

基于梯度提升决策树的城市轨道交通网络运营态势综合评估方法*

段力伟¹ 冉松民¹ 陈瑞雪¹ 邝雅静¹ 董少帅¹ 欧光继²

(1. 重庆交通大学交通运输学院, 400074, 重庆; 2. 重庆交通大学信息科学与工程学院, 400074, 重庆//第一作者, 副教授)

摘 要 以客流规模、运行效率、运营安全以及工作负荷作为一级指标,建立包含 20 个二级指标的城轨网络运营态势评估指标体系。对运营态势预警等级进行了划分,提出了以梯度提升决策树算法实现网络运营态势综合评估的研究思路。以重庆市轨道交通网络数据为研究对象,进行运营态势预警等级的综合评估验证,验证了网络运营态势综合评估方法的可行性。

关键词 城市轨道交通网络;运营态势;预警等级;梯度提升决策树

中图分类号 F530.7

DOI:10.16037/j.1007-869x.2022.08.007

Method of Comprehensive Evaluation on the Operation Situation of Urban Rail Transit Network Based on Gradient Boosting Decision Tree

DUAN Liwei, RAN Songmin, CHEN Ruixue, KUANG Yajing, DONG Shaoshuai, OU Guangji

Abstract Taking passenger flow scale, operation efficiency, operation safety and workload as primary indicators, an urban rail network operation situation evaluation indicator system comprising 20 secondary indicators is established. Operation situation early warning grades are divided, and the research idea of using Gradient Boosting Decision Tree algorithm to achieve a comprehensive evaluation of operation situation is proposed. Taking Chongqing rail transit network data as research object, by carrying out comprehensive assessment and verification of the operation situation early warning grades, the feasibility of the network operation situation comprehensive evaluation method is confirmed.

Key words urban rail transit network; operation situation; early warning grade; Gradient Boosting Decision Tree

First-author's address College of Traffic and Transportation, Chongqing Jiaotong University, 400074, Chongqing,

China

城市轨道交通(以下简为“城轨”)网络运营态势是对城轨网络的客流规模、运行效率、运营安全以及工作负荷的综合反映。城轨网络运营态势评估即通过分析影响城轨运行的各项因素得到城轨网络化运营管理与服务的水平,进而判断城轨网络是否处于安全稳定的运营状态。

目前,已有对城轨大客流预警需求的研究和分析^[1],通过确立预警指标和阈值进行城轨运营安全评估^[2],采用云模型对预警指标进行融合来确定预警阈值^[3],以及通过图像识别^[4]、支持向量机^[5-6]及时间序列分析法^[7]等技术或方法对可能带来安全隐患的短时大客流进行预警。而这些文献多为对客流预测的分析和预警方法的研究,仍缺少对城轨运营网络运营态势整体的综合评估分析,存在评估指标体系建设不够全面等问题。在已有研究的基础上,本文通过分析城轨网络运营态势的影响因素,确立多维度评估指标体系,划分了 4 个预警等级,综合使用梯度提升决策树算法对城轨运营网络进行运营态势等级评估。

1 指标体系和预警等级的建立

1.1 运营态势评估指标体系

按全面性、科学性、实用性的选取原则,城轨网络运营态势预警指标体系包含客流、效率、安全及工作负荷等 4 个一级指标,以及 20 个二级指标。

1.1.1 客流指标

客流指标主要涵盖影响运营态势的客流规模因素指标,以及由此导致的线网及站点的客流风险指标。

* 重庆市基础科学与前沿技术研究项目(cstc2017jcyjAX0218);重庆市教委科学技术研究项目(KJQN201800730)

客流指标包含 4 个二级指标: 全网客流总量 Q_1 , 即从全网各车站乘坐城轨到达目的地的客流总量; 全网换乘客流总量 Q_2 , 即全网所有换乘站内进行换乘的客流总量; 线路客流量不均衡系数 K_1 ; 站点进站负荷系数 K_2 。

其中:

$$K_1 = \frac{\sum_{i=1}^n \left(q_i - \frac{\sum_{i=1}^n q_i}{n} \right)}{n} \quad (1)$$

式中:
 n ——全网线路总数;
 q_i ——第 i 条线路在统计时间和统计范围内的客流量, $i = 1, 2, \dots, n$ 。

$$K_2 = \frac{\max(q_j)}{\frac{1}{u} \sum_{j=1}^u q_j} \quad (2)$$

式中:
 u ——全网车站总数;
 q_j ——统计时间内第 j 个车站的进站客流量, $j = 1, 2, \dots, u$ 。

1.1.2 效率指标

效率指标主要反映城轨网络运行效率, 其包含 4 个二级指标: 列车正点率 Z , 列车停站时间不均衡系数 K_3 , 一次滞留乘客数 Q_3 , AFC(自动售检票)系统的故障设备数 T_1 。

其中:

$$Z = \frac{2m - \sum m_f - \sum m_d}{2m} \times 100\% \quad (3)$$

式中:
 m ——统计时间内, 全网实际开行的列车数;
 m_f ——始发晚点列车数;
 m_d ——到达晚点列车数。

$$K_3 = \frac{\max t_{j, \text{停}}}{\frac{1}{u} \sum_{j=1}^u t_{j, \text{停}}} \quad (4)$$

式中:
 $t_{j, \text{停}}$ ——列车在第 j 个车站的停站时间。

1.1.3 安全指标

安全指标主要为影响列车运行安全和乘客乘车安全的各项因素指标。
安全指标包括 9 个二级指标: 列车临时停车总次数 E_1 , 列车发生故障的总次数 E_2 , 全网信号设备

发生故障的总次数 E_3 , 全网自动扶梯发生故障的总次数 E_4 , 全网供电系统发生故障总次数 E_5 , 全网屏蔽门发生故障总次数 E_6 , 网络中突然发生重大安全事故的总次数 E_7 , 人员纠纷等不安全事件发生总次数 E_8 , 安检发现危险物品的总检出次数 E_9 。

1.1.4 员工工作负荷指标

员工工作负荷指标为主要考虑影响列车运行安全关键岗位人员的工作时长因素指标。

员工工作负荷指标包含 3 个二级指标: 在统计时间内列车驾驶员平均工作时长 t_a , 驾驶员工作时长最大值 t_m , 行车调度员在统计时间内的平均工作时间占有率 t_h 。

其中:

$$t_a = \frac{\sum t_z}{N} \quad (5)$$

式中:
 t_z ——第 z 个驾驶员当日的工作时长;
 N ——全网列车驾驶员人数。

$$t_h = \frac{t_p}{t_j} \quad (6)$$

式中:
 t_p ——行调员在统计时间内的平均工作时长;
 t_j ——统计时间。

1.2 预警等级划分

根据运营态势评估指标体系及《城轨运营期间安全评估规范》, 将城轨网络运营态势预警等级划分为四级, 如表 1 所示。

表 1 城轨网络运营态势预警等级		
Tab.1 Urban rail network operation situation early warning grades		
等级	等级值	运营态势
I 级	1	运营状况良好, 无明显运营风险
II 级	2	客流短时聚集, 呈运营风险加剧趋势
III 级	3	客流通行速度下降, 关键设备故障
IV 级	4	突发严重事故, 客流拥挤加剧

2 运营态势综合评估模型的构建

2.1 梯度提升决策树的原理

GBDT(梯度提升决策树)算法由 Jerome Friedman 于 1999 年提出^[8], 其在训练过程中确定一系列合适的决策树作为弱学习器, 并将弱学习器线性组合成为强学习器后对测试样本进行判断。

对于训练集 $M = \{(x_1, y_1), (x_2, y_2), \dots, (x_N, y_N)\}$, GBDT 算法的具体模型表示形式为:

$$F_S(x) = \sum_{s=1}^S T(x; \Phi_s) \quad (7)$$

式中:

$F_S(x)$ ——预测函数;

x ——训练集数据;

$T(x; \Phi_s)$ ——第 s 个决策树变量;

Φ_s ——第 s 个决策树参数;

S ——决策树个数。

模型的迭代步骤如下:

步骤 1: 确定初始的提升树为 $F_0(x) = 0$ 。

步骤 2: 采用向前分布算法, 第 v 步的算法模型为:

$$F_v(x) = F_{v-1}(x) + T(x; \Phi_v) \quad (8)$$

式中:

$F_{v-1}(x)$ —— $v-1$ 步的预测函数模型。

步骤 3: 通过要求得到的预测值 $F(x)$ 使得损失函数 L 最小, 来确定第 v 个决策树参数 Φ_v , 即:

$$\Phi_v = \underset{\Phi_v}{\operatorname{argmin}} \sum_{i=1}^N L[y_i, F_{v-1}(x_i) + T(x_i; \Phi_v)] \quad (9)$$

式(9)中, L 为平方误差损失函数, 即:

$$L[y, F_{v-1}(x) + T(x; \Phi_v)] = [y - F_{v-1}(x) - T(x; \Phi_v)]^2 \quad (10)$$

通过不断迭代, 使得最后得到的输出结果不断逼近训练值。

2.2 基于 GBDT 算法的城轨运营态势评估

本文以训练集中的 4 个预警等级、2 个样本特征指标 (Q_1 及 E_1) 为例, 叙述 GBDT 算法模型通过拟合残差进行学习的过程。

城轨网络运营态势评估模型首先通过训练集 $M = \{(x_1, y_1), (x_2, y_2), \dots, (x_N, y_N)\}$ 找到目标函数。其中, x 为特征指标值, $x_i \in X \subseteq R_n$; y 为预警等级对应值, $y_i \in Y = \{1, 2, 3, 4\}$ 。随后, 使得 x 映射到 y 的损失函数 L 最小, 来确定决策树参数, 得到下一棵决策树; 最后, 将所有决策树节点值进行累加, 得到输出的评估结果。以 Q_1 、 E_1 指标为例, 根据初始数据生成决策树 1, 树的每一个节点相应有一个输出值; 然后, 根据树 1 的残差生成决策树 2, 当决策树节点值为 0 时, 结束迭代。将对应决策树节点值进行累加得到模型预测结果。过程见图 1, 如预警等级 I 的评估值 $y =$ 树 1 左节点 + 树 2 左节点 = 1。

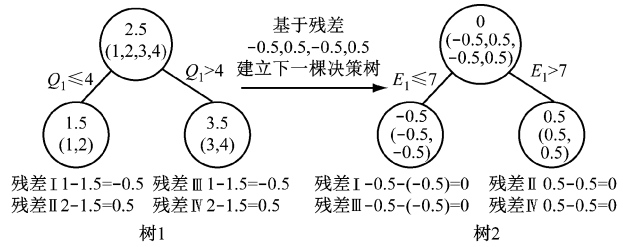


图 1 GBDT 算法通过拟合残差的迭代过程
Fig. 1 GBDT algorithm iteration process by fitting residuals

3 案例验证

根据城轨网络运营态势评估指标体系, 本文以重庆轨道交通客流量综合分析平台的数据为基础, 对评估模型进行验证。

选取全网客流总量 Q_1 、屏蔽门故障次数 E_6 、列车停站不均衡系数 K_1 以及行调员平均工作负荷占有率 t_h 作为典型评估指标, 绘制预警等级评估结果与典型评估指标的时间曲线如图 2 所示。

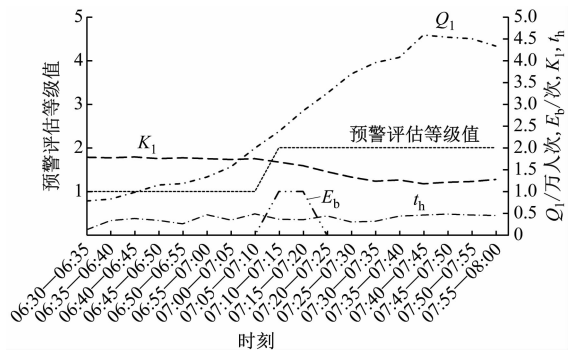


图 2 预警等级评估值与典型评估指标的时间曲线
Fig. 2 Time history curve of early warning grade evaluation values and typical evaluation indicators

图 2 可反映重庆城轨网络当日的运行态势。在运营初期(06:30—07:15), 城轨网络客流基本保持平稳, 预警评估等级维持在 I 级。在 07:15, 发生屏蔽门设备故障, 且客流量呈现递增趋势, 预警等级评估模型的评估等级变为 II 级。预警等级评估提升的原因并不是客流量的大幅上升, 而是运营过程中出现的设备故障。如果关键设备发生故障, 则会导致客流流速减缓, 并带来大量乘客滞留的风险。07:25 屏蔽门故障设备被修复, 但由于城轨网络运营时间临近早高峰, 客流量持续上升, 预警等级评估结果维持为 II 级, 城轨网络出现客流短时间聚集等状况, 有运营风险加剧的趋势。

城轨网络运营态势的预警等级在多维评估指

标的影响下是变化的。客流规模(Q_1 、 Q_2 、 K_1 、 K_2)并非判断运营态势的唯一因素,城轨网络运行效率(Z 、 K_3 、 Q_3 、 T_1)、安全状态($E_1 \sim E_9$)及员工工作负荷(t_a 、 t_m 、 t_h)都是影响城轨网络运营态势的必要因素。

4 结语

及时掌握城轨网络运营态势,进行多维度的综合评估与预警,是保障城轨网络运营安全、提高运营效率的有力保障。本文通过分析城轨网络运营态势的影响因素,提出了预警评估指标体系,基于梯度提升决策树算法建立城轨网络运营态势评估模型,并以重庆轨道交通网络为实例,进行了预警等级评估验证。验证结果表明:运营态势评估结果能基本反映城轨网络运营现状及发展趋势;模型具有可行性;进行城轨网络运营态势评估有助于保障城市轨道交通网络的运营安全,提高运营效率。

参考文献

[1] 王婵婵,张琦. 上海城市轨道交通日常大客流检测及预警需求研究[J]. 城市轨道交通研究,2018(10):11.
WANG Chanchan, ZHANG Qi. Requirements for daily large passenger flow detection and early warning in Shanghai urban rail transit[J]. Urban Mass Transit, 2018(10):11.

[2] 张晗. 城市轨道交通运营安全综合评估预警平台设计研究[D]. 北京:北京交通大学,2012.
ZHANG Han. Study on the design of urban rail transit operational safety assessment early warning platform[D]. Beijing: Beijing Jiaotong University, 2012.

[3] 潘杰. 城市轨道交通车站客流预警与控制研究[D]. 西安:长安大学,2019.

PAN Jie. Research on passenger flow early warning and control of urban rail transit station[D]. Xi'an: Chang'an University, 2019.

[4] 张霖. 北京城市轨道交通大客流辨识与安全状态评估技术及系统[D]. 北京:北京交通大学,2011.
ZHANG Lin. Techniques and system about identification of large scale passenger flow and assessment of security status of Beijing Mass Transit Railway [D]. Beijing: Beijing Jiaotong University, 2011.

[5] 张清泉. 轨道交通客流预警系统研究[D]. 重庆:重庆交通大学,2016.
ZHANG Qingquan. Research on the urban rail transit passenger flow early-warning system[D]. Chongqing: Chongqing Jiaotong University, 2016.

[6] 谢俏,叶红霞. 基于支持向量机的节假日进出站客流预测方法[J]. 城市轨道交通研究,2018(8):26.
XIE Qiao, YE Hongxia. Forecast for holiday passenger flow at urban rail transit station based on support vector machine model[J]. Urban Mass Transit, 2018(8):26.

[7] 潘罗敏. 地铁短时客流量预测预警研究[D]. 北京:首都经济贸易大学,2011.
PAN Luomin. Study on short-term passenger flow forecast and early warning of metro[D]. Beijing: Capital University of Economics and Business, 2011.

[8] FRIEDMAN J H. Greedy function approximation: a gradient boosting machine[J]. The Annals of Statistics, 2001,29(5):1189.

[9] 关菲菲,张帆,唐志远. 重庆轨道交通客流量分布特征[J]. 城市轨道交通研究,2020(5):97.
GUAN Feifei, ZHANG Fan, TANG Zhiyuan. Distribution characteristics of urban rail transit passenger flow in Chongqing City[J]. Urban Mass Transit, 2020(5):97.

(收稿日期:2020-05-31)

上海将新增地方债券支持轨道交通 21 号线、23 号线和 18 号线等重大项目建设

2022 年 7 月,上海市发行今年第二批地方政府债券共计 314.4 亿元。新增债券额度聚焦重大项目建设,优先保障自贸试验区临港新片区、虹桥国际开放枢纽等国家重点战略,以及市委、市政府重点推进的重大公共基础设施建设项目,积极支持各区推进补短板、促投资、稳增长、惠民生的建设项目,持续发挥地方政府债券带动扩大有效投资的作用。

上海市发展和改革委员会相关负责人表示,用好专项债券十分重要,一是为了加快补齐交通、能源、水利、生态环保、公共服务、城镇基础设施等领域的短板,增强新兴产业发展动能;二是稳住经济基本盘;三是发挥有效投资的放大效应。2020 年和 2021 年的地方政府专项债券共发行 1 666 亿元,项目覆盖国铁、市域铁路、轨道交通、污水处理、棚户区改造等多个领域,为 2020 年和 2021 年上海市固定资产投资分别增长 10.3%、8.1% 做出了积极贡献。

上海市财政局介绍了 7 月 18 日上海发行 314.4 亿元新增债券的基本情况和发行特点。申通集团、久事集团、城投集团三家项目单位介绍了本市轨道交通、市域铁路、污水设施领域规划建设情况,对本次将发行的轨道交通 13 号线西延、18 号线二期、21 号线一期、23 号线一期,泰和污水处理厂扩建,市域铁路嘉闵线等专项债券项目进行了重点推介,这些项目均已开工,总投资近 1 500 亿元。

(摘编自 2022-07-13《浦东发布》)