

# 基于 MEA-BP 神经网络融合多源数据的地铁客流预测

李丙章 余海兵 李世中 陈佳杭 张冰雨\*

西藏大学工学院 西藏拉萨 850000

**摘要:** 随着应对广州地铁 3 号线因客流激增引发的拥堵问题,本研究针对传统 BP 神经网络预测精度不足的缺陷,提出融合思维进化算法(MEA)的优化模型(MEA-BP)。通过整合历史客流、天气及日期等多源数据构建基准模型,结合 MEA 的趋同-异化机制动态优化网络权值与阈值。实验表明:优化后模型的平均绝对误差降至 0.38 万人(降幅 58%),平均相对误差由 8.5% 优化至 3.6%,峰值误差收敛至 4.6%,精准捕捉节假日、突发事件等复杂场景下的客流波动。该模型可为地铁运营部门提供高精度预测支持,优化运力调度并缓解高峰期拥堵,为城市轨道交通智能化管理提供了创新算法框架。

**关键词:** 地铁客流预测; BP 神经网络; 思维进化算法(MEA-BP); 预测精度

## 引言

地铁客流预测是城市轨道交通系统规划、设计和运营管理的重要组成部分。这对于实现交通资源合理配置,提高公共交通服务水平,促进城市可持续发展具有重要意义<sup>[1]</sup>。准确的客流预测不仅为运力分配、线路设计、车站功能部署等提供量化支持,而且为运力调控、应急处置、票价策略制定等运营层面的决策提供依据。因此,建立科学可靠的地铁客流预测模型成为城市交通研究的重要课题。

本文选取广州地铁 3 号线作为研究对象。该线路是广州地铁线网的一条主干线,客流强度高,全线功能多样,换乘节点密集,具有很强的代表性。系统分析 3 号线客流特性,有助于阐明城市核心轨道线路在不同时间、空间维度上的运行规律,并探究影响这些规律的机制。

研究发现地铁客流的变化受到多种因素的影响。宏观层面包括城市总体发展水平、人口密度分布、经济活动强度、公共交通系统结构等。中间层次主要体现在沿线车站周边功能区布局、就业及居住结构、商业设施密度、教育资源分布等空间因素上。微观层面包括天气条件(降雨、极端温度等)、假期计划、举办重大活动、道路交通状况、地面交通连接便利性等动态因素。这些因素相互交织,使得地铁客流在时间和空间上变化很大,增加了建立预测模型的复杂性。

考虑到上述复杂性,目前的地铁客流预测方法大致分为两类。一类是传统的统计方法,包括多元线性回归、时间序列分析(例如 ARIMA 模型)和卡尔曼滤波。这些方法的特点是模型结构清晰、可解释性强。这些方法适用于变量间

关系稳定的情况,但在处理非线性、强干扰数据时,其预测能力相对有限。二是现代智能方法,主要包括人工神经网络(ANN)、支持向量机(SVM)、随机森林(RF)、深度学习(LSTM 网络、图神经网络 GNN 等)。这些方法在处理高维、非线性、大规模数据方面表现出更强的适应性和预测精度。特别是与大数据平台和实时交通信息相结合,能够动态响应系统变化,具有强大的泛化能力<sup>[2]</sup>。

本文基于数据驱动的建模方法,结合广州地铁 3 号线实际运营数据,深入探究客流变化特征,构建多源异构数据客流预测模型。本研究充分考虑节假日效应、天气变量、地理空间信息等多维因素的交互影响,并引入先进的机器学习算法,提高模型在复杂情况下的适应性和预测精度。最终目的是为广州地铁的运营优化、服务提升以及未来扩建提供科学依据,并为其他大城市轨道交通系统建设初期或调整阶段提供方法借鉴和经验,推动我国城市轨道交通规划向智能化、精细化发展。

## 1. BP 神经网络客流预测分析

BP 神经网络是一种常见的人工神经网络,也称为反向传播神经网络。它可以用来解决分类、回归、聚类等问题。BP 神经网络是一种有监督学习算法,通过对输入样本进行训练,优化网络参数,从而得到一个能够对未知数据进行预测的模型。其核心思想是反向传播算法,它通过计算网络输出与目标输出之间的误差,然后将误差从输出层向输入层反向传递,并根据误差大小调整每个神经元的权值和偏置,使网络的输出与目标输出尽可能接近<sup>[3]</sup>。

本研究构建的BP神经网络模型采用三层拓扑结构,由输入层、单隐层和输出层构成。通过对比实验确定隐层节点数为10,输入层维度为31(对应前7日天气、日期及客流特征),输出层为单节点结构用于第8日客流预测。隐层采用Sigmoid型激活函数(logsig),输出层采用线性传递函数以提升模型非线性映射能力。数据采集涵盖2024年度广州地铁3号线全年的多源异构数据,包括日客流数据、气象参数及日期特征。在数据预处理阶段,对日期变量实施数值编码转换(周循环编码与月份数值化),并对气象参数进行离散化分类编码。采用滑动窗口法构建输入特征矩阵,训练集覆盖全年数据,测试集为2024年3-12月经时序划分的样本,实现7日特征到第8日客流的映射预测。针对数据质量问题,采用插值法处理缺失值,并应用Z-score标准化消除量纲影响。

基于MATLAB R2022b平台进行仿真实验,结果显示:模型预测相对误差分布区间为[2.45%,12.25%],均值为8.5%;绝对误差范围为[0.08,0.96]万人次,平均绝对误差0.66万人次。实验数据表明,该网络架构在客流时序预测中展现出较好的非线性逼近能力,能有效捕捉客流变化的潜在规律。然而,研究发现标准BP算法存在若干固有缺陷:首先,模型性能对初始权值和阈值具有显著敏感性,重复实验显示预测结果标准差达 $\pm 1.2%$ ;其次,梯度下降机制导致存在局部极小值收敛倾向,十次独立训练中40%的案例陷入次优解;再者,模型泛化能力与训练样本规模呈强正相关( $R^2=0.78$ ),当样本量低于200组时预测误差激增32%。此外,隐层节点数等关键超参数仍依赖经验公式( $H \approx \sqrt{(m+n)+a}$ ),缺乏严格的数学推导,交叉验证显示节点数在[8,12]区间外模型性能显著下降,存在结构敏感性风险。针对上述问题,后续研究拟引入思维进化算法等智能优化方法,对网络超参数进行全局寻优,同时结合贝叶斯正则化策略改进目标函数,以期提升模型的鲁棒性和预测精度。

## 2. MEA-BP 高精度预测地铁客流

思维进化算法(Mind Evolutionary Algorithm, MEA)是一种结合群体智能与进化策略的优化算法,通过模拟人类思维中的“趋同”与“异化”机制克服传统遗传算法(GA)的局部收敛缺陷。其核心流程包括:种群初始化:生成包含R个个体的初始种群,每个个体表征BP神经网络的权值与阈值组合;子群体划分:按适应度筛选优胜子群体(表现最优

个体集合)与临时子群体(记录竞争数据),形成M个优胜子群与T个临时子群;趋同操作:在子群内部通过正态分布更新权值,以当前最优个体为中心生成新个体,动态调整方差以平衡搜索速度与精度;异化操作:全局范围内替换低适应度子群,通过迁移优秀个体增强种群多样性,避免早熟收敛<sup>[415]</sup>。

本研究构建的MEA-BP神经网络模型采用层次化进化策略,设置种群规模为400(包含5个优胜子群和5个临时子群),进化迭代次数为200次。网络拓扑结构包含31个输入节点(对应7日历史客流数据)和10个隐含层节点,输出层为单日客流预测值。适应度函数采用复合指标 $S = \beta A + cB$ ,其中A表征隐含层复杂度,B反映权值矩阵维度, $\beta(0.65)$ 和 $c(0.35)$ 为经正交实验确定的权重系数,实现模型复杂度与泛化能力的优化平衡。

基于广州地铁3号线2024年运营数据构建训练集(1-2月)与测试集(3-12月),设置对照实验验证模型有效性。实验结果表明:MEA-BP模型MAE为0.38万人次,较传统BP模型(0.66万人次)降低58.3%;MAPE由8.5%优化至3.6%,峰值预测误差从12.25%收敛至4.6%。误差分布分析显示,MEA-BP模型的绝对误差标准差为0.11万人次,显著低于BP模型的0.29万人次,验证了其抗干扰能力的提升。典型场景预测效能验证:节假日效应:五一黄金周期间形成单日峰值98.74万人次,较基准日74.85万人次,增幅达32.0%,模型预测偏差率 $< 2.1%$ ;突发事件响应:演唱会举办日,6月15日客流激增46.2%、暴雨天气,6月6日客流下降23.4%,模型均实现异常波动特征的有效识别;趋势预测能力:测试期内月均客流增长率为4.3%( $R^2=0.97$ ),与广州市统计局公布的常住人口增长率(3.8%)具有显著相关性( $p < 0.01$ ),准确反映城市发展态势。研究表明,MEA-BP模型通过引入多模态进化机制,有效克服了传统神经网络易陷入局部最优的缺陷,其误差波动范围缩小61.9%,在复杂客流场景中展现出优越的泛化性能。研究成果可为城市轨道交通运营调度提供可靠的预测支持,并为高动态场景下的时间序列预测提供新的方法框架。

## 3. 结论与展望

本研究以广州地铁3号线为研究对象,提出一种基于心智进化算法(MEA)优化的BP神经网络客流预测模型(MEA-BP)。与现有的BP神经网络模型相比,验证了

MEA-BP 模型在客流预测任务中的有效性和优越性。我们的研究表明, MEA-BP 模型显著提高了预测精度。与传统 BP 模型相比, MEA-BP 模型的平均绝对误差 (MAE) 从 6600 降低到 3800, 降低了 58%。平均相对误差 (MAPE) 由 8.5% 优化至 3.6%, 最大误差由 12.25% 收敛至 4.6%。实验结果表明, MEA 算法在全局探索和局部收敛之间取得了良好的平衡, 能够有效提高 BP 神经网络的预测精度。

本研究的创新之处在于将 MEA 的“收敛-发散”机制引入到地铁客流预测领域。MEA 的“收敛”机制可以加速算法接近全局最优, “异化”机制可以防止模型陷入局部最优, 解决了传统 BP 神经网络对初始参数非常敏感, 容易陷入局部最优的问题。此外, MEA 动态调整权重搜索策略, 进一步增强模型在复杂数据环境下的适应性。本文结合多源数据的方法, 综合考虑历史客流数据、天气状况、日期效应等多维特征, 构建了适应高动态城市环境的预测框架, 有效提高了模型的实用性和泛化能力。

虽然 MEA-BP 模型在实验中表现良好, 但是在某些方面仍然存在改进的空间。首先, 实时性能仍然是本研究中面临的挑战。目前的 MEA-BP 模型在处理大规模实时数据时可能会面临一定的计算负担, 影响应用效率和实时性。其次, 随着地铁系统日益复杂, 客流预测的准确性不仅依赖于历史数据, 还与预测极端事件的能力密切相关, 例如节假日期间的大型事件或突发公共卫生事件, 对模型的鲁棒性提出了更高的要求。最后, 多模态数据的融合也是未来研究的重要方向。如何处理来自不同数据源 (例如车站内外摄像头、智能交通系统、社交媒体数据等) 的信息以及如何利用深度学习技术提高对复杂模式进行建模的能力是未来研究中需要解决的紧迫问题。

在未来的研究中, 随着深度学习技术的快速发展, 基于时间序列数据的 LSTM 网络和 Transformer 模型在序列数据的建模能力上具有显著优势, 可以更好地捕捉地铁客流的时间依赖性。将图神经网络 (GNN) 与地铁网络拓扑建模相结合, 可以进一步提高模型在复杂交通网络中的适用性。未来我们可以考虑结合多种数据源 (例如传感器数据、社会行

为数据、天气变化等) 来构建更全面的预测框架, 并使用深度学习和强化学习算法整合多模态数据, 以理解不同类型数据之间的复杂相互作用。同时, 针对突发事件、重大事件等极端事件提出有针对性的解决方案预测, 提高异常情况下客流预测模型的准确性和鲁棒性。

为了进一步提高模型在变化环境中的适应性, 我们开发了一种基于实时数据的自适应优化算法, 该算法可以根据客流波动的强度动态调整种群规模、迭代次数、MEA 的变异概率等参数。这增强了模型对不同操作场景的自主适应性。自适应机制的设计可以提高模型在高度动态和信息匮乏的环境中实时响应的能力, 增强其在实际作战中的适用性。

综上所述, 本研究提出的 MEA-BP 模型为地铁客流预测提供了有效的解决方案。通过将进化算法引入预测模型优化, 提高了客流预测的准确性和可靠性。未来研究将在现有成果的基础上, 进一步提升模型的实时性、多模态数据融合能力以及对极端现象的预测能力, 推动地铁客流预测技术朝着智能化、精细化方向发展。

#### 参考文献:

- [1] 叶森, 吕兴, 李盛杰, 等. 基于神经网络算法的城市轨道交通网弹性评估 [J/OL]. 交通运输工程与信息学报, 1-29[2025-04-08]. <https://doi.org/10.19961/j.cnki.1672-4747.2025.03.018>.
- [2] 董慧, 潘晓, 郭景峰, 等. 基于时空图神经网络的城市路网行程时间预测研究综述 [J]. 燕山大学学报, 2025, 49(02): 95-105
- [3] 张佳佳, 陈静. 深度学习交通流预测模型现状与未来 [J]. 天津职业技术师范大学学报, 2025, 35(01): 53-58+72.
- [4] Ding X and Liu Z and Chen W et al. The Forecast Algorithm of Stranded Passenger Flow Based on the Angle Cost of Urban Rail Line [J]. Journal of Computational and Theoretical Nanoscience, 2016, 13(9): 5724-5730(7).
- [5] Wu K, Wu Z, Ning C. Subway Passenger Forecast Based on System Dynamics [J]. Open Journal of Transportation Technologies, 2017, 6(3): 77-85