的特征变量作为预测模型的输入变量，得到多个特征变量与预测目标值的可用数据集，因变量与特征变量如下。

$$T_n \sim \begin{cases} [J_1,J_2,J_3,J_4,J_5,J_6,J_7,J_8,J_9,J_{10},T_{n-1}] & 1<n\leqslant 6 \\ [F_1,F_2,F_3,F_4,F_5,F_6,F_7,T_{n-1}] & 6<n\leqslant 9 \end{cases} \tag{3}$$

2.4　构建推算模型

利用随机森林回归的思想，构建基于随机森林的智能预测模型来得到最终的预测值，随机森林模型预测过程如图4所示，主要包含3部分内容。

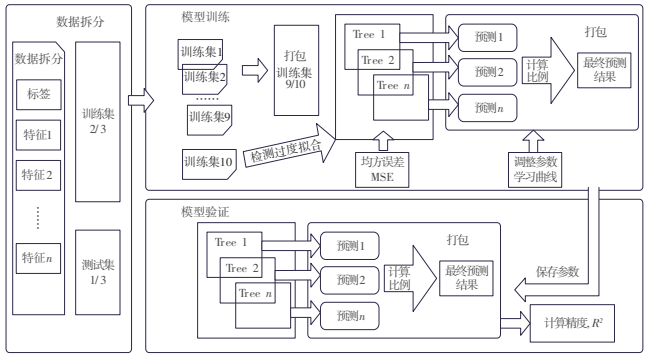


图4　随机森林模型预测过程

Fig.4　Prediction process of random forest model

（1）数据拆分。采用Bootstrap抽样方法，从原始数据集中随机抽取多个样本，每个样本都可以重复抽取。利用该方法将可用的数据集按照2∶1的比例有效地划分为训练集和测试集。训练集的主要作用是用于训练并建立模型，而测试集则用于评估已建立模型的预测性能，以确保模型的准确性和可靠性。

（2）模型训练。在模型训练过程中，训练集中的大部分数据被直接用于模型的拟合，通过构建多棵决策树来共同构成随机森林模型。同时，为了检测模型是否存在过拟合现象，预留出训练集中大约10%的数据作为验证集。每棵决策树在构建时，都会随机地从n个特征中选择一个固定数量的特征子集进行分裂，并且这些树都会尽可能地生长，不进行剪枝操作。这样，每棵树都会根据所选的特征子集和训练数据得出自己的预测结果。

在随机森林模型中，最终的预测结果是基于所有决策树的预测结果进行投票或平均得出的。对于回归问题，通常采用平均预测值作为最终预测结果。为了获取最佳模型参数，会根据学习曲线不断调整参数，并观察模型性能的变化。在评价回归模型的质量时，选用均方误差(MSE)作为标准，计算公式为

$$MSE=\frac{1}{m}\sum_{1}^{m}\left(y_{f_i}-y_{c_i}\right)^2 \quad i=1,2,\cdots,n \tag{4}$$

式中：m为样本数量；y_{f_i}为真实值；y_{C_i}为预测值。

在决策树每一个节点的选择中，MSE更小的节点将被视为回归质量高的节点。

（3）模型验证。测试集用于评估已经通过训练集训练好的模型的性能。为了确保评估结果的公正性和准确性，避免因为模型对测试数据的过拟合而导致评估结果偏高，不使用测试集来训练或调整模型，而是直接使用在训练阶段已经确定并保存下来的决策树和相应的参数。

当测试数据集通过已经训练好的模型进行预测后，可以得到一系列的预测结果。这些预测结果与实际的真实值进行比较，就可以计算出模型的精度。为了量化模型的性能，选用确定系数R^2(也称为决定系数或拟合优度)作为评估指标。计算公式为

$$R^2=1-\frac{\sum_{1}^{n}\left(\widehat{y_i}-y_i\right)^2}{\sum_{1}^{n}\left(\overline{y_i}-y_i\right)^2} \quad i=1,2,\cdots,n \tag{5}$$

式中：n为数据集的数据量；y_i为原始数据；$\widehat{y_i}$为预测数据；$\overline{y_i}$为原始数据均值。

R^2的取值范围为[0，1]，如果结果是0，说明模型拟合效果很差，如果结果是1，说明模型无错误。

3　模型验证及应用

3.1　模型验证

采集某铁路技术站2023年5月1日至2024年4月30日间由A站开来经某接车口终到本站3道的列车接车历史数据856条，经过数据预处理得到可用数据820条；由本站5道始发经某发车口开向B站的列车发车历史数据832条，经过数据预处理得到可用数据813条，模型验证结果如表3所示。

如果训练模型R^2值大于0.8，说明模型的训练效果比较好，根据表3的结果可以看出，训练集和测试集的R^2均接近1，说明模型误差很小，模型可以应用。

3.1.1　模型应用

将列车进路占用时段智能预测模型与SAM系统列车进路自动触发功能进行整合，将预测结果提供给SAM系统进行必要的自律运算和安全卡控。模型应用方案如图5所示。

表3　模型验证结果

Tab.3　Model verification results

特征条件	标签	平均值/min	R^2_TRAIN	R^2_TEST
J_1=通过	T_1	3.95	0.997	0.981
J_3=D类	T_2	2.52	0.995	0.960
J_6=不限速	T_3	5.25	0.996	0.972
J_7=良好	T_4	13.40	0.976	0.994
J_9=自触	T_5	1.23	0.982	0.911
J_{10}=停车	T_6	6.13	0.950	0.920
F_1=始发 F_3=D类 F_6=不限速 F_7=良好	T_7	1.50	0.987	0.943
	T_8	2.93	0.994	0.922
	T_9	4.25	0.989	0.956

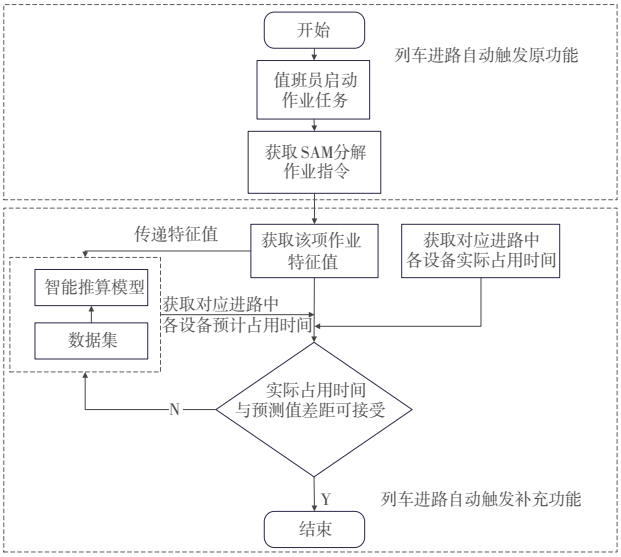


图5　模型应用方案

Fig.5　Model application scheme

（1）值班员在SAM系统上选定列车作业任务进行启动操作后，系统对列车作业计划进行分解，得到作业进路指令。

（2）将该项作业的特征值，传递给智能推算模型中得到每条进路的占用时段预测值。

（3）当进路排列完成，进行相应作业，并在列车作业完成后，记录实际列车作业时间，与预测值进行比对，计算误差率(可接受误差率范围由专家设定)，判断误差能否接受。

（4）如果误差率可接受，则结束预测；如果不满足误差率要求，则将数据更新至训练集对预测模型进行重新训练，得到优化后新的预测模型。

3.2　应用效果分析

2024年，中国铁路上海局集团有限公司某技术站上行作业场日均办理接车作业140列，发车作业162列，调车作业12项，本务机车作业50项。该技术站《站细》明确规定：接车时各方向不迟于列车到达前4 min开放接车进路信号，接发特快旅客列车时，不迟于列车到达、出发前8 min开放接车进路、发车进路信号。各方向不迟于列车到达前8 min，出发前5 min停止影响列车进路的调车作业。接发特快旅客列车时，不迟于列车通过前10 min停止调车作业。

预计实施前后各项指标对比如表4所示。目前只能依靠经验丰富的行车指挥人员根据站场状态、作业情况、联锁关系等条件估算列车进路占用时间，通过人工干预完成进路的选排。技术站作业系统实施列车进路占用时段智能预测后，将大幅降低行车作业人员的劳动强度，提高车站运输生产效率，为铁路技术站智能化提升打下坚实的基础。

表4　预计实施前后各项指标对比

Tab.4　Comparison of operation process before and after implementation

序号	变更内容	实施前	实施后	前后对比
1	推算进路占用时段	值班员凭经验选排	系统自动生成，值班员盯控	结果更精准
2	停调时机	值班员掌握	系统自动卡控	提高效率
3	进路排列	过多人工干预	自动化排列	提高自动化水平
4	作业风险	高	低	降低风险

4　结束语

深入梳理铁路技术站生产作业流程以及列车接发作业过程，确定影响进路占用时段的关键因素，以进路为单位进行细化分析，构建列车接发车进路智能预测模型，并以SAM系统日常积累的海量历史数据对智能预测进行训练验证，效果良好。目前正在进行智能预测模型与SAM系统的融合调试工作，预计应用后可以有效提升铁路运输行车指挥和运营控制水平，减轻行车指挥人员劳动强度并提升技术站的智能化水平。

参考文献：

[1]　李博宇，刘启钢，孙文桥，等. 5G技术与铁路技术站作业的融合应用探讨[J]. 铁道货运，2022，40(3)：15-19.
LI Boyu，LIU Qigang，SUN Wenqiao，et al. Discussion on the

Integration and Application of 5G Technology and Railway Station Operation[J]. Railway Freight Transport，2022，40(3)：15-19.

[2] 张　博，苗建瑞，豆　飞，等.基于时空状态网络的铁路中小型技术站进路不确定性规划研究[J].铁道运输与经济，2025，47(3)：140-150，160.
ZHANG Bo，MIAO Jianrui，DOU Fei，et al. Research on Uncertainty Planning of Routes in Small and Medium-sized Railway Technical Stations Based on Spatiotemporal State Network [J]. Railway Transport and Economy, 2025, 47(3): 140-150，160.

[3] 王　健，马　灿，苏阳平，等.基于SAM系统缩短株洲北站北发线行车间隔的研究[J].铁道通信信号，2021，57(6)：50-53.
WANG Jian，MA Can，SU Yangping，et al. Research on Minimizing Headway of Trains of Northbound Departure Track of Zhuzhou North Station Based on SAM System[J]. Railway Signalling & Communication，2021，57(6)：50-53.

[4] 刘　俊.普速铁路调度集中控制系统运用优化对策探讨[J].铁道运输与经济，2019，41(8)：1-5.
LIU Jun. A Tentative Study on the Optimizing Countermeasures for the Application of Centralized Traffic Control System of Conventional Railway[J]. Railway Transport and Economy，2019，41(8)：1-5.

[5] 张从立.提升徐州北编组站运输效率的思考与实践[J].铁道货运，2023，41(2)：1-5，27.
ZHANG Congli. Improvement of Transport Efficiency of Xuzhoubei Marshalling Station：Reflection and Practice[J]. Railway Freight Transport，2023，41(2)：1-5，27.

[6] 黄　平.数据驱动的高速铁路列车晚点传播机理及模型研究[D].成都：西南交通大学，2020.

[7] 蒋哲远，葛承宇，陈　超，等.基于xDeepFM的铁路货物运输时间预测[J].交通运输工程与信息学报，2022，20(1)：39-46，97.
JIANG Zheyuan，GE Chengyu，CHEN Chao，et al. Prediction of Rail Freight Time Using xDeepFM[J]. Journal of Transportation Engineering and Information，2022，20(1)：39-46，97.

[8] 丁洪轩.基于数据驱动的铁路集装箱运输时间预测研究[D].北京：北京交通大学，2023.
DING Hongxuan. Research on Railway Container Transportation Time Prediction Based on Data-Driven [D]. Beijing: Beijing Jiaotong University, 2023.

[9] 文　超，李　津，李忠灿，等.机器学习在铁路列车调度调整中的应用综述[J].交通运输工程与信息学报，2022，20(1)：1-14.
WEN Chao，LI Jin，LI Zhongcan，et al. Review on Application of Machine Learning Methods to Railway Train Operation Adjustment[J]. Journal of Transportation Engineering and Information，2022，20(1)：1-14.

[10] 邓　薪.基于随机森林算法的铁路货物运到时间预测研究[D].北京：北京交通大学，2020.
DENG Qi. Research on Railway Freight Arrival Time Prediction Based on Random Forest Algorithm [D]. Beijing: Beijing Jiaotong University, 2020.

[11] 王建英.智能调度集中系统研究[J].铁道通信信号，2022，58(7)：7-12.
WANG Jianying. Study on Intelligent Centralized Traffic Control System[J]. Railway Signalling & Communication，2022，58(7)：7-12.

[12] 周书乐，张子群，周家梁.提高车站通过能力的接近锁闭方法研究[J].铁道运输与经济，2023，45(12)：36-43.
ZHOU Shule，ZHANG Ziqun，ZHOU Jialiang. Variable Approach Locking Method for Improving Station Passing Capacity[J]. Railway Transport and Economy，2023，45(12)：36-43.

[13] 张天祥，张　巍，王道奇，等.车流径路调整下编组计划与技术站作业能力及作业效率耦合机理研究[J].铁道运输与经济，2025，47(4)：117-127.
ZHANG Tianxiang，ZHANG Wei，WANG Daoqi，et al. Coupling Mechanism between Formation Plan and Technical Station Operation Capacity and Efficiency under Train Route Adjustment[J]. Railway Transport and Economy，2025，47(4)：117-127.

（责任编辑：汪沁萱）

Intelligent Prediction Method for Train Route Occupation Period in Technical Stations

LI Ruichen[1]，SHANG Linyu[1]，LIU Yiming[1]，YANG Yujia[2]

(1. Signal & Communication Research Institute，China Academy of Railway Sciences Corporation Limited，Beijing 100081，China；
2. Safety Quality Management Department，Beijing Huatie Information Technology Corporation Limited，Beijing 100081，China)

Abstract: When the occupation period of train receiving and dispatching routes is accurately predicted in technical stations, it can provide auxiliary decision-making for dispatchers. This is significant for improving railway transportation command and operation control levels, as well as enhancing the overall carrying capacity of stations. Therefore, this paper first analyzed the process of train receiving and dispatching operations in technical stations to sort out the factors influencing the occupation period of receiving and dispatching routes. Second, this paper proposed the overall scheme, algorithm selection, and data preprocessing methods of the intelligent prediction model. An intelligent prediction model was also constructed for the train route occupation period. Then, actual data was used to verify the intelligent prediction model, and the results show an ideal training effect of the model. Finally, a scheme for integration and optimization was proposed based on the current status of the automatic route trigger function in the synthetic automation of marshalling (SAM) system.

Keywords: Technical Station; Train Receiving and Departure; Route Occupancy Period; Intelligent Prediction; Random Forest