

融合数据分析的工程成本超支预警模型与决策支持系统

陈思琦 刘诗妍 白琰玺

西安市物联网应用实验室 陕西西安 710000

摘要: 本文构建了基于多源数据融合的工程成本超支智能预警体系, 通过 BIM-ERP 集成实现施工全过程数据采集, 采用 LSTM-蒙特卡洛混合模型进行成本趋势预测, 结合数字孪生技术建立成本镜像系统。系统包含动态特征提取、多级预警触发、智能决策优化三大核心模块, 在实际工程中实现成本偏差提前 6 个月预警, 关键路径成本管控精度达 92%。

关键词: 工程成本管控; 混合预测模型; 数字孪生; 智能决策系统; 多源数据融合

1 数据驱动型成本监控框架设计

1.1 工程成本数据全生命周期采集技术

在工程项目实施过程中, 成本数据的完整性直接影响管控模型的精确性。本研究提出基于 BIM+ERP 架构的多源数据融合方案, 通过 IFC 标准实现建筑信息模型与企业管理系统的数据贯通。在施工准备阶段, 利用参数化建模技术将合同清单分解为 WBS 工作包, 并嵌入材料价格波动指数与劳务市场基准费率, 构建三维成本基准体。

现场数据采集采用边缘计算节点部署方案, 在塔吊、施工电梯等关键设备加装物联网传感器, 实时捕获混凝土浇筑方量、钢筋绑扎进度等物理量数据。针对非结构化文本信息, 研发基于 BERT 架构的施工日志解析引擎, 通过语义角色标注技术提取“窝工 4 小时”、“泵车故障”等事件型数据, 并与进度款支付凭证形成交叉验证链。特别地, 建立动态成本基准线校准机制, 当实际进度偏差超过 5% 时, 自动触发 SVM 分类器重新标定预算分配权重。

1.2 实时数据清洗与质量管控策略

原始工程数据存在 30%-45% 的噪声比例, 需构建三层过滤体系进行质量治理。首层部署改进型孤立森林算法, 通过构建随机超平面划分特征空间, 实现对异常成本项的秒级筛查。

针对施工日志中的缺失数据, 创新应用时空克里金插值法。该方法融合 GPS 轨迹数据与工序标准工时库, 在空间维度建立作业面网格化模型, 时间维度采用 ARIMA 模型预测缺失时段可能发生的工程量。

为解决跨阶段数据断层问题, 设计基于知识图谱的关联验证引擎。将设计变更、签证单、认价单等文档转化为

RDF 三元组, 通过 SPARQL 查询语言检测“变更内容未计价”、“措施费重复计取”等逻辑矛盾。在某商业综合体项目中, 成功拦截 17 项隐蔽性成本风险, 涉及金额 2300 余万元。该框架已封装为 Docker 镜像, 支持 Kubernetes 集群部署, 具备横向扩展能力。

2 超支风险特征智能提取方法

2.1 多维度风险因子量化建模

工程成本超支的显性诱因往往隐藏于地质条件、供应链波动及政策变更等非线性因素中。本研究构建三维地质建模引擎, 将钻孔数据与物探报告映射至八叉树空间结构, 通过克里金插值生成连续地质曲面, 量化岩层参数误差对桩基工程成本的影响系数。实验表明, 当岩层倾角预测偏差超过 8° 时, 嵌岩桩施工成本将产生 12%-18% 的级联增量。

针对材料价格波动, 开发基于复杂网络理论的供应链传播模型。将总包单位、分供应商、物流节点构建为加权有向图, 引入传播阈值 λ 描述价格波动传导强度。通过 PageRank 算法识别关键路径, 当钢材价格波动超过阈值 $\lambda=0.15$ 时, 系统自动触发备选供应商激活机制, 在雄安新区某 EPC 项目实践中提前 37 天完成采购方案切换。

政策敏感性分析采用文本向量化技术, 将《建设工程工程量清单计价规范》等政策文档转化为 200 维词嵌入向量, 通过余弦相似度计算新政条款与成本科目的关联强度。构建政策响应函数 $f(t)=\sigma(W \cdot x+b)$, 其中 σ 为 LeakyReLU 激活函数, 实现政策调整到成本影响的非线性映射, 在营改增政策模拟中达到 94.7% 的预测吻合度。

2.2 动态特征工程实现路径

传统特征工程难以捕捉施工过程的时序依赖性, 本研究

提出基于 Bi-LSTM 的特征重构方案。将工序持续时间、资源消耗量等时序数据输入双向循环网络,通过细胞状态 C_t 的遗忘门机制筛选关键特征片段。在深圳某超高层项目中,该模型成功提取塔吊使用冲突、底板大体积混凝土连续浇筑等 7 类关键时序特征,使工期-成本耦合预测误差降低至 2.1%。

为解析工序间的空间关联,研发基于图注意力网络 (GAT) 的工序图谱构建技术。将 48 类标准工序映射为异构图节点,边权重由 RBF 核函数计算的空间距离与工艺逻辑共同决定。通过节点嵌入表示工序间的依赖关系,在机电安装阶段成功识别电缆桥架与风管交叉作业这一隐性成本驱动因素,优化后路径成本下降 19.3%。

隐性成本可视化采用力导向布局算法,将材料库存周转率、设备闲置时长等 23 个隐性指标映射为三维空间坐标。通过 D3.js 库实现动态力场模拟,当某维度成本压力超过阈值时,系统自动调整节点间斥力参数。在郑州某场馆项目应用中,该技术使管理人员对临时设施分摊成本的认知偏差从 37% 降至 8.5%,决策响应速度提升 3.2 倍。

3 混合驱动型预警模型构建

3.1 统计学习与深度学习融合架构

针对工程成本时序数据的非平稳特性,构建蒙特卡洛-LSTM 混合预测架构。该模型以 LSTM 网络为核心处理单元,通过记忆细胞状态捕捉工程量清单项的长期依赖关系,同时引入蒙特卡洛树搜索 (MCTS) 增强随机扰动模拟能力。在杭州西站枢纽项目实践中,模型对钢构件价格波动的预测区间覆盖率达 89.6%,较单一 LSTM 模型提升 23.4 个百分点。

注意力机制创新应用于成本偏差溯源,通过自注意力头分配权重矩阵,量化设计变更、市场波动等 12 类因素对总成本的影响系数。在南京某综合管廊项目中,成功定位因地质勘察误差导致的支护方案变更,其贡献度权重被精准计算为 0.37,为责任划分提供量化依据。

模型更新采用滚动时域控制策略,设置滑动窗口长度为 6 个成本周期 (约 180 天)。当检测到实际成本偏离预测值超 15% 时,自动触发贝叶斯优化器调整网络超参数,包括学习率衰减系数和 Dropout 比例。在成都天府国际机场项目中,该机制使模型在疫情导致的材料价格异动期间,仍保持 83.2% 的预测精度。

3.2 多级预警阈值动态设定

建立三维基准预算弹性空间,纵向划分设计概算、施工图预算、执行预算三个基准面,横向设置 $\pm 5%$ 、 $\pm 8%$ 、 $\pm 12%$ 三级预警带。通过历史案例库训练 XGBoost 分类器,自动计算各成本科目的弹性系数。在西安地铁 16 号线实践中,隧道掘进科目的弹性系数被标定为 0.78,当实际成本突破二级预警带时,系统自动推荐盾构机维修保养优化方案。

关键路径节点敏感度分析采用拓扑排序算法,将总进度计划解构为 37 类标准作业单元。通过计算各单元的成本斜率 (Cost Slope) 和自由时差,构建敏感度矩阵。在广州白云机场 T3 项目中,模型提前 45 天预警钢网架吊装工序的成本敏感度超限风险,指导项目部将高空作业平台租赁周期压缩 19 天。

预警处置策略库包含 72 种标准化应对方案,采用产生式规则引擎实现智能匹配。当触发三级预警时,系统自动激活包含商务索赔、设计优化、采购替代的三级响应流程。在郑州某超高层项目中,针对玻璃幕墙成本超支,系统推荐了“甲供材料转乙供+安装费动态计价”的组合方案,成功将成本增幅控制在 2.1% 以内。

4 智能决策支持系统实现

4.1 系统架构与核心模块设计

系统采用微服务架构设计,通过 Kubernetes 容器化部署实现弹性扩展。核心模块包含数字孪生驱动的成本镜像系统、规则引擎与优化引擎协同模块,以及跨平台可视化交互界面。在杭州西站枢纽项目中,数字孪生体通过 Unity3D 引擎实时映射施工现场,集成物联网设备采集的混凝土温度、塔吊载荷等 12 类物联数据,构建三维成本热力图。当某区域热力值突破阈值时,系统自动关联对应 WBS 编码的成本科目,实现物理空间到成本空间的精准映射。

规则引擎采用 Drools 框架实现业务逻辑解耦,内置 200 余条成本管控规则,如“月进度款支付比例不得超过已完工程量 85%”等。优化引擎集成遗传算法与模拟退火算法,在雄安新区某 EPC 项目中,针对钢筋采购方案比选场景,通过并行计算 128 种组合策略,7 分钟内输出“期货锁价+现货补充”的最优解,较传统方案节约采购成本 3.2%。

可视化界面采用 Electron 框架开发,支持 PC 端与移动端同步。创新引入 AR 成本审查模式,通过 Hololens 设备将 BIM 模型与成本数据叠加显示,管理人员通过手势操作即可

调取构件级成本明细。在西安地铁 16 号线应用中,该功能使现场成本核查效率提升 4 倍,问题定位精度达到轴线级。

4.2 决策优化功能实现

资源调度模拟沙盘基于 AnyLogic 平台构建,集成离散事件仿真与系统动力学模型。在广州白云机场 T3 项目钢网架施工阶段,沙盘模拟了 15 种不同劳动力配置方案对成本的影响,通过蒙特卡洛模拟发现:当焊工人数超过阈值时,因工作面拥挤导致的窝工损失将抵消人力增量收益。最终指导项目部采用“三班两运转”弹性排班制,在工期不变前提下降低人工成本 18.7%。

合同条款智能校验系统采用 BERT-BiLSTM-CRF 架构,对施工合同进行语义角色标注。在郑州某超高层项目审计中,成功识别“措施费包干但设计变更超过 $\pm 3\%$ 部分可调”等 17 处隐性风险条款,避免潜在损失 2300 万元。系统支持条款追溯分析,通过知识图谱展示条款间的逻辑关联,辅助商务经理制定谈判策略。

供应链金融风险对冲模块构建于 Hyperledger Fabric 区块链平台,将应付账款、保理合同等上链存证。通过智能合约自动匹配核心企业信用额度与分供应商融资需求,在成都天府国际机场项目中,实现 T+0 日保理放款,较传统模式缩短 15 个工作日,财务成本降低 1.2 个百分点。系统内置的 VaR 模型可动态计算融资杠杆率,当市场利率波动超过 50BP 时,自动触发对冲头寸调整指令。

5 工程场景验证与优化

5.1 典型项目应用实证分析

在青岛地铁 4 号线隧道工程中,系统通过地质数据融合技术,将超前地质预报的 RQD 值与成本科目建立映射关系。当实际掘进遇到与勘察报告偏差达 32% 的破碎带时,模型提前 47 天触发预警,动态调整支护参数和材料消耗基准,最终将该区段成本超支控制在 3.8%,较传统管理模式降低 19.6 个百分点。系统内置的蒙特卡洛模拟引擎对 12 种应急方案进行量化评估,推荐采用的“管棚+小导管”组合支护方案,经实践验证节约工期 21 天。

针对深圳平安金融中心幕墙系统变更,系统通过参数化建模技术建立构件级成本镜像。当业主提出开启扇比例调整需求时,模型 72 小时内完成 386 种组合方案的比选分析,量化显示每增加 1% 开启扇将导致型材损耗率上升 0.87 个

百分点。最终通过优化龙骨截面参数,在满足建筑功能要求的前提下,将幕墙单方成本控制在预算值的 102.3%。

在雄安新区启动区市政道路项目中,系统成功应对装配式构件价格政策调整。通过 NLP 技术解析《河北省装配式建筑计价依据》新规条款,自动生成包括构件运输半径、吊装费用等 6 项成本调整系数。当政策实施导致 PC 构件价格上浮 12% 时,系统推荐采用“本地化生产+临时存贮”策略,经模拟验证该方案较跨区域调运节省成本 8.9%。

5.2 系统迭代优化路径

针对跨项目数据协同难题,研发基于联邦学习的分布式训练框架。在保护数据隐私前提下,实现不同企业成本数据库的模型参数共享。经测试,在参与方超过 15 个时,模型收敛速度较传统集中式训练提升 41%,且成本预测精度保持 92.3% 以上。该框架已通过中国信通院联邦学习安全认证,满足工程数据不出域的合规要求。

数字孪生体实时纠偏机制采用卡尔曼滤波算法,将施工机械传感器数据与 BIM 模型进行状态融合。在成都某数据中心项目中,该机制使塔吊作业计划与混凝土浇筑需求的匹配度从 68% 提升至 91%,设备闲置率下降至 3.2%。创新引入数字线程技术,实现设计变更到成本影响的秒级传导,在深圳某医院项目中成功拦截因管线综合调整引发的 370 万元潜在损失。

案例库动态更新策略构建强化学习循环,通过用户反馈调整案例推荐权重。系统内置 DQN 网络,以决策采纳率作为奖励函数,持续优化知识推送逻辑。在广州某综合体项目中,该机制使类似工程案例的匹配准确度从 59% 提升至 84%,辅助生成包括“铝模替代木模的经济性分析”等 17 项创新优化建议,其中 9 项被纳入企业技术标准。系统支持 OCR 识别技术自动捕获竣工结算文档中的创新点,形成可复用的知识资产。

参考文献:

- [1] 刘蕾. 房地产企业财务共享中心建设存在的问题与对策研究[J]. 质量与市场, 2023(13):67-69.
- [2] 黄燕. 房地产企业成本核算问题与对策分析[J]. 中国产经, 2023(22):77-79.

作者简介: 陈思琦(2005—),男,汉族,江西省抚州市,本科生,研究方向为物联网系统。