

文章编号:1672-4747(2022)01-0001-14

## 机器学习在铁路列车调度调整中的应用综述

文超<sup>1</sup>,李津<sup>2</sup>,李忠灿<sup>1</sup>,智利军<sup>3</sup>,田锐<sup>3</sup>,宋邵杰<sup>4</sup>

(1. 西南交通大学, 交通运输与物流学院, 成都 611756; 2. 中土集团福州勘察设计研究院有限公司, 福州 350013; 3. 中国国家铁路集团有限公司调度指挥中心, 北京 100844; 4. 中国铁路广州局集团有限公司调度所, 广州 510088)

**摘要:**随着铁路信息化进程的加快,铁路大数据与人工智能技术有力地支撑了我国铁路的高质量发展。本文回顾了近10年机器学习方法在铁路列车调度调整领域的主要研究成果,将列车调度调整问题分为列车晚点状态分析与评估、列车晚点传播预测和列车运行调整智能化决策三个方面,分别总结和分析了机器学习方法在上述三方面的应用情况。在列车晚点状态分析与评估方面,既有研究主要集中于传统统计分析,其描述和预测性能往往有限。在列车晚点传播预测方面,传统机器学习和深度学习方法被应用于晚点致因、晚点持续时长、列车晚点状态演化及晚点恢复预测问题的建模。在列车运行调整智能化决策方面,既有研究主要侧重于运用强化学习、模糊神经网络方法建模,机器学习方法应用于列车调度辅助决策系统仍是研究的主要难点和关键。在归纳了既有研究特征的基础上,展望了机器学习方法在铁路调度调整研究方面的最新动向。以深度学习为代表的高级机器学习应用于列车调度调整智能决策将是未来的发展重点。

**关键词:**智能铁路;列车运行调整;列车晚点;机器学习

中图分类号:U292

文献标志码:A

DOI:10.19961/j.cnki.1672-4747.2021.04.021

## Review on Application of Machine Learning Methods to Railway Train Operation Adjustment

WEN Chao<sup>1</sup>, LI Jin<sup>2</sup>, LI Zhong-can<sup>1</sup>, ZHI Li-jun<sup>3</sup>, TIAN Rui<sup>3</sup>, SONG Shao-jie<sup>4</sup>

(1. School of Transportation & Logistics, Southwest Jiaotong University, Chengdu 611756, China; 2. CCECC Fuzhou Survey & Design Institute Co. Ltd., Fuzhou 350013, China; 3. Dispatching Center, China National Railway Group Limited, Beijing 100844, China; 4. Dispatching Department, China Railway Guangzhou Group Co. Ltd., Guangzhou 510088, China)

**Abstract:** With the rapid development of railway informatization, big data and artificial intelligence technology are providing strong support for the high-quality development of China's railway system. The major studies of the application of machine learning methods to railway train operation adjustment in the past 10 years are reviewed. The problem of train operation adjustment is divided into three topics: analysis and evaluation of the train delay state, prediction of train delay propagation, and train operation adjustment and intelligent decision-making. The application of machine learning methods in these three topics is summarized and analyzed. Existing research on train delay analysis and evaluation mainly focuses on traditional statistical analysis, and its description and prediction performance are often limited. Traditional machine learning and deep learning methods have been ap-

收稿日期:2021-04-16

录用日期:2021-09-18

网络首发:2021-09-24

审稿日期:2021-04-16~04-17; 05-30~06-02; 07-09~07-16; 09-11~09-18

基金项目:国家自然科学基金项目(71871188, U1834209);中国国家铁路集团有限公司科技研究计划项目(P2020X016);山东省自然科学基金项目(ZR2019PG008)

作者简介:文超(1984—),男,博士,副教授,研究方向为铁路运输组织优化、铁路大数据应用,E-mail:wenchao@swjtu.cn

引文格式: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 Zhong-can, 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.

plied to the modeling of the causes and duration of train delay, the evolution of train delays, and the prediction of train delay recovery. The existing research on the intelligent decision-making for train operation adjustment mainly focuses on applying reinforcement learning and fuzzy neural networks to establish models, while the application of machine learning in train dispatching support systems development remains the main difficulty and the key problem. The characteristics of the existing research are summarized. On this basis, this paper looks forward to the latest trend of using machine learning methods in the research of railway train operation adjustment. The application of advanced machine learning methods represented by deep learning in intelligent decision-making for train operation adjustment will be highly focused in the future.

**Key words:** intelligent railway; train operation adjustment; train delay; machine learning

## 0 引言

铁路调度指挥是铁路运营管理和列车运行控制的中枢,担负着组织指挥铁路列车运行和日常生产活动的重要任务。列车调度员需要根据各种实时信息和预定的规则来实现相应的调整目标,而基于调度员经验的传统人工为主的调度方法已经难以适应当前复杂的运输组织需求。智能铁路已成为世界铁路发展的重点方向,智能行车调度是我国铁路当前及未来一定时期内发展的关键技术。中国工程院“智能高铁战略研究(2035)”重大咨询项目认为:智能行车调度是我国智能高铁中期(2021—2025)要实现的目标<sup>[1]</sup>。为了保障铁路运输服务的安全性、连续性、高效性与准时性,亟需高效准确的决策支持方法来提升调度员决策水平和效率。

实现海量数据可挖掘、设备状态可诊断、行车安全可预警、复杂路网运营变化可感知、发展趋势可推断、辅助决策可支撑,提升铁路的智能决策水平,提高决策效率是智能铁路亟待解决的关键科学问题,是“交通强国、铁路先行”战略的迫切需要和必然要求。数据科学及人工智能给解决铁路运输组织大规模和实时性复杂决策支持问题带来了新的契机。机器学习方法作为人工智能的核心,无需以先验知识为基础,而是从数据中发现规律并构造模型来逼近铁路运输生产实际情况,已经在铁路运输领域凸显了优势<sup>[2,3]</sup>。推理与推荐系统、计算机视觉、知识表达与模式识别等人工智能方法和技术已经在我国铁路运输安全、客票、旅客服务等领域取得了初步应用。

列车调度调整一直以来都是制约调度指挥自动化的瓶颈问题<sup>[4]</sup>。目前普遍认为基于机器学习方法的调度决策方案比经验决策和传统数学优化模型更具实用性和科学性<sup>[5,6]</sup>,但当前世界铁路调

度指挥智能化还有很长的路要走。加速机器学习方法在铁路列车调度指挥中的应用进程,提高调度决策的智能化和自动化水平,将有效助力我国铁路行业发展。为全面梳理以机器学习方法在铁路列车调度指挥中的研究现状,明确未来研究的重点和方向,本文系统总结了近10年来机器学习方法在铁路列车调度调整方面的研究成果,以期对相关研究及发展实践提供一定参考。

## 1 相关基本概念

### 1.1 铁路列车调度调整流程

在列车实际运行过程中,铁路系统内外部的各种随机因素干扰,会导致列车运行的实际状态偏离预定值、运行秩序紊乱的情况时有发生。列车调度员是管辖区段行车工作的统一指挥者,负责及时收集并分析晚点列车相关的各种信息,根据列车运行态势分析与评估列车晚点原因和状态,预测和估计列车晚点演化态势,并根据经验和规则制定相应的列车运行调整策略,组织列车尽快恢复正常运行秩序。图1所示为铁路列车调度调整的一般流程。

### 1.2 铁路列车调度调整的支持数据

如图1所示,列车调度员在进行调度调整及制定列车运行调整决策时,需要以海量的列车运行信息为基础。图2所示为铁路列车调度调整的数据,铁路列车运行控制系统、行车安全监控系统等调度指挥边界接口系统(图中①)采集并提供基础数据,提取列车运行实时状态、固定和移动设备运用状态、列车运行环境信息等调度决策信息(图中②),上述信息被汇总到铁路调度指挥系统(图中③)供调度员决策使用。综合分析、语义识别等方法已经广泛应用于铁路列车调度数据的获取、筛

选、聚合、可视化等数据融合过程<sup>[7-9]</sup>。

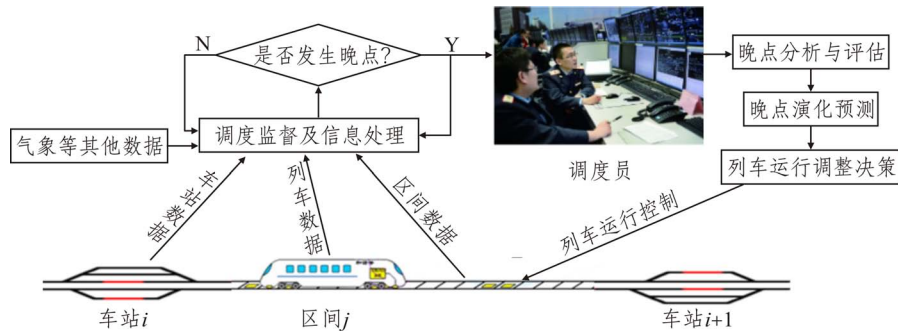


图1 铁路列车调度调整的一般流程

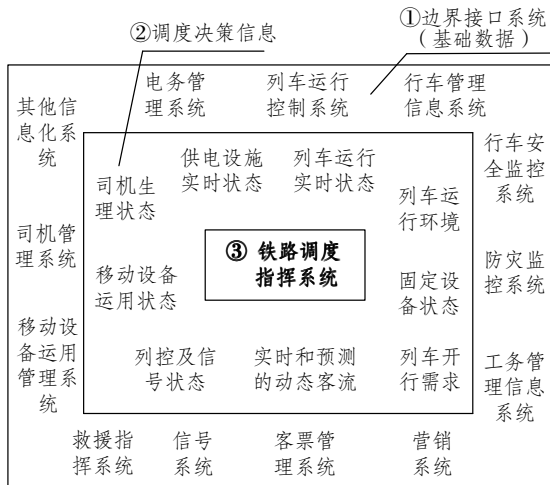


图2 铁路列车调度调整的数据

### 1.3 机器学习方法分类

机器学习方法是实现人工智能的重要途径,被广泛应用于多个领域,其核心目标是使用算法来解析数据,进而基于数据模拟人类的学习、决策和预测行为。机器学习的主要研究内容为让机器自动从数据中学习规则并不断改善自身性能。根据学习策略的不同,机器学习方法一般可以分为有监督学习、无监督学习、强化学习<sup>[10]</sup>。深度学习是机器学习领域中一个新的研究方向,其概念源于模拟人脑进行分析学习的神经网络研究。

有监督学习的数据集包含了样本特征与标签,在训练过程中,需要选择合适的函数模型,通过计算模型的预测值与真实值之间的误差来优化模型参数。常见的有监督学习算法有线性回归、逻辑回归、支持向量机、支持向量回归、随机森林等。无监督学习的数据样本集仅包含特征,算法需要自行发现数据的模态来优化模型。常见的无监督学习算法有主成分分析、K均值算法、关联分析算法等。强化学习也称为增强学习,其算法通

过与环境的交互来学习解决问题的策略,进而达成回报最大化或实现特定目标。常见的强化学习算法有Q-learning、近端策略优化算法等。近年来,学者们提出了大量先进的深度学习算法,并在诸多领域取得了远超先前机器学习技术的分析效果,常见的深度学习算法有卷积神经网络、循环神经网络、递归神经网络等。

## 2 机器学习方法在铁路调度调整中的应用

本文重点分析机器学习方法在调度调整三个主要方面的应用研究,具体为:

(1)列车晚点状态分析与评估。旨在评估和阐释列车历史运行信息,使用数据聚合和数据挖掘技术,帮助调度员了解列车运行晚点的一般规律。

(2)列车晚点传播预测。着重于运用机器学习方法研究铁路列车晚点的传播过程,包括列车晚点状态和恢复的预测。

(3)列车运行调整智能化决策。综合相关行车调度业务规则、机器学习算法和计算建模程序实现调度决策的自动制定。

### 2.1 列车晚点状态分析与评估研究

通过对线路晚点规律的掌握,列车调度员可以在晚点发生时正确识别晚点恢复的关键影响因素,从而制定相应的对策,实现列车延误的快速消解。晚点状态分析与评估是列车晚点预测及列车运行调整的理论基础,其主要研究内容为:基于铁路列车运行实绩数据,提取并分类统计列车晚点、初始晚点、连带晚点,运用描述性统计、绘制分布曲线、关联、聚类等方法探索数据蕴含的规律,从宏观上探明高速列车晚点分布的基本规律,如列车的晚点时空、时长分布规律等,并将这些规律应用于列车运行过程控制及晚点预测中,调度员可

以根据列车在区间的历史和实时状态以及列车在后续车站、区间的晚点分布规律,对列车的运行状态进行预判。

现代统计模型是最简单和基本的统计机器学习方法,也是晚点分析最常用的方法,其目的在于对数据进行聚合和总结,帮助调度员直观地了解特定车站、列车或区间的晚点总体或详细信息。描述性统计方法<sup>[11,12]</sup>、分布拟合<sup>[13,14]</sup>等都被应用于晚点致因、晚点分布、连带晚点的建模。基于列车运行实绩数据,对数正态分布、韦伯分布、伽马分布等统计模型被用于拟合高速列车的晚点致因及影响列车数和影响总时间<sup>[15,16]</sup>。

为了挖掘深层次的列车晚点规律,一些更高级的机器学习方法被应用于晚点分析中。Marković等<sup>[17]</sup>率先使用支持向量回归建立列车晚点和铁路系统的各种特征(基础设施、时刻表和列车)的功能关系,旨在帮助调度员评估铁路系统的各种变化对列车晚点的影响。Murali<sup>[18]</sup>基于模拟的列车晚点数据,提出了一种晚点估计方法,该方法定义了列车晚点与列车组合、运行参数和网络拓扑之间的函数关系。Lee等<sup>[19]</sup>提出了基于机器学习和数据挖掘的监督决策树方法,用于发现影响连带晚点的关键因素,该模型由数据预处理与分析、决策树构建、关键晚点因子的分析、延迟时空拓扑分析四个阶段组成。Cerreto等<sup>[20]</sup>使用K-Means聚类方法识别了丹麦哥本哈根以北的一条高铁线路周期性晚点模式,通过分析经常性晚点列车的运行规律,查找列车重复晚点原因。张琦等<sup>[21]</sup>提出了一种高速铁路列车连带晚点的特征识别方法,该方法综合考虑了连带晚点的实际值、预测值和影响值等因素,将列车晚点分为严重、潜在、消散、一般四种类型,提高了列车连带晚点的辨识度。

列车的运行过程是一个时变的过程,受到大量外界因素的干扰以及行车人员决策的双重影响。列车晚点的分布具有较大的随机性,但从历史数据能够挖掘相关设备、天气、作业组织造成晚点的相应概率是关键,这样就能推导并形成列车在相关影响因素下的一般规律,从而为行车指挥提供一定的指导。由于概率统计模型多是基于单变量分析,且模型形式也比较简单,因此其描述和预测性能往往很有限。目前,主要是运用现代统计、聚类等机器学习方法研究列车晚点状态分析

与评估问题,未见神经网络、深度学习等高级机器学习方法的应用。

## 2.2 列车晚点传播预测研究

### 2.2.1 列车晚点传播过程分析

晚点传播预测是研究者非常关注的问题,其主要内容为:根据列车的历史和当前运行状态,通过研究事件和时间驱动下列车晚点状态的演化过程,预测列车未来在各站的到达/出发时刻以及区间运行时间。准确的晚点传播预测可以帮助调度员更好地预判和估计列车运行态势、相关调度决策的预期效果、列车晚点可恢复的程度等,进而通过制定合理的调度决策更高效地实现调整目标。当前列车晚点传播预测的研究主要集中在晚点致因及持续时长预测、晚点状态演化预测、晚点恢复预测等几个方面,也是国内外学者研究铁路列车调度问题时最热衷、成果产出最丰富的领域<sup>[22-24]</sup>。

晚点发展与列车运行环境的影响密切相关,晚点传播的过程受到当前列车的晚点程度、可利用的冗余时间以及调度工作人员采取的运行调整策略三方面的影响。当调度工作人员采取积极的运行调整措施时,可以有效地缓解晚点程度以达到晚点时长缩短甚至回归正点状态。

列车晚点传播除了上述纵向传播过程外,还具有横向影响传播特征。列车晚点状态通过列车间的相互制约关系横向传播,前行列车的晚点导致后行列车的连带晚点<sup>[14]</sup>。要研究晚点的传播过程并实现晚点的预测,需要先探明列车状态间的时空依赖关系,从已知列车运行实绩中建立晚点传播模型,预测未来的晚点状态。图3(a)为列车状态横纵向影响的示意图,图中时刻 $t$ 对应的纵轴左边是已经兑现的运行图,列车 $n-1$ 在车站 $S_2$ 的晚点可能引起列车 $n-1$ 在 $S_3$ 、 $S_4$ 及以后各站的晚点(纵向传播),也可能引起列车 $n$ 及后行列车在各站的晚点(横向传播),此时需要根据列车 $n-1$ 及其前行列车的已知状态去预测列车 $n-1$ 及其后行列车的晚点状态;图3(b)中时间轴右侧的晚点状态预测需同时考虑其前方多状态的影响,图中箭头为列车状态的顺向推导和演化,分别代表了列车晚点的横向传播和纵向传播的迭代和推演过程。近年来,机器学习方法被广泛应用于捕捉列车间的相互制约关系、列车运行状态与运行环境之间的关系、列车运行时间-空间作用过程等研究中。

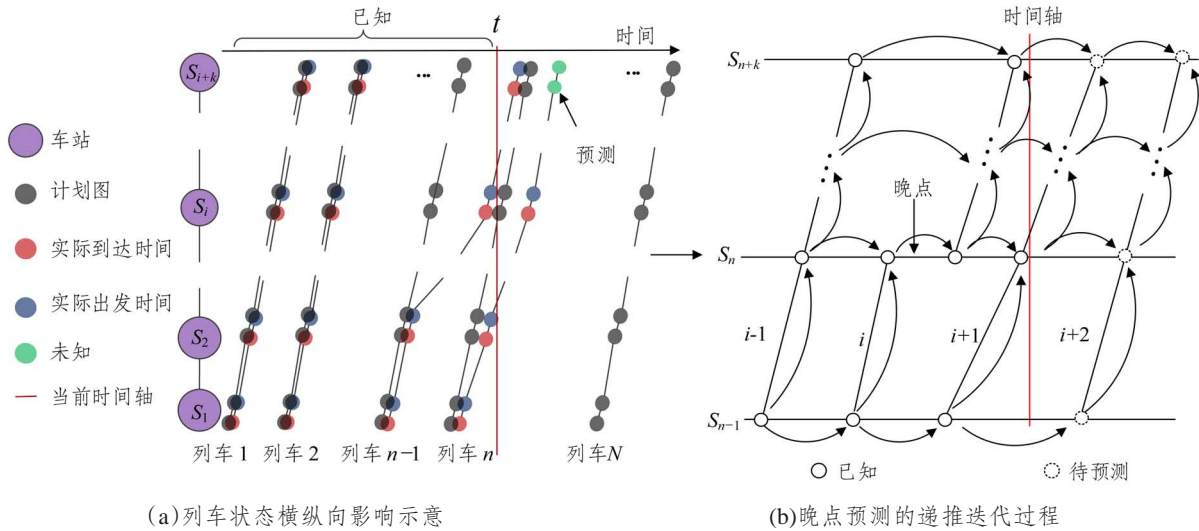


图3 考虑多列车的列车运行状态时空依赖关系及晚点传播过程预测

### 2.2.2 列车晚点影响预测

实现铁路行车故障的影响预测将能够为调度员预测列车晚点提供基础依据,可以帮助调度人员估计线路恢复正常运行的时间,并适当地重新安排列车运行。Huang 等<sup>[25]</sup>提取了故障影响晚点列车运行序列,运用 K-Means 聚类算法,根据故障的强度、发生时段及行车间隔将列车晚点故障聚类为 4 个不同的类别,可以用于列车晚点致因的特征分类。Oneto 等<sup>[26]</sup>以意大利铁路网实际维修记录、外部天气数据和运营商的经验为研究基础,应用决策树方法构建了一个基于规则的故障恢复时间预测模型,该模型具有足够的可解释性,有助于列车调度员合理评估局部铁路网络的可用性。Zilko 等<sup>[27]</sup>利用非参数贝叶斯网络建立了一个概率模型来估计铁路中断持续时间,该模型在很大程度上依赖于由历史数据生成的因变量的经验分布准确性。为了适应实时更新的信息,进一步提出了基于 Copula-Bayesian 网络方法的故障时间预测模型<sup>[28]</sup>,该模型通过建立中断长度与各影响因素之间的依赖关系来产生准确预测。汤轶雄等<sup>[29]</sup>以初始晚点时间、影响列车数、晚点致因为自变量,总晚点时间为因变量,运用支持向量回归模型实现了故障的晚点时长预测,能够为调度员的调度决策提供行车故障可能持续时长的信息支持。综上,在列车晚点致因分析方面,既有研究仍较为缺乏,没有对不同致因类型引起的晚点影响进行细分,缺乏不同类型晚点情况下的晚点状态描述和晚点程度的量化标准,需要应用机器学习方法对晚点致因和不同类型初始晚点影

响建模进行进一步研究,以推进列车运行精细化管理的实施。

### 2.2.3 列车晚点状态演化预测

列车晚点状态演化的预测,通常是以预测列车在车站到发时刻为突破口,基于列车运行数据挖掘列车运行状态影响要素对列车运行过程的作用,建立相应的机器学习模型以实现列车运行状态的推演和预测。线性回归方法被广泛应用于晚点预测问题建模,Wang 等<sup>[30]</sup>基于列车运行历史数据和实时信息建立了两种线性回归模型,使用列车在车站的出发时间估计列车晚点。Li 等<sup>[31]</sup>运用参数回归模型和非参数回归模型预测了列车短时停站时间。Guo 等<sup>[32]</sup>将列车运行看作是一系列的离散事件,基于京沪高速铁路 5 个车站的运行实绩建立了晚点预测的线性回归模型。近年来,支持向量机、随机森林等机器学习模型被初步尝试用来分析和预测晚点<sup>[17]</sup>。Kecman 和 Goverde<sup>[33]</sup>运用统计学习技术,构建了三种列车过程时间预测的全局模型:稳健回归、回归树模型和随机森林模型,并基于线性回归的鲁棒性,对特定的列车、车站或区段的局部模型进行了校准。Pongnumkul 等<sup>[34]</sup>以泰国国家铁路六个月的列车历史旅行时间数据为研究基础,使用最邻近节点算法实现了晚点旅客列车到站时间预测,该模型的缺点是难以应用于大型数据集。马尔科夫模型和贝叶斯网是列车晚点状态预测研究领域应用最广泛的方法。运用基于列车运行数据的马尔科夫模型预测晚点,其重点是状态转移矩阵的构建<sup>[35]</sup>,Barta 等<sup>[36]</sup>利用大量的历史晚点数据,提出了马尔科夫链模型

用以预测货运列车在连续车站的晚点情况。Gaurav等<sup>[37]</sup>建立了一个N阶马尔可夫晚点预测框架,使用随机森林回归和岭回归作为预测模型进行了实验。Corman等<sup>[38]</sup>构建了列车晚点时间预测的贝叶斯网络模型,实现了30分钟以内的列车晚点时间预测,揭示列车运行的动态特性。Oneto等<sup>[39]</sup>提出了一种基于深度极值机的列车晚点预测系统来预测晚点,该系统考虑了外部天气数据的影响。线性回归、梯度提升回归树、决策树、随机森林等4种机器学习方法被用来预测考虑天气影响下的列车晚点,模型比较结果显示随机森林的预测精度最高<sup>[40]</sup>。在神经网络模型应用方面,Yaghini等<sup>[41]</sup>提出了一种高精度的人工神经网络模型来预测伊朗铁路客运列车的晚点,并将3种不同的数据输入方式及3种体系结构的预测结果与决策树和多元逻辑回归等常用预测方法进行了比较。Oneto等<sup>[42,43]</sup>将列车晚点预测映射为多元回归问题,建立了一种基于前馈神经网络的极限学习机晚点预测模型,并进一步通过引入外部天气数据的影响以及阈值调整技术对模型进行了改进。张琦等<sup>[21]</sup>构建了基于小波神经网络的列车连带晚点预测模型,利用列车晚点波动的线性组合方程及其结构向量进行列车连带晚点影响值的量化。孙略添等<sup>[44]</sup>应用径向基函数神经网络对技术站列车晚点时间进行精确预测,同时指出该方法适用于晚点历史数据较多、大规模、对运到期限要求较高的技术站。

近年来,深度学习的方法被逐步应用于列车晚点预测建模。Oneto等<sup>[45]</sup>提出了基于机器学习方法的晚点预测系统Train Delay Prediction Systems (TDPS),该系统运用大数据技术和深度极限学习机算法,集成了异构数据源。通过从大量历史列车运行数据中提取信息,训练列车晚点预测模型,建立了基于数据驱动的动态列车晚点预测系统。根据列车运行状态更新训练数据集,TDPS实现了列车运行状态的实时预测。Wen等<sup>[46]</sup>基于列车运行实绩,长短记忆深度学习模型被用于预测列车晚点时间,挖掘列车运行各要素对列车运行的作用规律,开启了深度学习用于列车晚点预测的研究。Huang等<sup>[47]</sup>提出了一种结合三维卷积神经网络、长短期记忆神经网络和全连接神经网络结构的深度学习的方法,针对4条具有不同运行特征的线路进行了预测分析,结果表明该方法具有较高的

精度和较强的鲁棒性。

另外,列车晚点影响是晚点严重程度的重要度量指标。建立晚点影响模型能够有效地预测晚点的发生概率以及影响程度,可以协助列车调度员进行晚点预测和判断晚点传播的影响范围,从而制定合理的列车运行调整方案。典型机器学习方法,如极端梯度提升被用来预测晚点影响列车数、支持向量回归模型被用来预测晚点总影响时间,对武广高铁列车晚点影响的测试和验证表明模型能够为调度员估计晚点影响,度量晚点严重程度提供依据<sup>[48]</sup>。

综上,列车运行晚点状态的变化受到外界运行环境和铁路系统内部各要素的综合影响,是一个非常复杂的过程,列车运行晚点预测所要考虑的因素众多。一系列的统计机器学习方法和以卷积神经网络、长短期记忆神经网络等为代表的高级机器学习方法已经广泛应用于列车晚点时间的预测。已有研究建立了一套用于列车晚点预测的高级机器学习模型体系,经过对不同线路、不同数据量情况下的模型验证发现,高级机器学习方法取得了较好的晚点预测效果,但各类机器学习方法的适用条件和应用情景、所考虑的列车运行不同特征量、推广可行性等还是值得深入研究的问题,还没有得到通用性较好的模型和算法。更为重要的是,现有的研究主要是基于静态数据的,没有基于实时数据进行列车晚点状态动态推演的高效算法。

#### 2.2.4 列车晚点恢复预测

列车晚点恢复建模能够帮助调度员掌握相关调度策略的晚点恢复效果、冗余时间利用情况等,实现对列车晚点恢复的预测,提高调度调整效果。列车晚点恢复建模的重点是建立冗余时间利用模型,难点是提取冗余时间的分布规律及冗余时间的利用效率,最大限度地利用冗余时间消解晚点。

Khadilkar<sup>[49]</sup>研究了晚点分布概率,通过分析历史数据知印度铁路平均晚点恢复率为0.13 min/km,并在晚点恢复模型中以此值表示晚点恢复的能力,但是这个平均值很难反映列车在每个区间、车站的晚点恢复能力,这将影响模型的预测能力。Yang等<sup>[50]</sup>运用统计方法建立了列车晚点期望模型,并提出了铁路冗余时间冲突分配的方法。Steven等<sup>[23]</sup>用统计模型分析了晚点列车的冗余时间利用情况,但在分析列车晚点恢复时,假定了冗余

时间利用率最大的理想情况。Jiang等<sup>[51]</sup>基于武广高铁的列车运行实绩,比较了多类机器学习模型用于初始晚点恢复预测的效果,建立了有最佳效果的初始晚点恢复预测的随机森林模型。Şahin等<sup>[52]</sup>使用从土耳其国家铁路收集到的历史数据,建立了一个马尔可夫链模型估计列车晚点传播和恢复。Huang等<sup>[53]</sup>基于武广高铁列车运行数据建立了缓冲时间分配的数据驱动模型;岭回归机器学习模型被用来解析考虑车站缓冲时间方案、区间缓冲时间方案、晚点严重程度等要素的列车晚点恢复规律;根据缓冲时间的利用率,该模型重新分配缓冲时间,为缓冲时间的优化配置提供了新的解决思路。所提出的机器学习模型综合考虑了运行图的执行效果指标,如缓冲时间利用率和晚点概率等。Martin<sup>[54]</sup>提出了运用预测推理和机器学习方法来提高铁路系统可靠性,认为运用贝叶斯推理进行晚点预测时可以考虑冗余时间的布局,但该研究并没有阐释具体方法应用。胡雨欣等<sup>[55]</sup>以高速列车初始晚点时间、站停冗余时间和区间冗余时间等为变量,使用多层感知器和循环神经网络建立了高速列车晚点恢复时间预测机器学习模型,预测精度对比发现循环神经网络的效果要优于多层感知器模型。

在采用机器学习方法研究冗余时间利用问题时,需要使用大量的晚点恢复数据和冗余时间运用数据,而这些数据的获得又有一定难度,因此,基于机器学习的列车晚点恢复预测研究仍较缺乏。挖掘不同晚点情景下的冗余时间恢复效率、基于冗余时间利用数据建立冗余时间利用与重布局的数据驱动模型将是研究的重点方向。运用机器学习方法建立车站及区间晚点时间恢复模型,探明车站、区间的冗余时间布局方案对于一定晚点的恢复能力具有重要的实际意义,有助于发现不同调度调整策略作用下列车运行晚点恢复及效果,能够为列车运行调度决策提供决策支持。将列车运行实绩与列车调度命令数据结合、结构化数据与半结构化数据融合建模是提高列车晚点恢复决策和实现调度知识自动化要解决的关键问题。

### 2.3 列车运行调整智能化决策研究

对于列车运行调整决策问题的研究主要集中在列车运行冲突消解、列车运行计划调整、动态网络调度等方面。国内外既有研究主要集中于运用数学优化模型求解列车运行冲突检测及消解、运行

图编制优化问题<sup>[56,57]</sup>,建立以最小晚点影响<sup>[58]</sup>、最小列车晚点时间<sup>[59]</sup>、最短列车运行时间<sup>[60]</sup>等为优化目标的列车运行调整模型,但这些模型必须在计算性能和解决方案质量之间进行权衡。同时,由于列车运行调整决策模型涉及的约束条件多、参数多、优化目标多,使得精准建模和求解极为困难。列车运行调整智能化决策研究的主要内容如表1所示。

表1 列车运行调整智能化决策研究的主要内容

常用策略	问题类型	优化目标
变更到发时刻		最小化晚点列车数
变更停站时间	晚点恢复	最小化旅客总晚点时间
变更股道使用方案	冲突检测	最小化运营成本
变更停站	与消解	最小化总晚点时间
变更列车次序	运行图优化	最小化总旅行时间
变更区间运行时间		

列车运行调度调整过程常被作为马尔科夫决策过程来处理,通过推导列车的运行状态变化规律得到可能的调度决策方案<sup>[61]</sup>。Dündar<sup>[62]</sup>基于调度人员在10个工作日内解决的331个冲突记录,建立了一个神经网络模型来模拟列车调度员进行冲突消解。Oneto等<sup>[63]</sup>以列车晚点和相关代价最小化为目标,研究了列车越行预测的问题,并基于随机森林方法开发了一个混合预测模型。上述研究基于调度员历史决策合理的假设,着重于调度员历史决策方案的再现,没有考虑到调度决策制定过程中众多的动态变化因素的影响,因此,其决策方案的优化性能有限。强化学习结合了动态规划和监督学习的原理,尤其适用于解决列车运行调整问题。Šemrov等<sup>[64]</sup>提出了一种基于Q-learning算法的单线铁路列车运行调整方法,其基本原理为:通过算法与环境交互,获得并解释其从环境中获得增强信号,进而选择使所获得的奖励之和最大化的动作,并以此逐步生成合理的列车运行调整方案,模型的奖励函数为使列车总晚点最小。Khadilkar<sup>[65]</sup>将该研究扩展到双线铁路上,并以印度铁路网的两条线路为例,验证了该算法在实际中的适用性,结果表明:这种方法可以在与启发式方法相当的计算时间内处理大型调度问题实例,同时具有更好的实用性与决策质量。模糊神经网络作为模糊理论和神经网络的融合,吸取了模糊逻辑和神经网络的优点,部分避免了两者的缺点。Sun等<sup>[66]</sup>通过对高速铁路列车运行控制程序的详细分析,提出了一种基于模糊神经网络

的列车运行调整模型,并通过引入变步长的改进BP (Back Propagation)神经网络算法,完成了自适应神经网络模糊推理系统的训练工作。Agent方法作为一种解决动态环境中不确定性问题的有力技术,在交通系统领域的应用正在迅速兴起。Narayanawami等<sup>[67]</sup>提出了一种结合Agent和无监督学习方法的动态调度模型,通过引入一个基于实时系统参数的动态调度计算框架对模型进行改进,并将其与混合整数线性规划方法进行比较。实验结果表明:Agent算法在求解复杂铁路网络调度问题时,其计算时间较精确的模型有很大的减少,体现出其在解决大规模复杂的调度问题方面的潜力。D-Agent方法被用来研究冲突消解并支持调度员基于多源信息及异构数据进行列车冲突检测,支持列车运行调度决策<sup>[68,69]</sup>。

瑞士铁路运营公司在既有调度指挥系统的基础上,自主研发了瑞士铁路调度指挥控制系统(Railway Control System, RCS),与既有调度指挥系统形成功能上的兼容互补,是目前最先进的列车运行与调度管理系统<sup>[70]</sup>。RCS系统根据历史运行数据、当前列车运行状态、列车运行可能的运行情景等对列车运行进行精确预测,并生成无运行

冲突的调度决策,该系统为瑞士提高路网利用效率和运输服务质量做出了重要贡献。这也让世界看到了基于铁路实时和历史数据,运用机器学习实现调度决策知识自动化、支持调度决策及列车运行的应用前景。

从上述研究分析可以看出,机器学习方法为铁路列车运行调整智能化决策提供了有效的解决渠道,已经成为铁路智能化调度决策研究的主要方法,将先进的机器学习方法应用于调度策略效果评估、调度决策知识自动化仍然是研究的主要难点和关键,但目前利用机器学习算法实现调度策略的优选、智能化调度策略生成等的研究仍比较缺乏,构建列车调度调整的自动化知识库,建立列车调度调整策略的推荐系统,为调度员实时列车调度提供决策支持是亟待解决的问题。

### 3 既有研究的特征及研究动向

#### 3.1 既有研究的特征

综上所述,机器学习方法已经广泛应用于铁路调度指挥领域,在晚点分析、预测分析、智能调度模型构建等方面取得了较为丰硕的研究成果(见表2)。

表2 近十年机器学习方法在铁路列车运行调整领域研究文献汇总

文献来源	发表时间/年	研究问题	方法(系统)
文献[11]	2013	晚点特征分析	统计模型
文献[12]	2015	晚点特征分析及拟合	统计模型
文献[13]	2017	晚点特征分析及拟合	统计模型
文献[14]	2020	列车连带晚点分析及拟合	统计模型
文献[15]	2017	列车晚点致因及影响拟合	统计模型
文献[16]	2018	列车晚点致因及影响拟合	统计模型
文献[17]	2015	晚点特征建模	支持向量回归
文献[18]	2010	晚点特征建模	线性回归
文献[19]	2016	晚点致因分析	决策树
文献[20]	2018	周期性晚点规律分析	K-Means 聚类
文献[21]	2019	连带晚点分析、特征识别	小波神经网络
文献[22]	2018	晚点传播过程分析	统计模型
文献[23]	2019	晚点演化及恢复过程	统计模型
文献[24]	2019	晚点演化数据分析	统计模型
文献[25]	2019	晚点致因分类	K-Means 聚类
文献[26]	2019	故障恢复时间预测	决策树
文献[27]	2014	铁路中断持续时间预测	非参数贝叶斯网络
文献[28]	2016	铁路中断持续时间预测	Copula-贝叶斯网络
文献[29]	2019	铁路故障晚点时间预测	支持向量机
文献[30]	2015	晚点预测	线性回归
文献[31]	2016	短时停站预测	参数回归与非参数回归

续表2

文献来源	发表时间/年	研究问题	方法(系统)
文献[32]	2015	晚点预测	线性回归
文献[33]	2015	列车运行和停站时间预测	稳健回归、回归树、随机森林
文献[34]	2014	旅客列车到站时间预测	K最近邻分类
文献[35]	2015	列车晚点不确定性建模、晚点预测	非平稳马尔可夫链
文献[36]	2012	晚点传播模型、货运列车晚点预测	马尔可夫链
文献[37]	2018	列车晚点预测	N阶马尔可夫模型
文献[38]	2018	列车晚点预测	马尔科夫模型
文献[39]	2016	列车晚点预测	核方法、极限学习机、随机森林
文献[40]	2021	列车晚点预测	线性回归、梯度提升回归、决策树、随机森林
文献[41]	2013	旅客列车晚点预测	神经网络、决策树、多分类逻辑回归
文献[42]	2016	列车晚点预测	极限学习机
文献[43]	2017	列车晚点预测	阈值调整技术、深浅层极限学习机
文献[44]	2018	列车晚点时间预测	径向基函数(RBF)神经网络
文献[45]	2018	列车晚点预测	深度极限学习机
文献[46]	2020	列车晚点预测	长短记忆深度学习
文献[47]	2020	晚点预测	3D-CNN、LSTM、FCNN
文献[48]	2020	晚点影响程度预测	极端梯度提升算法
文献[49]	2016	列车晚点恢复预测	统计模型
文献[50]	2020	列车晚点恢复、冗余时间布局模型	统计模型
文献[51]	2019	初始晚点恢复预测	多重线性回归、支持向量机、神经网络、随机森林
文献[52]	2017	晚点传播模型、晚点预测	马尔可夫链
文献[53]	2019	冗余时间布局模型	岭回归
文献[54]	2016	冗余时间布局模型	贝叶斯推理
文献[55]	2020	列车运行晚点恢复预测	多层感知器、循环神经网络
文献[61]	2017	列车运行调整策略	马尔科夫模型
文献[62]	2013	冲突消解决策模拟	神经网络
文献[63]	2019	列车越行预测及决策	随机森林
文献[64]	2016	列车运行调整策略	Q-learning 算法
文献[65]	2018	列车运行调整策略	Q-learning 算法
文献[66]	2010	列车运行调整策略	模糊神经网络
文献[67]	2015	列车运行调整策略	Agent 方法、关联分析算法
文献[68]	2016	冲突消解决策	D-Agent
文献[69]	2017	冲突消解决策	D-Agent

研究侧重点和相关文献的发表趋势如图4、图5所示。

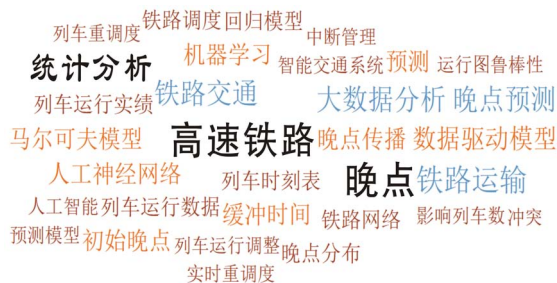
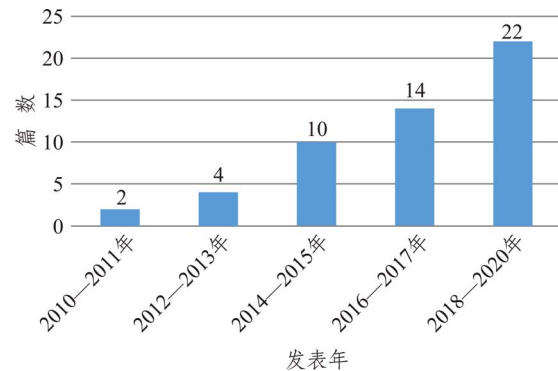
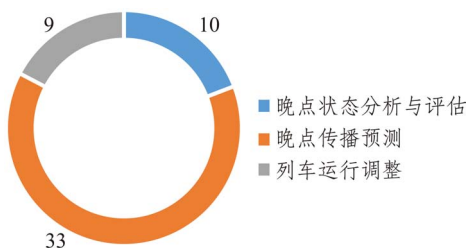


图4 相关研究关键词云



(a) 相关文献的发表趋势



(b) 相关文献研究的侧重点

图5 既有研究特征分析

通过对上述研究的回顾,可知机器学习应用于铁路列车调度调整问题的既有研究呈现以下特征及存在的问题:

(1)机器学习方法应用于铁路调度调整问题基于的数据以列车运行计划运行图、实际运行图为主,气象数据、到发线运用数据、设备故障数据等作为补充数据在部分研究中得以应用并有效提高了模型精度。在既有研究中,相关数据不全面是主要的因素之一,比如有列车运行数据而没有设备运用数据、有客流数据而没有气象数据等,所建立的模型存在较大的缺陷。

(2)从图5(a)可以看出,机器学习方法在铁路列车调度调整问题的研究呈现较为明显的增长趋势,尤其伴随大数据技术和深度学习方法的快速发展,深度学习应用于列车运行晚点智能预测与列车运行调整决策方面的成果已经逐步显现,可以预期将有较为丰富的研究成果产出。

(3)从图4和图5(b)可以看出,高速铁路调度指挥是研究的重点领域、晚点预测是当前的研究热点问题,传统机器学习、人工神经网络、大数据分析以及统计分析等方法已经在既有研究中得到了广泛的应用。

(4)在列车晚点分析方面,机器学习方法(如关联、聚类等)表现出较大的潜力<sup>[71]</sup>,可以用于挖掘大数据中隐含的列车晚点机理,但当前该领域的研究仍是以传统统计分析为主导,处理的数据也较为有限,运用深度学习方法挖掘更大规模和高维度列车运行相关复合数据的工作还有待加强。

(5)列车晚点致因及持续时长预测受到铁路系统内外部诸多因素的影响,需要针对不同的故障类型收集故障报告、天气情况、相关调度命令等信息。受限于数据获取及处理的难度,且目前设备故障数据、天气数据等与列车运行数据的粒度和采集间隔相差较大,增加了建模的难度,降低了模型的精度。

(6)现有运用数学优化方法和一般机器学习方法研究列车晚点状态演化及时间的预测问题已经很充分,运用深度学习分析列车之间的相互作用关系并建立列车晚点精确预测模型,是列车晚点智能化预测领域新的研究趋势。相关模型与算法在晚点预测方面虽然已经体现了较好的性能,但列车分类型晚点传播预测的研究还明显偏少,相应的标志性成果还有待深化,预测不同程度晚点情况下的发展态势从而辅助调度员的调度策略制定研究是当前研究的短板。

(7)在列车运行调整智能化决策方面,国内外学者已经提出了多种结合机器学习方法的模型解决调度优化问题,但相关研究仍较为缺乏,将更多既有的高性能预测模型应用于辅助决策系统构建是下一步需要重点研究的内容。既有研究的重点多是放在模型构建和模拟验证上,还未有研究成果在实际应用中部署和实施,如何将机器学习模型应用于智能化调度指挥是亟待解决的难题。

### 3.2 研究动向

“数据+算力+算法”引领的智能制造,带来了工具革命,也带来了以智能化提高决策科学性、精准化为标志的决策革命。通过人工智能等手段优化决策的准确性、及时性、科学性已成为重要创新途径<sup>[72]</sup>。通过铁路智能化提高铁路运输组织效率、保障铁路运输安全已成为各国铁路发展的必由之路<sup>[73]</sup>。法国、德国、美国、加拿大、日本、韩国等国家铁路相继提出了数字化与智能化发展的战略规划。实现列车运行态势智能感知和冲突自动消解,支撑调度智能化是世界智能铁路发展的迫切需求。阿里研究院发布的《2020年十大科技趋势》中第一个趋势是“人工智能从感知智能向认知智能演进”<sup>[74]</sup>。铁路运输系统“智能+”决策革命是铁路智能化发展的必然趋势,结合铁路行车组织理论、经验和既有技术,基于铁路多源数据,以数据科学、人工智能为主要手段的智能调度集中系统,是实现铁路行车组织理论和技术革新的有效途径。

铁路运输系统尤其是列车运行过程是一个随机+可控的系统,从列车运行实绩及相关数据中探索并发现随机干扰事件的发生概率、相关调度调整策略可能的效果是铁路调度调整决策的关键问题。而机器学习方法尤其是以深度学习为代表的高级机器学习方法在挖掘大规模数据并发现其规

律方面具有显著优势,其在铁路列车运行调整领域还有大量工作要开展,如:

(1)引入多源异构数据构建机器学习模型。当前大部分既有机器学习模型都是基于列车运行所产生的结构化数据(如车站的到达和出发时间)构建的。近年来随着铁路信息化建设的不断发展,与列车运行相关的大量数据也得到了有效的记录和保存,其中包含了诸多不同来源的结构化与非结构化数据。利用先进的数据分析方法,完成各种类型数据的清洗与加工,基于多源异构数据构建不同预测目标的机器学习模型是未来研究的重点。以列车运行图数据为基本数据,考虑列车运行所涉及的气象、信联闭设备、到发线运用、车底运用等数据,将能够挖掘更丰富的列车运行规律,建立的列车运行调度调整模型将更能贴近列车运行实际,更好地指导调度指挥决策。

(2)列车晚点程度量化和分级能够为调度员掌握全局运输态势提供依据。研究列车晚点分级模型,确定列车晚点程度并实现不同程度晚点的预测将能够提升铁路调度调整的精细化水平。

(3)列车运行晚点预测依然是研究的热点和难点。高效准确的预测是机器学习模型应用于调度生产实际的关键所在,虽然当前 Boosting、Bagging 和 Stacking 等典型集成学习技术已经得到了较为广泛的应用,但不断改进算法、提升模型的计算能力和精度仍是一个重要的研究方向。列车运行晚点分类预测对于提升调度决策水平具有重要意义,分别建立初始晚点、连带晚点的预测模型,能够辅助调度员根据一定的初始晚点制定合理的调度策略以尽可能控制连带晚点及其影响程度和影响范围。列车运行晚点的实时预测是一个对模型抽象能力和求解速度要求很高的问题,需要不断寻求能够尽可能贴近运输生产实际的高效模型和算法。

(4)列车运行调度调整自动化是重要的发展趋势。基于列车运行数据,建立列车运行调度调整的数据驱动模型,挖掘列车运行调度调整的模式及自动化知识,根据历史、当前及未来可能的运输态势实现列车运行调度策略的自动优选,将能够辅助调度员的调度调整决策,提高列车运行调度调整的决策质量并逐步实现调度调整的自动化。

(5)建立基于机器学习的铁路调度调整智能

化系统。以列车运行相关历史数据和实时数据为基础,以基于机器学习的列车调度调整系列模型为底层模型,实现铁路列车晚点可视化、晚点预测、晚点恢复、调度策略智能化优选等。基于机器学习的铁路调度调整智能化系统将是我国铁路智能调度的重要保障。

(6)探索高级机器学习方法的更多应用。当前机器学习方法在列车运行调整智能化决策方面的应用还较为有限,并且还有很多先进的机器学习算法尚未使用或应用不足,需要针对调度指挥工作实际需要,基于不同机器学习算法特性开发更多的模型与方法,并尝试将其与既有铁路调度系统相结合,尽可能贴近列车运行实际过程,考虑尽可能多的因素,以提升铁路调度指挥决策与控制自动化水平,降低相关作业人员负荷,提高铁路运输效率和服务质量。

#### 4 结束语

从大规模复杂路网建设过渡到精细化管理的网络化运营已是我国铁路运营的重要特征,实现海量数据可挖掘、设备状态可诊断、行车安全可预警、复杂路网运营变化可感知、发展趋势可推断、辅助决策可支撑,提升铁路的智能决策水平,提高决策效率是智能铁路亟待解决的关键科学问题。铁路运输系统海量数据信源多样,列车运行过程中产生的大量列车运行实际数据尚未被充分挖掘和运用,铁路列车运行调整仍然是以个人经验为基础的人工决策为主。如何基于海量多源数据,运用数据科学和人工智能提升铁路调度指挥决策的效度,提升决策水平、降低作业人员工作负荷是铁路智能调度决策面临的巨大挑战。

本文全面回顾了近10年机器学习方法在铁路列车运行调整领域应用的主要研究成果,将列车运行调整问题分为列车晚点状态分析与评估、列车晚点传播预测和列车运行调整智能化决策三个方面,分别总结和分析了机器学习方法在上述三方面问题的应用情况。总结了既有研究的特征,并在此基础上展望了机器学习方法在铁路列车运行调整研究方面的最新动向。

机器学习方法能够克服传统数学模型难以应用于生产实践的问题,在当前国家大数据战略的驱动下,基于机器学习的调度指挥自动化系统必将为铁路行车指挥决策提供有效支撑。近10年来

机器学习方法在铁路列车运行调整方面的研究呈现较为明显的增长趋势,深度学习应用于列车运行晚点智能预测与列车运行调整决策方面的成果已经逐步显现。高级机器学习和深度学习的列车运行调整智能决策理论以及列车运行调整知识自动化等智能化理论将是未来的发展重点。

### 参考文献

- [1] “智能高铁战略研究(2035)”通过工程院结题评审[EB/OL].[2020-01-15]<http://www.chinarevit.com/revit-55316-1-1.html>.
- [2] GRAVES A, MOHAMED A R, HINTON G. Speech recognition with deep recurrent neural networks[C]//2013 IEEE International Conference on Acoustics, Speech and Signal Processing. Vancouver, Canada: IEEE, 2013: 6645-6649.
- [3] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[EB/OL]. (2015-04-10) [2021-03-15] <https://arxiv.org/abs/1409.1556>.
- [4] 赵雁飞. 行车指挥自动化几个关键问题的研究[D]. 北京:北方交通大学, 2000.
- [5] 中国科学院. 2019全球人工智能发展白皮书(2019)[R]. 北京:中国科学院, 2019.
- [6] 赛迪顾问股份有限公司. 2018年中国大数据产业生态白皮书[R]. 北京:赛迪顾问股份有限公司, 2018.
- [7] KECMAN P, GOVERDE R. Process mining approach for recovery of realized train paths and route conflict identification[C]//Transportation Research Board TRB: 92nd Annual meeting. Waslaington, USA: TRB, 2013.
- [8] 廖志林, 唐伟忠. 铁路调度综合管理平台应用研究[J]. 铁路计算机应用, 2018, 27(4): 25-28.
- [9] KIM K, SEOL S, KONG S H. High-speed train navigation system based on multi-sensor data fusion and map matching algorithm[J]. International Journal of Control, Automation and Systems, 2015, 13(3): 503-512.
- [10] 雷明. 机器学习——原理、算法与应用[M]. 北京:清华大学出版社, 2019.
- [11] GRAFFAGNINO T. Ensuring timetable stability with train traffic data[J]. Computers in Railways XIII: Computer System Design and Operation in the Railway and Other Transit Systems, 2013,127: 427-438.
- [12] 刘岩, 郭竞文, 罗常津, 等. 列车运行实绩大数据分析及应用前景展望[J]. 中国铁路, 2015(6): 70-73.
- [13] WEN C, LI Z, LESSAN J, et al. Statistical investigation on train primary delay based on real records: evidence from Wuhan - Guangzhou HSR[J]. International Journal of Rail Transportation, 2017, 5(3): 170-189.
- [14] 徐传玲, 文超, 胡瑞, 等. 高速铁路列车连带晚点产生机理及其判定[J]. 交通运输工程与信息学报, 2020, 18(4): 31-37.
- [15] 庄河, 文超, 李忠灿, 等. 基于高速列车运行实绩的致因-初始晚点时长分布模型[J]. 铁道学报, 2017, 39(9): 25-31.
- [16] 文超, 李忠灿, 黄平, 等. 高速铁路初始晚点致因-影响列车数分布模型[J]. 西南交通大学学报, 2018, 53(6): 1261-1269.
- [17] MARKOVIĆ N, MILINKOVIĆ S, TIKHONOV K S, et al. Analyzing passenger train arrival delays with support vector regression[J]. Transportation Research Part C: Emerging Technologies, 2015, 56: 251-262.
- [18] MURALI P, DESSOUKY M, ORDÓÑEZ F, et al. A delay estimation technique for single and double-track railroads[J]. Transportation Research Part E: Logistics and Transportation Review, 2010, 46(4): 483-495.
- [19] LEE W H, YEN L H, CHOU C M. A delay root cause discovery and timetable adjustment model for enhancing the punctuality of railway services[J]. Transportation Research Part C: Emerging Technologies, 2016, 73: 49-64.
- [20] CERRETO F, NIELSEN B F, NIELSEN O A, et al. Application of data clustering to railway delay pattern recognition[J/OL]. Journal of Advanced Transportation, 2018: 18(2018-04-29 ). <https://doi.org/10.1155/2018/6164534>.
- [21] 张琦, 陈峰, 张涛, 等. 高速铁路列车连带晚点的智能预测及特征识别[J]. 自动化学报, 2019, 45(12): 2251-2259.
- [22] CHEBOTAREV V, DAVYDOV B, KABLUKOVA K. Probabilistic model of delay propagation along the train flow[J]. Probabilistic Modeling in System Engineering, IntechOpen, 2018: 171-193.
- [23] HARROD S, CERRETO F, NIELSEN O A. A closed form railway line delay propagation model[J]. Transportation Research Part C: Emerging Technologies, 2019, 102: 189-209.
- [24] HARROD S, POURNARAS G, NIELSEN B F. Distribution fitting for very large railway delay data sets with discrete values[J]. Danish Journal of Transportation Research - Dansk Tidsskrift for Transportforskning, 2019, 1: 1-12.
- [25] HUANG P, WEN C, PENG QY, et al. Modeling the Influence of Disturbances in High-Speed Railway Systems [J/OL]. Journal of Advanced Transportation, 2019: 13. (2019-05-19).<https://doi.org/10.1155/2019/8639589>.

- [26] ONETO L, BUSELLI I, SANETTI P, et al. Restoration time prediction in large scale railway networks: big data and Interpretability[C]//Recent Advances in Big Data and Deep Learning: Proceedings of the INNS Big Data and Deep Learning Conference INNSBDDL2019. Genova, Italy: Springer, 2019, 1: 136-141.
- [27] ZILKO A, HANEA A, KUROWICKA D, et al. Non-parametric Bayesian network to forecast railway disruption lengths[C]//PROC. 2nd International Conference on Railway Technology: Research, Development and Maintenance (Railways 2014). Ajaccio, France: PROC, 2014.
- [28] ZILKO A A, KUROWICKA D, GOVERDE R M. Modeling railway disruption lengths with copula Bayesian networks[J]. Transportation Research Part C: Emerging Technologies, 2016, 68: 350-368.
- [29] 汤轶雄, 徐传玲, 文超, 等. 高铁故障晚点时间预测的支持向量回归模型[J]. 中国安全科学学报, 2019, 29(S2): 18-23.
- [30] WANG R, WORK D B. Data driven approaches for passenger train delay estimation[C]// IEEE. 2015 IEEE 18th International Conference on Intelligent Transportation Systems. Gran Canaria, Spain: IEEE, 2015: 535-540.
- [31] LI D, DAAMEN W, GOVERDE R M. Estimation of train dwell time at short stops based on track occupation event data: a study at a Dutch railway station[J]. Journal of Advanced Transportation, 2016, 50(5): 877-896.
- [32] GUO J, MENG L, KECCMAN P, et al. Modeling delay relations based on historical train monitoring data: a Chinese high-speed railway case[C]//6th International Conference on Railway Operations Modelling and Analysis. Tokyo, Japan: ROMA, 2015.
- [33] KECCMAN P, GOVERDE R M. Predictive modelling of running and dwell times in railway traffic[J]. Public Transport, 2015, 7(3): 295-319.
- [34] PONGNUMKUL S, PECHPRASARN T, KUNASETH N, et al. Improving arrival time prediction of Thailand's passenger trains using historical travel times[C]//11th International Joint Conference on Computer Science and Software Engineering. Chonburi, Thailand: JCSSE, 2014: 307-312.
- [35] KECCMAN P, CORMAN F, MENG L. Train delay evolution as a stochastic process[C]//6th International Conference on Railway Operations Modelling and Analysis Rail. Tokyo, Japan: ROMA, 2015.
- [36] BARTA J, RIZZOLI A E, SALANI M, et al. Statistical modelling of delays in a rail freight transportation network[C]// Proceedings of the 2012 Winter Simulation Conference (WSC). Berlin, Germany: WSC, 2012: 1-12.
- [37] GAURAV R, SRIVASTAVA B. Estimating train delays in a large rail network using a zero shot markov model [C]//2018 21st International Conference on Intelligent Transportation Systems ITSC. Maui, USA: ITSC, 2018: 1221-1226.
- [38] CORMAN F, KECCMAN P. Stochastic prediction of train delays in real-time using Bayesian networks[J]. Transportation Research Part C: Emerging Technologies, 2018, 95: 599-615.
- [39] ONETO L, FUMEO E, CLERICO G, et al. Advanced analytics for train delay prediction systems by including exogenous weather data[C]//2016 IEEE International Conference on Data Science and Advanced Analytics. Montreal, Canada: DSAA, 2016: 458-467.
- [40] ARSHAD M, AHMED M. Train delay estimation in indian railways by including weather factors through machine learning techniques[J]. Recent Advances in Computer Science and Communications, 2021, 14(4): 1300-1307.
- [41] YAGHINI M, KHOSHRAFTAR M M, SEYEDABADI M. Railway passenger train delay prediction via neural network model[J]. Journal of Advanced Transportation, 2013, 47(3): 355-368.
- [42] ONETO L, FUMEO E, CLERICO G, et al. Delay prediction system for large-scale railway networks based on big data analytics[C]//INNS Conference on Big Data. Thessaloniki, Greece: INNS, 2016: 139-150.
- [43] ONETO L, FUMEO E, CLERICO G, et al. Dynamic delay predictions for large-scale railway networks: deep and shallow extreme learning machines tuned via thresholdout[J]. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2017, 47(10): 2754-2767.
- [44] 孙略添, 宋瑞, 何世伟, 等. 技术站列车晚点时间预测方法[J]. 北京交通大学学报, 2018, 42(1): 94.
- [45] ONETO L, FUMEO E, CLERICO G, et al. Train delay prediction systems: a big data analytics perspective[J]. Big Data Research, 2018, 11: 54-64.
- [46] WEN C, MOU W, HUANG P, et al. A predictive model of train delays on a railway line[J]. Journal of Forecasting, 2020, 39(3): 470-488.
- [47] HUANG P, WEN C, FU L, et al. A deep learning approach for multi-attribute data: a study of train delay prediction in railway systems[J]. Information Sciences, 2020, 516: 234-253.
- [48] LI Z, HUANG P, WEN C, et al. Predictive models for influence of primary delays using high-speed train operation records[J]. Journal of Forecasting, 2020, 39(8): 1198-1212.

- [49] KHADILKAR H. Data-enabled stochastic modeling for evaluating schedule robustness of railway networks[J]. *Transportation Science*, 2016, 51(4): 1161-1176.
- [50] YANG X, HOU Y, LI L. Buffer time allocation according to train delay expectation at stations[J]. *International Journal of Rail Transportation*, 2020, 8(3): 249-263.
- [51] JIANG C, HUANG P, LESSAN J, et al. Forecasting primary delay recovery of high-speed railway using multiple linear regression, supporting vector machine, artificial neural network, and random forest regression[J]. *Canadian Journal of Civil Engineering*, 2019, 46(5): 353-363.
- [52] ŞAHIN I. Markov chain model for delay distribution in train schedules: assessing the effectiveness of time allowances[J]. *Journal of Rail Transport Planning & Management*, 2017, 7(3): 101-113.
- [53] HUANG P, WEN C, PENG Q Y, et al. A data-driven time supplements allocation model for train operations on high-speed railways[J]. *International Journal of Rail Transportation*, 2019, 7(2): 140-157.
- [54] MARTIN L J. Predictive reasoning and machine learning for the enhancement of reliability in railway systems [C]//International Conference on Reliability, Safety, and Security of Railway Systems. Paris, France. 2016: 178-188.
- [55] 胡雨欣, 彭其渊, 鲁工圆, 等. 基于初始晚点和冗余时间的列车晚点恢复时间预测模型[J]. *交通运输工程与信息学报*, 2020, 18(2): 93-102.
- [56] CACCHIANI V, HUISMAN D, KIDD M, et al. An overview of recovery models and algorithms for real-time railway rescheduling[J]. *Transportation Research Part B: Methodological*, 2014, 63: 15-37.
- [57] CADARSO L, MARÍN Á. Recovery of disruptions in rapid Transit networks with origin-destination demand [J]. *Procedia-Social and Behavioral Sciences*, 2014, 111: 528-537.
- [58] 占曙光, 赵军, 彭其渊, 等. 高速铁路区间能力全失效条件下列车运行实时调整研究[J]. *铁道学报*, 2015, 37(11): 1-9.
- [59] KRASEMANN J T. Design of an effective algorithm for fast response to the re-scheduling of railway traffic during disturbances[J]. *Transportation Research Part C: Emerging Technologies*, 2012, 20(1): 62-78.
- [60] 王涛, 张琦, 赵宏涛, 等. 基于替代图的列车运行调整计划编制及优化方法[J]. *中国铁道科学*, 2013, 34(5): 126-133.
- [61] 庄河, 何世伟, 戴杨斌. 高速铁路列车运行调整的模式及其策略优化方法[J]. *中国铁道科学*, 2017, 38(2): 118-126.
- [62] DÜNDAR S, ŞAHIN I. Train re-scheduling with genetic algorithms and artificial neural networks for single-track railways[J]. *Transportation Research Part C: Emerging Technologies*, 2013, 27: 1-15.
- [63] ONETO L, BUSELLI I, LULLI A, et al. Train overtaking prediction in railway networks: a big data perspective[C]//INNS Big Data and Deep Learning conference. Genova, Italy: Springer, 2019: 142-151.
- [64] ŠEMROV D, MARSETIČ R, ŽURA M, et al. Reinforcement learning approach for train rescheduling on a single-track railway[J]. *Transportation Research Part B: Methodological*, 2016, 86: 250-267.
- [65] KHADILKAR H. A scalable reinforcement learning algorithm for scheduling railway lines[J]. *IEEE Transactions on Intelligent Transportation Systems*, 2018, 20(2): 727-736.
- [66] SUN J, WANG S. Intelligent train dispatching for high-speed railway based on fuzzy neural network[C]//2010 2nd International Conference on Advanced Computer Control. Shenyang: CACC, 2010: 222-225.
- [67] NARAYANASWAMI S, RANGARAJ N. A MAS architecture for dynamic, realtime rescheduling and learning applied to railway transportation[J]. *Expert Systems with Applications*, 2015, 42(5): 2638-2656.
- [68] ZHU T, MERA J M, SUAREZ B, et al. An agent-based support system for railway station dispatching[J]. *Expert Systems with Applications*, 2016, 61: 39-52.
- [69] ZHU T, DE PEDRO J M M S. Railway traffic conflict detection via a state transition prediction approach[J]. *IEEE Transactions on Intelligent Transportation Systems*, 2017, 18(5): 1268-1278.
- [70] 唐涛, 李开成, 宿帅, 等. 瑞士铁路调度指挥系统特点分析及对我国启示[J]. *中国铁路*, 2019(11): 18-23.
- [71] GHOFRANI F, HE Q, GOVERDE R M, et al. Recent applications of big data analytics in railway transportation systems: a survey[J]. *Transportation Research Part C: Emerging Technologies*, 2018, 90: 226-246.
- [72] 阿里研究院. 从工具革命到决策革命——通向智能制造的转型之路[R]. 杭州: 阿里研究院, 2019.
- [73] 李平, 邵赛, 薛蕊, 等. 国外铁路数字化与智能化发展趋势研究[J]. *中国铁路*, 2019(02): 25-31.
- [74] 阿里巴巴达摩院. 达摩院 2020 十大科技趋势[R]. 杭州: 阿里巴巴达摩院, 2020.