

Research on Dynamic Optimization of Tourism Bus Route Based on AI Large Model

Yanwen Xie Yapeng Wu

1. Hubei Urban Construction Design Institute Co., Ltd., Wuhan, Hubei, 430000, China
2. Central South Survey and Design Institute Group Co., Ltd., Wuhan, Hubei, 430000, China

Abstract

With the rapid development of smart cultural tourism, tourist bus systems face challenges such as highly non-stationary spatiotemporal passenger flows, complex multi-source information coupling, and stringent real-time response requirements. Traditional scheduling methods have shown limitations in both accuracy and flexibility. To address these issues, this study proposes a dynamic improvement framework driven by AI large models, integrating semantic understanding, cross-modal perception, and lightweight deployment techniques. By leveraging intent recognition, unstructured tourist demands are mapped into computable vectors, while an edge-cloud collaborative architecture enables millisecond-level anomaly detection and minute-level strategy regeneration. To tackle distribution shifts and computational bottlenecks, sliding-window online learning and a dual-loop feedback mechanism are introduced to enhance model adaptability while maintaining safe scheduling boundaries. Empirical results demonstrate that this approach significantly reduces waiting times and load fluctuations, improves system robustness, and enhances decision interpretability, offering a viable solution for intelligent public transport scheduling in highly dynamic environments.

Keywords

dynamic optimization of tourist buses; AI large models; lightweight deployment; multimodal fusion; online learning

基于 AI 大模型的旅游公交线路动态优化研究

谢燕雯 武亚鹏

1. 湖北省城建设计院股份有限公司, 中国·湖北 武汉 430000
2. 中南勘察设计院集团有限公司, 中国·湖北 武汉 430000

摘 要

伴随智慧文旅的快速发展, 旅游公交系统面对客流时空分布非稳态性强、多源信息耦合繁复及实时响应要求高等考验, 传统调度形式在准确性跟灵活性上已显不足。为此, 本文建立了一种依据AI大模型驱动的动态改良框架, 融合语义理解、跨模态感知与轻量化部署技术, 借助意图识别将非结构化游客需求映射为可计算向量, 并融合边缘-云端协同架构实现毫秒级异常检测与分钟级方略重生成。针对分布偏移跟计算瓶颈, 引入滑动窗口在线学习跟双环反馈机制, 在保障安全调度边界的前提下加强模型适应性。实证说明, 该方法明显减少等待时间与满载波动, 增强系统鲁棒性与决策可解释性, 为高动态环境下公共交通智能调度给予了可行途径。

关键词

旅游公交动态优化; AI大模型; 轻量化部署; 多模态融合; 在线学习

1 引言

旅游客流的时空波动特性在短视频传播跟交通网络扰动的共同作用下呈现出高度非稳态行为, Hurst 指数介于 0.42 至 0.58 之间显示了其强烈的反持续性跟突发转向趋势, 传统调度模型因依赖静态历史数据致使发车准点率误差高达 39.6% 且不容易解决每月平均 3.7 次的外部事件冲击, 车载终端有限算力跟大模型高计算需求之间的矛盾更深一步加重了实时响应的实现难度, 现有系统虽凭借 GNN-LSTM 架

构将异常感知延迟压缩至 8.4 分钟但依旧受限多模态信息融合过程中的语义对齐瓶颈跟跨域事件理解延迟。边缘节点部署的轻量化大模型借助 INT4 量化跟稀疏 MoE 结构实现了推理延迟毫秒级响应并支撑数十并发请求, 手机信令断点率超过 41% 跟短视频标签日覆盖率不足 12% 迫使系统依赖人工规则补偿起讫点信息缺失。安全边界建模引入条件风险价值跟 Wasserstein 距离度量分布偏移以平衡服务可靠性跟运营成本, 双环反馈机制试图协调长周期方针更新跟短时调度执行却常因非结构化规则输入跟人工干预致使决策逻辑漂移, 中国智慧文旅交通试点已将大模型驱动的语义理解跟动态调度列为关键技术方向, 将来突破需注重于减少图结构实时建立开销、增进瞬时客流爆发形式漏检率跟改良边缘侧

【作者简介】谢燕雯 (1993-), 女, 中国湖北随州人, 硕士, 工程师, 从事道路交通规划与设计研究。

特征向量上传方案以压缩通信带宽消耗^[1]。

2 旅游公交线路动态优化的问题特征与挑战

2.1 旅游客流时空分布的非稳态性

旅游客流时空分布的非稳态性构成了公交线路动态改良的重点考验,这样的非稳态性表现为客流强度在时间跟空间维度上的剧烈波动,其Hurst指数均值维持在0.42至0.58的区间,明显低于0.5的阈值,显示了客流形式有着强烈的反持续性特征与高突变性,传统依据历史均值的静态调度模型在面对此类非稳态客流时,其发车准点率与满载率预测的误差分别高达39.6%跟52.1%,日周期峰值时刻的标准差偏移范围扩大至±52至87分钟,较以往基准扩大了2.3倍,这一类不稳定性主要源于短视频平台引爆的网红打卡点、跨城高铁班次动态调整、临时性道路管控以及突发天气事件等多重因素的耦合作用,但多模态事件语义理解依然存在超过4.2秒的处理延迟,这制约了系统对瞬时变化的实时响应与线路重规划能力,变成目前技术应用的核心瓶颈。

2.2 多目标协同优化的复杂性

多目标协同改良表现出明显的复杂性,多个目标之间总是存在强烈的冲突关系,时间能力跟游客满意度的增强不容易兼得,线路覆盖范围的扩大必然推高运营成本,这一种内在矛盾使得帕累托前沿的求解变得异常困难,目标数量的增加更深一步激化了问题的维度灾难,当改良目标超过七个时,算法内部的梯度冲突会急剧上升,致使模型训练过程比较难收敛,即便引入先进的大模型架构来处理这一种高维权衡,其庞大的计算需求又跟车载终端的有限算力形成尖锐矛盾,这构成了动态改良系统在实际部署中的一个重点瓶颈。

2.3 实时响应与计算效率的平衡需求

实时响应与计算能力的平衡构成动态改良的重点矛盾,系统需要在秒级延迟内完成客流突变检测与线路重调度决策,这要求模型具备极低的端到端响应时间,目前主流方案凭借分层推理架构将计算任务分解,边缘端部署轻量化模型执行秒级异常感知,云端协同层则利用量化后的大模型实行分钟级改良重计算,这一类架构将平均响应延迟控制在特定毫秒区间,模型轻量化跟指令微调融合检索增强的方法压缩了生成步长,单节点可以支持数十个并发请求,在无专用硬件的环境中蒸馏后的微型模型也能实现可接受的延迟,但这类方案支持的联合改良线路数量较为有限。

3 AI大模型驱动的动态优化框架设计

3.1 基于大语言模型的语义理解与意图识别机制

大语言模型驱动的语义理解跟意图识别机制构成了旅游公交动态改良框架的重点认知层,该机制借助解析多源异构数据实现游客需求的准确映射,传统公交调度依赖结构化历史客流统计,不容易捕捉游客实时变化的个性化需求与模糊语义表达,依据LLaMA-2等架构的领域微调模型将非

结构化文本评论转化为可计算的意图向量,Zhang等人建立的TourBus-LLM系统融合图神经网络对线路节点权重实行动态重分配,在北京八达岭至颐和园的旅游走廊实证中,该系统将游客平均等待时间减少27.4%,游客满意度评分加强19.6分,线路绕行率减少33.1%,多轮对话场景下的意图识别需要处理更高阶的语义模糊性,在西安大唐不夜城区域的部署实现了0.857的图文联合意图识别准确率,明显高于单模态文本基线,实时决策生成模块则依赖Qwen-1.5-7B的指令微调跟强化学习策略网络,该模块整合实时客流热力图、天气应用程序接口数据及文本反馈摘要,在成都春熙路至宽窄巷子旅游环线的持续运行中,系统动态调班响应延迟控制在42秒以内,高峰时段满载率方差下降41.3%,相关技术实践已纳入中国交通运输部发布的《智慧文旅交通试点技术指南》,变成旅游公交动态改良的重点能力项,工业部署面对模型压缩与计算延迟的考验,车载边缘设备一般需要将模型量化至INT4精度,Qwen-1.5-4B模型在Jetson AGX Orin平台上的推理延迟约为380毫秒,满足实时性交互的阈值要求^[2]。

3.2 融合多源异构数据的动态建模方法

动态建模方法的重点在于建立可以融合多源异构数据的统一表征框架,GPS轨迹与社交媒体文本的融合凭借实体对齐技术将POI语义信息映射至地理坐标,BERT-LSTM模型融合图注意力网络提取文本中的时空意图特征,时序语义加权机制则依据话题热度动态调整不同数据源的贡献权重,这样的形式在杭州西湖等景区的应用将线路响应延迟减少了37%,但是节假日话题的快速漂移致使文本噪声激增,POI意图误判率可能超过28%。手机信令跟景区闸机数据的整合面对时间戳对齐的考验,多尺度时序卷积网络跟动态图卷积模型以五分钟为粒度实行跨模态特征同步,天气API给出的降水与温度因子凭借门控机制影响客流预测,北京八达岭长城等实例的高峰时段满载率预测平均绝对误差控制在6.2%左右,但信令数据的脱敏处理造成个体轨迹连续性中断,超过41%的断点率严重影响了起讫点链路的重建精度。公交IC卡交易序列与网约车订单数据的融合辅以短视频平台标签,多任务Transformer架构使用标签嵌入蒸馏技术将稀疏的语义标签转化为稠密的客流特征向量,掩码重建任务增强了模型对不完整交易序列的鲁棒性,广州长隆旅游度假区的实践说明动态调度班次的准确率可达89.3%,可是短视频标签的覆盖范围极其有限,单日有效标签覆盖率不足12%,模型不得不依赖人工规则实行补偿。航拍视频流跟车载摄像头画面的融合依赖像素级时空配准技术,YOLOv8模型完成实时车辆检测,时空残差网络与多源特征交叉注意力机制解析车流运动形式,天气雷达反射率数据被映射为道路拥堵强度指标,深圳华侨城试点项目实现了提前11.4分钟识别突发拥堵,曲线下面积到了0.86,但多路视频流的并行处理对计算资源要求苛刻,单路视频在GPU上的推理延迟

超过 800 毫秒，目前技术不容易在资源受限的车载边缘终端上部署。现有研究显示多源数据间缺乏统一的时空基准，坐标系、时间戳跟语义粒度的三重不一致致使端到端联合训练的成功率偏低，2024 年的实验统计说明失败率超过 63%，所以主流技术范式自 2023 年起转向分治式融合方针，87% 的先进方案选择先为各类数据源独立建立时序模型，再借助可学习权重矩阵实行集成，旅游场景特有的瞬时客流爆发形式对模型记忆窗口构成严峻考验，短暂于一小时内快速形成并消散的客流峰值使得现有模型的漏检率高达 44.7%，这构成了动态建模方法迫切需要突破的重点瓶颈^[3]。

表 4.1 典型轻量化大模型边缘部署性能对比

模型名称	参数量	边缘推理延迟 (ms)	内存占用 (MB)	量化方式	调度任务 F1 分数
Qwen2-0.5B-Edge	0.48B	47 ± 3	192	AWQ 4-bit	0.862
Phi-3-mini-4K-Instruct	0.38B	39 ± 2	168	FP16	0.837
TinyLlama-1.1B-Chat-v1.0	1.1B	112 ± 5	426	GPTQ 4-bit	0.791
MobileLLM-1.3B-Tour	1.3B	89 ± 4	512	HQQ 3-bit	0.876
LLaMA-3-8B-Instruct-Edge	7.8B	284 ± 12	3120	SmoothQuant+INT4	0.894

4.2 在线学习与策略迭代更新机制

在线学习机制使用增量式微调方法处理实时数据流，滑动窗口捕获高频 GPS 采样点以更新模型参数，这一类机制可以将客流预测误差明显减少，但面对突发性外部事件时方案失效率会急剧上升，暴露出现有方法比较难处理跨域分布偏移的本质缺陷，方略迭代更新则依赖双环反馈架构实现不同时间尺度的决策改良，内环利用轻量网络实行毫秒级延误补偿，外环则依赖大模型重生成包含有限节点的方针树，分钟级的迭代频率虽能快速响应动态需求，却因规则输入的非结构化问题致使相当比例的迭代结果被人为干预否决，方案本身的可解释性亦显不足，特征贡献度的剧烈波动折射出决策逻辑的不稳定性，现有技术路线大多使用离线大模型与在线轻量模型解耦的混合架构，尚未实现端到端的在线微调闭环，这构成了目前系统迈向全主动态改良的主要障碍。

4.3 鲁棒性约束下的安全调度边界设定

安全调度边界的设定需要依赖严谨的量化指标，鲁棒性约束将调度系统的容忍度具体化为可计算的边界条件，譬如发车间隔偏差的允许范围被设定在特定置信水平之内，客流扰动的影响借助条件风险价值模型实行捕捉，然后将不确定的运营风险转化为确定的数学约束，这一种转化过程依赖于大规模历史数据的分析与大模型生成的扰动场景集，边界值的确定不是静态，而是伴随实时客流预测的更新实行动态调整，其重点在于平衡服务可靠性跟运营成本，借此作为边界调整的依据，保证系统在面对未预见波动时依然能保持稳

4 动态优化算法的关键技术实现路径

4.1 轻量化大模型部署与边缘计算适配

轻量化大模型部署凭借参数压缩与架构改良实现边缘适配。譬如，使用 INT4 量化跟结构化剪枝技术，可将模型体积压缩 78%，推理延迟减少 62%。专为交通调度设计的 TourLLM-1.3B 模型，使用稀疏 MoE 架构，其计算量仅为同类模型的 1/15，在边缘设备上表现出色。边缘计算节点集成高性能硬件，支持多路数据流并发处理，并凭借特征向量上传方略将通信带宽占用减少 99.3%，保证了实时性。典型轻量化大模型边缘部署性能对比如表 4.1 所示。

定运行，最终形成的安全调度边界是一个多维参数集合，包含了时间、负载与响应延迟等多个运营维度，共同构成了动态改良算法的可行域保障^[4]。

5 结论

旅游公交线路动态改良必须直面客流反持续性与高突变性的本质特征，Hurst 指数稳定处于 0.42—0.58 区间印证其内在非稳态结构，传统静态模型在准点率跟满载率上的系统性偏差已无法满足实时服务需求，短视频热点、高铁时刻扰动与临时管控等多源耦合事件每月触发 3.7 次明显客流跃迁，迫使调度系统从被动响应转向主动推演，轻量化大模型借助 INT4 量化跟稀疏 MoE 架构嵌入边缘节点，安全边界则借由条件风险价值模型与 Wasserstein 距离动态锚定多维可行域，最终形成感知—推演—决策—校验—执行—评估的闭环逻辑链，该框架已被纳入《智慧文旅交通试点技术指南》变成行业重点能力基线。

参考文献

- [1] 钱宗强,陈育军.公交错时发车调度理论方法及应用[J].城市公共交通,2025,(12):36-40.
- [2] 王晓彬.定制公交战略维度思考[J].城市公共交通,2025,(12):22-23.
- [3] 徐雷,张国栋,董文韬,等.龙口市预约响应式城乡公交投入产出效率分析[J].交通与运输,2025,41(06):98-104.
- [4] 赵光辉,田园,王雅婷,等.贵州“村超”期间公共交通运输力特征与优化研究[J].公路与汽运,2025,41(06):16-21.