



XXXX

大语言模型辅助的室内多源定位智能调度与连续性控制方法

李成

(中国电信集团有限公司本部、北京、邮编: 100033) 牛锐(中电信无人科技(江苏)有限公司、南京、邮编: 210019) 邵震(中国电信股份有限公司上海研究院、上海、邮编: 201315)

摘要: 针对大型复杂室内场景多技术协同难、跨场景切换滞后、定位精度不稳定和服务连续性不足等问题,参考3GPP(3rd Generation Partnership Project,第三代合作伙伴计划)定位框架并进行平台级扩展,提出大语言模型辅助的室内多源定位智能调度与连续性控制方法。方法以蜂窝、Wi-Fi等测量信息为主,并可扩展适配蓝牙/UWB(Ultra Wide Band,超宽带)和惯性传感等多源数据,通过时空对齐、信号语义编码、场景识别、约束化调度和融合解算实现动态源选择。基于40万平方米商业综合体8个楼层实测数据验证,结果表明:场景识别加权准确率达到94.3%,开阔无遮挡LOS(Line-of-Sight,视距)中庭场景定位均方根误差为0.75 m,跨场景定位连续性达到88%,平均定位误差较固定权重融合方法降低49.5%。本方法可提升复杂室内场景下定位服务的连续性、稳定性和跨场景适配能力。

关键词: 室内定位; 多源感知; 大语言模型; 场景识别; 连续性控制; 5G-Advanced(5G-A)

中图分类号: TN929.5

文献标志码: A

doi: 10.11959/j.issn.1000-0801.

Large-language-model-assisted intelligent scheduling and continuity control for indoor multi-source positioning

Li Cheng(China Telecommunications Corporation, Beijing, Postal Code: 100033) Niu Rui

(China Telecom Unmanned Technology (Jiangsu) Co., Ltd., Nanjing, Postal Code: 210019)

Shaozhen(China Telecom Co., Ltd. Shanghai Research Institute, Shanghai, Postal Code:

201315)

Abstract: To address the difficulties of multi-source positioning coordination, delayed cross-scenario handover, and insufficient service continuity in complex indoor environments such as large commercial complexes, a large language model assisted scheduling and continuity control method for indoor multi-source positioning was proposed inspired by the 3GPP positioning framework and extended at the platform level. Cellular and Wi-Fi measurements were primarily aggregated on the location service platform, with extensible adaptation to Bluetooth/UWB and inertial sensing

收稿日期: 2026-08-22; 修回日期: 2026-09-14

基金项目: 无线通感一体化可靠感知理论及关键技术研究(U25B2040)

Foundation Items: Research on Reliable Sensing Theory and Key Technologies for Wireless Integrated Sensing and Communication(U25B2040)



data. Dynamic source selection was achieved through spatio-temporal alignment, signal semantic encoding, scenario recognition, constrained scheduling, and fusion calculation. The method was evaluated with field data collected from eight floors of a 400 000 m² commercial complex. The results showed that the weighted scenario recognition accuracy reached 94.3%, the root mean square error in open LOS atrium scenarios was 0.75 m, the cross-scenario positioning continuity reached 88%, and the average positioning error was reduced by 49.5% compared with the fixed-weight fusion method. The proposed method improves the continuity, stability, and cross-scenario adaptability of indoor positioning services.

Key words: indoor positioning, multi-source perception, large language model, scenario recognition, continuity control, 5G-advanced

1 引言

商业综合体、交通枢纽、大型医院等复杂室内场景，建筑遮挡复杂、人员密集、无线多径效应突出，是高精度连续定位的难点场景^[1-3]。当前5G定位依托3GPP标准^[4-7]，支持E-CID（Enhanced Cell-ID，增强型小区识别码）、UT-DOA（Uplink Time Difference of Arrival，上行到达时间差）、Multi-RTT（Multiple Round-Trip Time，多往返时间）等多种主流定位方案；行业同步配套Wi-Fi RTT（Round-Trip Time，往返时延）、指纹定位等非3GPP短距定位技术，可融合蓝牙、UWB及终端惯导数据实现协同定位。由于各类定位技术在精度、时延、功耗、部署难度上各有优劣，单一方案无法适配动态复杂的室内环境，必须通过多技术融合调度与联合解算，保障定位服务稳定连续。

在智能算法层面，传统机器学习与深度模型可实现定位特征提取与建模^[8-9]，但存在明显短板，不仅依赖人工特征工程与场景专属数据，跨场景易出现数据偏移，还存在可解释性弱、小样本迁移能力不足、落地难度大等问题。当前大模型^[10-11]、链式推理^[12]、高效微调技术^[13-14]为定位技术升级提供了新路径，可实现场景理解、策略优化与异常纠错，无需替代传统定位解算模块，能够低成本完成领域适配。目前大模型赋能多源室内融合定位仍处于探索阶段，缺乏成熟可落地

的一体化方案，本文聚焦复杂室内5G+Wi-Fi融合定位痛点，结合轻量化大模型技术，构建全新智能融合定位体系，弥补现有技术短板。

2 室内跨场景智能定位框架

本文以商场混合覆盖场景为例，同时部署5G定位平台、Wi-Fi定位平台和融合定位管理平台，其中5G定位平台包括LMF（Location Management Function，位置管理功能）与GMLC（Gateway Mobile Location Center，网关移动位置中心）；而网络分成5G基站主覆盖区、Wi-Fi主覆盖区及两种信号重叠覆盖区三种区域。终端沿行进轨迹移动时，融合定位管理平台综合实时测量信号、运动轨迹、覆盖网格、平台能力和终端硬件能力等特征，由LLM（Large Language Model，大语言模型）辅助决策模块生成保持、预切换或回切建议，并经规则约束后下发执行，从而保障定位服务连续性。整个系统分为定位功能层与大模型中枢调度层，工作机理如图1所示：

定位功能层，终端会持续扫描蜂窝基站、Wi-Fi AP（Access Point，无线接入点）信号分别上报至5G定位平台、Wi-Fi定位平台；两套定位平台独立运行对应定位算法（基站侧E-CID/UT-DOA/Multi-RTT、Wi-Fi侧RTT/指纹），并实时向融合定位管理平台同步原始测量数据、当前定位结果。

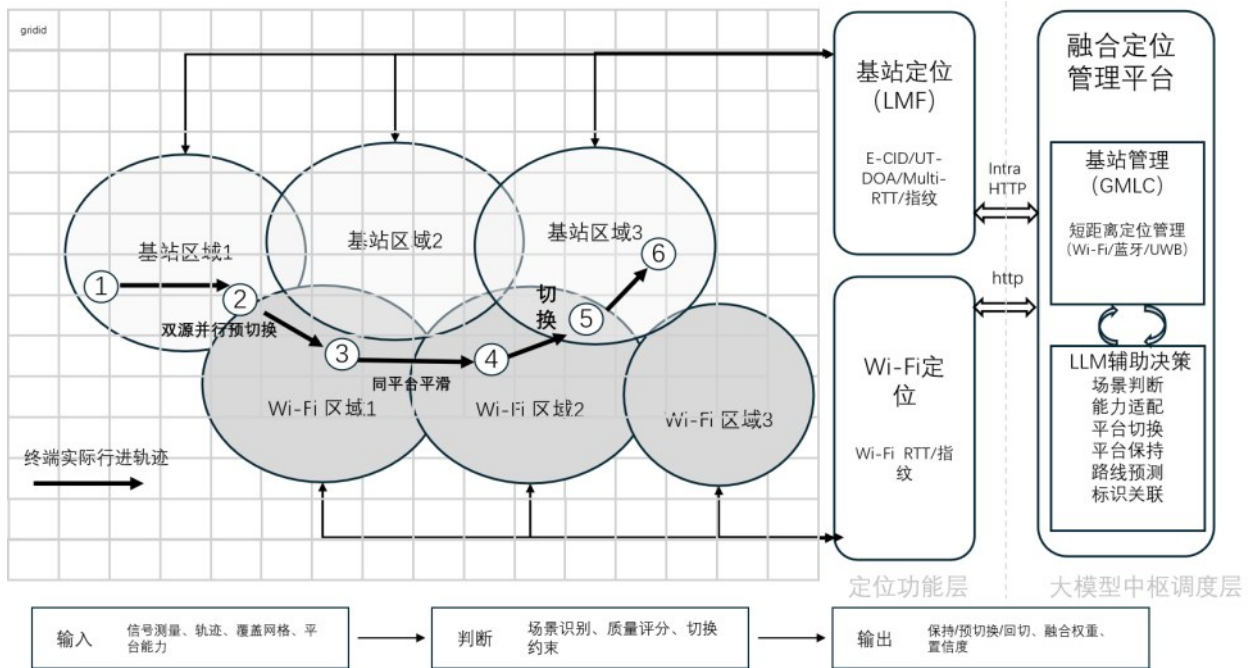


图1 融合定位跨平台切换全流程图

大模型中枢调度层，调用专项训练的 LLM 辅助决策模块，输入了终端实时信号测量向量、终端运动轨迹（速度/方向）、各平台定位精度/覆盖能力、场内网格化区域边界、终端硬件适配能力等，进行处理后输出“持续服务、跨平台切换、同区域保持、预调度待机”四类调度指令。同时，根据终端行进轨迹提前预判相邻覆盖区域，能实现切换预加载、标识预关联，解决跨 5G 与 Wi-Fi 的定位中断、坐标跳变等问题，完成

定位的全域连续。

2.1 融合定位跨平台切换全流程

以图1为例来说明整个工作流程，图中节点1~节点6是融合定位跨平台切换全流程涵盖6个典型节点，各节点的区域特征、数据输入、区域路径预测与调度动作如表1所示：

2.2 LLM辅助决策的处理架构

为实现 5G 定位与 Wi-Fi 定位之间的智能平滑切换，引入的 LLM 辅助决策是基于时序信号趋

表1 融合定位跨平台切换全流程中的节点情况说明

	节点1	节点2	节点3	节点4	节点5	节点6
区域特征	基站区域1内（仅5G信号）	基站区域1与Wi-Fi区域1重叠交界	Wi-Fi区域1核心覆盖区	Wi-Fi区域1与Wi-Fi区域2重叠交界	Wi-Fi区域2与基站区域3重叠交界	基站区域3核心覆盖区
数据输入	仅5G测量数据+终端运动状态	蜂窝基站信号+Wi-Fi区域1 AP信号+轨迹运动数据	Wi-Fi区域1信号为主残留微弱基站杂波	Wi-Fi区域1、2两类AP信号	Wi-Fi区域2 AP信号+基站区域3 5G信号+轨迹方向数据	基站区域3蜂窝信号为主Wi-Fi信号衰减至阈值以下
区域路径预测	下一步途经基站区域2、Wi-Fi区域1	快速离开基站区域1长期驻留Wi-Fi区域1后续途经基站区域2、Wi-Fi区域2	持续在Wi-Fi区域1行进下一步抵达Wi-Fi区域2、基站区域2交界	完全驶入Wi-Fi区域2周边含基站区域2、Wi-Fi区域3	即将脱离Wi-Fi主覆盖长期进入基站区域3蜂窝覆盖区	稳定驻留基站区域3周边可覆盖基站区域2、Wi-Fi区域3
调度动作	维持5G定位等待Wi-Fi信号触发	执行跨平台切换：向Wi-Fi平台发起准入短时并行双平台平滑过渡	Wi-Fi定位预调度基站平台待机监听提前同步终端标识	同类型区域保持服务：跨Wi-Fi小区持续服务AP间信号平滑融合	二次跨平台切换：向5G基站平台发起准入短时双平台融合消除跳点	维持蜂窝定位循环采集与决策闭环全程保障无间断定位



势分析、机器学习规则提取与大语言模型多维语义推理的协同机制，涉及了信号连续监测、时序趋势分析、ML规则提取、LLM多维语义推理与切换执行评估五个递进阶段，各阶段相应的工作内容如图2所示：

切换决策以时序趋势分析为核心驱动力，实现了从信号监测到切换执行的全链路智能化。然而，切换决策的准确性高度依赖于对当前所处室内场景的准确识别。不同室内场景具有显著差异化的信号传播特性与定位源可用性，若场景识别出现偏差，将直接导致切换时机误判与定位服务中断。因此，需要进一步构建精细化室内定位场景体系，将清洗处理后的测量数据进行场景化分类识别，提供给LLM进行实时推理决策。以下为执行实例：

ü 系统提示词：

你是一个智能定位调度助手，专门负责室内多源定位系统。你的任务是分析当前的信号环境和终端移动状态，然后输出最优的定位调度决策。请依次评估各信号源质量、判断当前场景类型并给出调度决策。仅输出简明、可审计的决策依据，不输出完整内部推理过程。

ü 用户提示词：

当前信号测量数据（5分钟滑动窗口，每个窗口60秒）

| 窗口 | 时间戳 | 5G RSRP (dBm) | 5G RSRP 趋势 | Wi-Fi RSSI (dBm) | Wi-Fi 趋势 | 可见小区数 | 可见 AP 数 |

| --- | ----- | ----- | ----- | ----- | ----- | ----- |

| T-4 | {ts0} | {rsrp0} | {trend0} | {rssi0} | {wtrend0} | {cell0} | {ap0} |

| ... | ... | ... | ... | ... | ... | ... |

| T0 | {ts4} | {rsrp4} | {trend4} | {rssi4} | {wtrend4} | {cell4} | {ap4} |

场景识别结果

- 当前场景：{scenario_label}（置信度：{confidence}）

- 场景类型：{LOS/NLOS 分类}（*注：LOS 指视距，NLOS 指非视距*）

终端移动状态

- 预估速度：{speed} m/s

- 移动方向：{direction}

- 当前网格区域：{region_id}

平台能力约束

- 5G 定位可用性：{5g_available}（精度：{5g_acc} 米）

- Wi-Fi 定位可用性：{wifi_available}（精度：{wifi_acc} 米）

- 切换开销（耗时）：{handover_cost} 毫秒

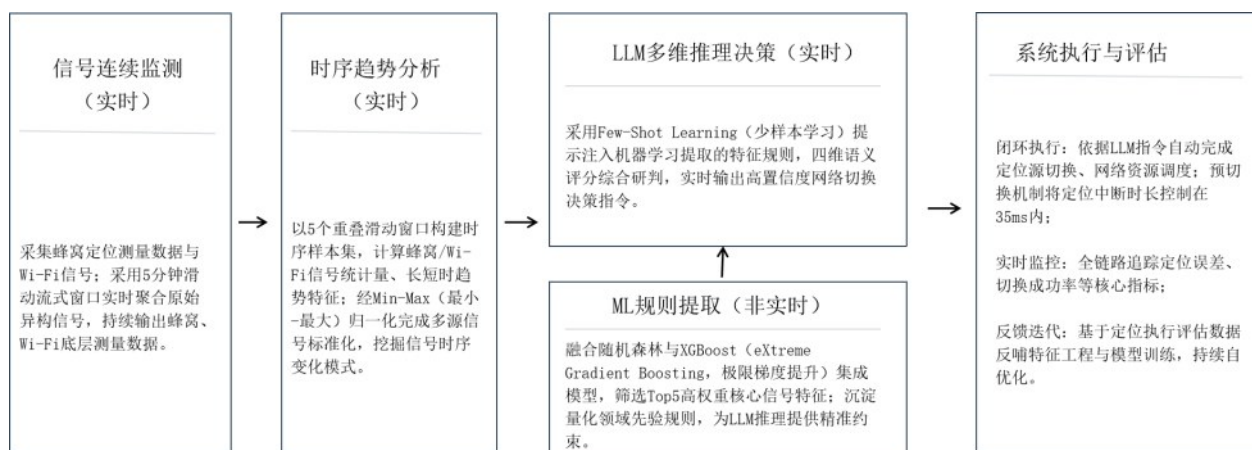


图2 处理系统总体流程架构

任务

基于以上信息，确定最优的调度动作。请以 JSON 格式输出。

时序信号趋势分析采用 3 窗/5 窗滑动线性回归计算 RSRP 和 Wi-Fi RSSI (Received Signal Strength Indication, 接收信号强度指示) 的变化率 (dB/分钟), 作为"5G RSRP 趋势"和"Wi-Fi 趋势"字段填入用户提示词; ML 规则提取基于历史决策数据训练的轻量级随机森林模型, 输出 Top-3 候选调度策略及其置信度, 作为 LLM 推理的参考输入。

2.3 测量数据的场景化分类识别

本阶段工作是以标准化的场景分类知识库为基础, 并结合信号覆盖密度、无线传播特性、空间结构属性三个工程维度, 去构建覆盖商业综合体全工况的精细化室内定位场景体系。

(一) 场景划分核心维度与量化阈值的定义

本研究场景判定标准由目标商业综合体实测数据量化得到, 把室内商场特征主要场景划分 12 类全覆盖室内定位场景作为 LLM 智能决策的标准化知识库。

实测数据集统计特征是采用分位数分析与工程经验相结合的方法确定场景量化判定阈值。具体而言, 以实测数据中基站/AP 数量的 25% 分位数与 75% 分位数作为稀疏/密集划分的统计边界, 以 RSRP (Reference Signal Receiving Power, 参考信号接收功率) 标准差的 IQR (Interquartile Range, 四分位距) 作为传播环境分类依据, 经

交叉验证后确定最终阈值。各维度量化判定标准如下:

(二) 精细化场景分类与 LLM 调度策略的匹配

基于上述指标, 先由信号覆盖密度与传播环境两个量化维度形成 12 类信号场景, 再以空间结构属性作为语义标签和辅助约束。所划分的 12 类精细化场景是在实测数据中空间标注识别的 5 类基础功能区 (商场入口、中央广场、店铺通道、配电房、电梯厅) 基础上, 再结合信号覆盖密度与传播环境维度进行细粒度交叉划分得到的。5 类功能区对应粗粒度空间语义标签, 12 类场景则对应细粒度信号-空间联合决策单元, 二者形成"功能区→场景类→调度策略"的层级映射关系。每类场景绑定专属 LLM 调度规则、最优定位平台与冗余备用策略, 全面覆盖室内常规工况与极端复杂工况, 具体场景参数与调度方案如表 3 所示。此外, 针对相邻场景边界处的模糊切换问题, 采用了滑动窗口置信度平滑机制。当连续 3 个观测窗口的场景分类结果出现跨类跳变时, 触发 LLM 二次确认推理, 结合终端移动轨迹方向与速度信息判定是否为真实场景切换, 有效抑制场景边界的乒乓切换现象。本次试验选取了一个基于 40 万平方米 8 个楼层的商业综合体, 其涵盖了 12 种场景。

(三) 采集数据的精细化场景分类识别

同步采集到的 5G、Wi-Fi 数据在进行时空对齐后, 清洗异常值和计算多维度统计融合特征, 通过定制模板将数值信号转为带领域先验的自然

表 2 室内无线场景分类判定维度及阈值对照表

核心维度	阈值要求
信号覆盖密度	1. 蜂窝基站判定: 实测可检测 5G/LTE 基站数量 25% 分位数=2、75% 分位数=4; 基站 ≥ 3 为蜂窝密集场景, ≤ 2 为蜂窝稀疏场景 2. Wi-Fi AP 判定: 实测可扫描有效 AP 数量 25% 分位数=2、75% 分位数=5; 有效 AP ≥ 3 为 Wi-Fi 密集场景, ≤ 2 为 Wi-Fi 稀疏场景
无线传播特性	以 RSRP 标准差划分三类传播场景: 1. 开阔 LOS 场景: 标准差 < 8 dBm, 信号稳定、无明显多径干扰; 2. 半 NLOS (Non Line-of-Sight, 非视距) 遮挡场景: 标准差 8~15 dBm, 轻质隔断/货架造成局部遮挡; 3. 重度 NLOS 遮挡场景: 标准差 > 15 dBm, 承重墙/金属结构/电梯井带来强遮挡与严重多径干扰
空间结构属性	依托建筑 CAD (Computer-Aided Design, 计算机辅助设计) 图纸+POI (Point of Interest, 兴趣点) 语义标注, 划分 4 类标准室内空间形态, 作为场景分类空间约束: 商场开阔公共区、连片商铺区、密闭功能区、狭长通道死角



表3 面向LLM智能调度的精细化室内定位场景分类表

场景编号	场景名称	核心量化判定条件	商场典型对应区域	LLM最优调度定位平台	冗余备用策略
S1	蜂窝+Wi-Fi双密集开阔LOS场景	可测蜂窝基站 ≥ 3 ；可扫Wi-Fi AP ≥ 3 ；空间开阔无硬质遮挡，多节点LOS视距，信号稳定	商场中庭、首层主广场、大型展厅、影院大厅	Wi-Fi RTT多边测距、5G UT-DoA融合定位	双源RSSI指纹融合定位
S2	蜂窝密集-Wi-Fi稀疏开阔LOS场景	可测蜂窝基站 ≥ 3 ；可扫Wi-Fi AP ≤ 2 ；开阔视距环境，蜂窝信号质量较好	商场临街落地窗廊道、室外衔接通道、沿街开阔商铺	5G UTDaA定位平台	蜂窝RSSI指纹定位
S3	Wi-Fi密集-蜂窝稀疏开阔LOS场景	可测蜂窝基站 ≤ 2 ；可扫Wi-Fi AP ≥ 3 ；室内开阔无遮挡，Wi-Fi信号覆盖充足	室内步行街、美食广场大厅、商超主通道、室内休闲区	Wi-Fi RTT高精度定位、Wi-Fi指纹定位	IMU（Inertial Measurement Unit，惯性测量单元）惯性融合定位
S4	双密集半遮挡NLOS场景	可测蜂窝基站 ≥ 3 ；可扫Wi-Fi AP ≥ 3 ；轻质隔断/矮货架半遮挡，信号小幅波动	中庭周边商铺门口、文创轻隔断街区、连锁品牌外廊	RTT+指纹融合定位、蜂窝/Wi-Fi混合定位	动态权重自适应融合定位
S5	蜂窝密集-Wi-Fi稀疏半遮挡场景	可测蜂窝基站 ≥ 3 ；可扫Wi-Fi AP ≤ 2 ；玻璃/轻质隔断局部遮挡	临街玻璃隔断商铺、走廊外侧精品店铺	5G UTDaA蜂窝定位	蜂窝小区加权定位
S6	Wi-Fi密集-蜂窝稀疏半遮挡场景	可测蜂窝基站 ≤ 2 ；可扫Wi-Fi AP ≥ 3 ；布艺/货架轻度遮挡	服装区主走廊、书店阅览区、商超外围通道	Wi-Fi RTT+指纹补偿定位	蓝牙AOA（Angle of Arrival，到达角）辅助定位
S7	双密集重度遮挡NLOS场景	可测蜂窝基站 ≥ 3 ；可扫Wi-Fi AP ≥ 3 ；承重墙/金属结构重度遮挡，多径干扰显著	独立密闭商铺、电梯厅、墙体分隔功能区、设备间外围	蜂窝+Wi-Fi双源指纹融合定位	IMU惯性融合定位
S8	蜂窝密集-Wi-Fi稀疏重度遮挡场景	可测蜂窝基站 ≥ 3 ；可扫Wi-Fi AP ≤ 2 ；外墙厚墙体完全遮挡Wi-Fi信号	临街密闭商铺、后厨操作间、外墙内侧封闭区域	蜂窝高精度指纹定位	蜂窝E-CID冗余定位
S9	Wi-Fi密集-蜂窝稀疏重度遮挡场景	可测蜂窝基站 ≤ 2 ；可扫Wi-Fi AP ≥ 3 ；密闭空间金属屏蔽、重度遮挡	影院放映厅、服装试衣间、超市仓储区、地下车库隔间	离线Wi-Fi指纹融合定位	IMU+气压高度融合定位
S10	双稀疏开阔LOS场景	可测蜂窝基站 ≤ 2 ；可扫Wi-Fi AP ≤ 2 ；空间开阔无遮挡，基准节点稀缺	超长消防通道、空置展厅、顶层空旷备用区域	近邻AP/小区区域粗定位	IMU轨迹推演定位
S11	双稀疏半遮挡NLOS场景	可测蜂窝基站 ≤ 2 ；可扫Wi-Fi AP ≤ 2 ；局部轻微遮挡，信号不稳定	走廊尽头小型商铺、楼梯间休息区、建筑边角区域	小区/AP区域识别定位	历史轨迹约束定位
S12	双稀疏重度密闭NLOS场景	可测蜂窝基站 ≤ 2 ；可扫Wi-Fi AP ≤ 2 ；完全密闭遮挡，信号间歇性丢失	设备机房、弱电间、卫生间隔间、车库偏远死角	惯性轨迹模糊定位、区域分区定位	时序平滑补偿定位

语义文本，并结合场景阈值规则构造少样本提示送入大模型，经链式推理完成三层维度联合场景判定，低置信度时采用多窗口投票优化分类稳定性，最终实现基于采集数据的精细化场景分类的实时识别。

（四）多平台协同的预判式切换机制引入

融合定位跨平台切换全流程需要将场景识别结果与信号切换决策有机融合，形成端到端的智能定位解算体系。而5G与非3GPP定位平台相互独立，存在服务接续不足的问题，设计了基于异

构标识关联的多平台协同机制。为降低敏感信息暴露风险，系统不直接在策略层使用IMSI（International Mobile Subscriber Identity，国际移动用户识别码）或明文MAC（Media Access Control，媒体访问控制）地址，而采用脱敏用户会话标识、Wi-Fi MAC哈希值等匿名标识构建动态映射关系。同时，LMF或定位平台维护区域级定位能力表，记录各区域可用定位体制、覆盖范围、精度能力和切换代价，为策略调度提供约束条件。系统采用预判式切换机制，由LLM辅助决策模

块根据场景变化趋势、信号质量劣化状态和平台能力约束生成预切换建议，再由规则校验模块确认是否执行，以减少被动切换导致的定位中断和轨迹跳变。具体而言，LLM 辅助决策模块基于终端当前运动轨迹的速度与方向，预判终端即将进入的相邻覆盖区域，并提前分析该区域的网络能力与终端匹配度，在信号质量劣化前完成平台准入与标识预关联，避免切换窗口期内因能力不匹配导致切换失败。其中，轨迹预判采用线性外推法，基于终端当前速度和方向，预测终端在 $T+\Delta t$ 时刻 ($\Delta t=30$ 秒，即一个切换准备周期) 的预计位置，结合区域网格地图确定预到达区域。能力预匹配查询 LMF 维护的区域级定位能力表，获取预到达区域的可用定位体制和覆盖能力，提前完成平台准入认证和标识预关联。

3 大模型轻量化微调与自适应融合算法

3.1 基座大模型特性

基座大模型在通用自然语言理解与生成任务上表现较好，但在信号定位与切换决策这一垂直领域存在显著的知识盲区。具体表现为：

· 信号传播物理规律缺失：通用 LLM 缺乏对 RSRP 空间衰减特性、NLOS 遮挡导致的信号波动、多径效应等无线传播物理机制的深度理解，难以区分信号暂时抖动与持续性恶化。

· 场景指纹语义认知不足：商场入口、中央广场、店铺通道、配电房、电梯厅等室内场景具有独特的信号指纹特征（如电梯厅物理小区数量 >10 且 RSRP 标准差 >10 ），通用 LLM 无法从数值信号中自动归纳这些场景规则。

· 时序趋势语义理解盲区：信号切换决策本质上是 5 个基于时序连续观测滑动窗口（以 1 分钟为窗口聚合，共 5 分钟观测时长）的时序趋势分析问题。通用 LLM 虽具备一定的数值比较能力，但在 RSRP 三窗趋势 $<-8\text{dB}$ 且 Wi-Fi RSSI 趋势 $>+3\text{dB}$ 且长期累积衰减 $>15\text{dB}$ 这样的多维联合推理上表现不稳定。

3.2 LoRA 轻量微调实验

一般通用大模型缺乏室内定位领域专业知识，直接零样本推理存在精度低、稳定性差等问题，需要通过微调将信号物理规律和场景先验知识注入模型参数，实现了模型行为与业务需求的对齐，确保模型在推理过程中严格遵循多维评分 $\rightarrow$ 综合决策 $\rightarrow$ 结构化输出的规范流程，弥补通用 LLM 与特定信号定位任务之间的领域鸿沟。但若采用全量微调，又会存在参数量大、算力成本高、部署困难等缺陷，本文采用 LoRA (Low-Rank Adaptation, 低秩适配) 方案来解决这些问题。

3.2.1 场景切换实验

信号切换决策是一个时序趋势分析+多维信号联合评估问题。传统方法依赖固定阈值规则（例如 $-100\text{ dBm} < \text{RSRP} < -70\text{ dBm}$ ），但固定阈值无法适应不同场景的信号特征差异。本方案将 LLM 作为切换决策引擎，利用微调数据及提示模板中注入的信号传播先验规则，对时序信号特征进行语义化联合判断，在蜂窝和 Wi-Fi 信号的“交叉窗口”做出最优切换决策。LLM 在决策时，会综合考量表 4 中的 4 个核心维度，确保决策具备明确的物理可解释性：

表 4 模型推理决策分析维度表

维度	推理依据	物理含义
信号强度趋势	蜂窝 RSRP 的 3 窗/5 窗滑动趋势	终端远离基站 $\rightarrow$ 信号衰减，是切换的核心驱动力
Wi-Fi 可靠性	RSSI 绝对值、AP 密度、Wi-Fi 趋势	Wi-Fi 定位精度是否满足服务等级要求
长期衰减	5 窗累积 RSRP 变化 (RSRP)	区分临时波动和持续衰减，避免过早切换
切换时机	蜂窝衰减+Wi-Fi 增强的交汇点	最优切换点：蜂窝已不可靠+Wi-Fi 已成熟



表5呈现了一个典型的连续时间窗口推理过程，体现了LLM决策逻辑，其中输入序列数据为接收连续5个时间窗口的数据（每窗口60秒）：本文将线性回归斜率定义为信号变化率，负值表示衰减，正值表示增强。

以“RSRP变化率=-7.6 dB/min（持续衰减）；Wi-Fi RSSI=-56.1 dBm；可见AP数=5；5窗累计RSRP衰减=17.6 dB；切换必要性=高；决策：跨平台切换”作为决策依据。表5整体呈现蜂窝信号持续衰减、Wi-Fi信号持续增强的趋势，LLM结合平台能力约束将T4识别为候选切换点，在规则校验通过后执行切换。

使用多个方法进行了微调实验，对比结果显示：本方案采用的“Qwen2.5-7B+LoRA”方法，在切换决策任务上主要提升体现在F1（0.8400）和Recall（0.8680），对需要切换样本的识别能力更强，适合用于复杂场景下的辅助调度：

为了进一步验证LLM各组件对定位性能的贡献，于是采用消融实验来进行验证，其中关于

实验参数如下：模型版本Qwen2.5-7B-Instruct，LoRA秩 $r=16$ ，学习率 $2e-4$ ，训练3轮，4bit量化，随机种子42，温度0.1，结构化输出采用JSON Schema校验，推理失败时回退为规则策略。JSON Schema输出样例如下：

```
{
  "$schema": "http://json-schema.org/draft-07/
schema#",
  "type": "object",
  "properties": {
    "decision_basis": {
      "type": "string",
      "description": "Concise and auditable decision
basis"
    },
    "decision": {
      "type": "string",
      "enum": ["maintain", "cross_platform_hando-
ver", "same_region_keep", "pre_schedule_standby"],
```

表5 典型的连续时间窗口推理过程表

窗口	RSRP均值	RSRP变化率/(dB/min)	Wi-Fi RSSI	Wi-Fi变化率/(dB/min)	小区数	AP数
T0	-85.2 dBm	+0.5	-72.1 dBm	+1.2	8	1
T1	-89.1 dBm	-2.2	-67.6 dBm	+2.1	7	2
T2	-93.5 dBm	-3.8	-64.8 dBm	+3.5	6	3
T3	-96.3 dBm	-4.9	-60.2 dBm	+2.3	5	4
T4	-102.8 dBm	-7.6	-56.1 dBm	+4.8	4	5

表6 LoRA微调实验结果对比表

方法	Accuracy	Precision	Recall	F1-Score	规则一致性评分
Qwen2.5-7B+LoRA（本方案）	0.8550	0.8140	0.8680	0.8400	4.3
Random Forest	0.7882	0.6667	0.5272	0.5888	2.8
XGBoost	0.7905	0.6266	0.6720	0.6485	2.9
Qwen2.5-7B(Zero-shot)	0.5620	0.4820	0.7520	0.5874	3.5
LLaMA-3-8B(Zero-shot)	0.5710	0.4930	0.6980	0.5778	3.4
GPT-4o-mini(Zero-shot)	0.6750	0.6061	0.8620	0.7120	4.2
LLaMA-3-8B+LoRA	0.7880	0.7450	0.8030	0.7730	3.9
Mistral-7B+LoRA	0.8020	0.7580	0.8260	0.7810	4.0

各方法采用相同的数据划分、Prompt框架与结构化输出约束。表6中的Qwen2.5-7B(Zero-shot)对应仅使用数值键值对、未注入场景语义的配置；LoRA组在相同测试集上进行评估。

```
"description": "Scheduling action: maintain current service, handover across platforms, keep within same region, or pre-schedule standby"
```

```
},
```

```
"target_platform": {
```

```
"type": "string",
```

```
"enum": ["5G", "Wi-Fi", "hybrid", "none"],
```

```
"description": "Target positioning platform after decision"
```

```
},
```

```
"confidence": {
```

```
"type": "number",
```

```
"minimum": 0.0,
```

```
"maximum": 1.0,
```

```
"description": "Confidence score of the decision"
```

```
},
```

```
"handover_timing": {
```

```
"type": "string",
```

```
"enum": ["immediate", "delayed", "prepared"],
```

```
"description": "Timing of handover execution"
```

```
}
```

```
},
```

```
"required": ["decision", "target_platform", "confidence", "decision_basis"]
```

```
}
```

实验结果如表 7 所示:

由表 7 可知: (1) 对比 A 与 B, 传统 ML 较固

定阈值规则的准确率提升 6.48%, F1 提升 0.32%。

(2) 对比 C 与 D, 语义化文本编码使准确率提升 11.30%、F1 提升 7.76%。(3) 对比 D 与 E, LoRA 微调使准确率提升 14.80%、F1 提升 13.60%, 切换时延降低 30 ms。(4) 对比 E 与 F, 规则约束使连续性提升 2.8%。完整方案 F 相较基线 A, 准确率相对提升 13.1%, 连续性提升 11.5%, 说明文本化编码、领域微调和规则校验均带来可测增益。

同时, 无文本化实验实际上仍将数值转换为最简文本键值对形式, 将数值信号以简洁的键值对文本形式输入 (如 RSRP=-85.2, RSSI=-72.1, Cells=8, aps=1), 与有文本化实验的差异在于是否包含场景语义描述、物理规律解释和结构化决策依据引导。而 C 组与 D 组的性能差异主要归因于语义信息注入, 两组实验均使用相同的数值格式 (保留两位小数), 有文本化组仅在数值键值对基础上附加场景语义描述和物理规律提示, 两组的数值 token 分布完全一致。

继续采用融合权重计算方法、连续性指标来统计切换时延与定位误差, 去验证 LLM 辅助调度对定位精度和连续性的提升效果:

由表 8 可知, LOS 场景下预切换机制可实现零中断切换; 重度遮挡场景因信号恶化迅速, 预切换提前量不足导致平均中断 35ms, 表 7 中的 280 ms 为固定阈值方法的切换时延, 与中断时长口径不同, 二者不作直接比较。为进一步细化不同传播环境下的定位性能, 按开阔无遮挡 LOS 中

表 7 消融实验结果对比

实验组	配置	精度	F1	切换时延/ms	连续性/%
A	固定阈值规则	0.7234	0.5856	280	76.5
B	Random Forest	0.7882	0.5888	220	82.1
C	Qwen2.5-7B Zero-shot (无文本化)	0.5620	0.5874	190	79.3
D	Qwen2.5-7B Zero-shot (有文本化)	0.6750	0.6650	185	83.5
E	Qwen2.5-7B+LoRA (无规则约束)	0.8230	0.8010	155	85.2
F	完整方案	0.8550	0.8400	120	88.0



表8 不同场景类型切换时延统计

场景类型	切换次数	平均时延/ ms	P95时延/ ms	平均中断 时长/ms
LOS 开阔	45	120	180	0
半遮挡	38	156	230	12
重度遮挡	22	210	320	35
总体	105	152	280	12

庭、单层隔断轻度NLOS 商铺/走廊和多层遮挡/设备稀疏边角区三类场景分别统计定位误差,结果如表9所示:

本文采用加权最小二乘融合: $\hat{X} = (H^T W H)^{-1} H^T W Z$, 其中, $\hat{X}$ 为融合定位坐标, Z 为各定位源观测向量, H 为观测矩阵, $W = \text{diag}(w_1, w_2, \dots, w_n)$ 为权重矩阵。归一化初始权重为 $W_i = (1/\sigma_i^2) / \sum_j (1/\sigma_j^2)$, 满足 $\sum_i W_i = 1$ 。

动态权重更新综合RSSI方差、SNR以及多径指标。第*i*个定位源在*k*时刻的权重更新为 $w_i(k) = \alpha \cdot f(Q_i(k)) + (1 - \alpha) \cdot w_i(k-1)$, 其中 $f(Q_i(k))$ 为归一化信号质量评分, α 为指数平滑系数, 本实验取0.3。更新后执行 $w_i(k) \square w_i(k) / \sum_j w_j(k)$, 以保持权重和为1。

为了验证LLM模块在权重分配中的独立贡献, 本文增设了一组消融实验作为基线(Baseline)。该基线方法移除了LLM场景先验推理环节, 仅保留基于RSSI方差、SNR及多径指标的传统动态权重更新机制。相比之下LLM不直接输出物理坐标, 而是作为“语义-环境感知引擎”将非结构化的文本/逻辑推理转化为结构化的环境先验参数。LLM根据输入信号特征推理当前场

景, 输出场景标签后系统查表获取各传感器先验方差 σ_{prior}^2 , 作为初始权重再由衰减函数动态更新。

固定权重融合方法在三类场景下的RMSE分别为2.00 m、3.20 m和6.20 m; 仅基于传统信号质量指标的无LLM动态权重基线分别为1.15 m、2.10 m和4.50 m; 本文方法分别降至0.75 m、1.70 m和3.70 m。按39%、51%和10%的样本权重计算, 三种方法的整体加权RMSE分别为3.03 m、1.97 m和1.53 m。本文方法相较固定权重融合降低49.5%, 相较无LLM动态权重基线进一步降低22.3%。结果表明, LLM提供的语义级环境先验能够在传统信号质量动态加权基础上带来独立增益。

3.2.2 场景识别分类实验

室内场景识别分类是利用智能手机或物联网设备感知到的多源射频(RF)信号特征, 构建“环境射频指纹”。由于不同物理空间(如商场入口、中央广场、配电房等)在建筑结构、空间开阔度及电磁环境上存在显著差异, 其接收到的蜂窝网络与Wi-Fi信号在强度、数量、波动性及多径效应上会呈现出独特的统计规律。本次试验数据采集使用2部红米K70智能手机, 用于不同楼层的并行采集, 一般按1 m间隔设置指纹采样点, 部分楼层补充2 m间隔样本进行对比验证, 每个采集点停留3-5秒后取统计均值作为一个指纹样本; 信号采集频率为1 Hz, 每采样点采集5G RSRP、Wi-Fi RSSI及可见基站/AP数量等特征; 数据集按楼层进行划分, 以防止楼层间信号

表9 不同场景类型定位精度对比

场景类型	固定权重 RMSE/m	无LLM动态权重 RMSE/m	本文方法 RMSE/m	较固定权重降 低/%	较无LLM基线 降低/%
开阔无遮挡LOS中庭	2.00	1.15	0.75	62.5	34.8
单层隔断轻度NLOS 商铺/走廊	3.20	2.10	1.70	46.9	19.0
多层遮挡/设备稀疏边角区	6.20	4.50	3.70	40.3	17.8
整体加权平均	3.03	1.97	1.53	49.5	22.3

整体加权平均按各场景样本数占比计算, 权重系数分别为开阔LOS约39%、轻度NLOS约51%、重度遮挡约10%。

泄露, 训练集、验证集、测试集比例约为 7:1:2; 场景标注真值由人工实地标注结合 CAD 图纸区域划分确定。表 10 为进行微调后模型在场景识别分类下表现效果:

为验证该模型的场景识别能力, 将所使用的方法与固定权重融合、通用零样本 LLM 方案开展对比测试, 统计不同定位误差阈值下的样本累积分布情况。本实验场景识别任务基于第 2.3 节所述 12 类精细化场景分类体系 (S1-S12) 开展, 测试样本覆盖全部 12 类场景, 每类场景样本数依据实际分布比例抽样, 确保评估结果与真实工况一致。

由图 3 可知, 本文所使用的 LLM 微调模型的定位误差整体显著优于对比算法。

3.3 LLM 方法的规模化应用价值分析

传统 ML 方法在每个新增项目中均需执行完整的数据采集、标注、特征工程、模型训练与部署测试流程, 以下人工成本为基于团队既有实施流程的工程情景估算, 并非行业统一成本基准。一般人工成本大致为数据采集约 5 人天、人工标注约 3 人天、特征工程约 2 人天、模型训练与调

优约 3 人天、部署测试约 1 人天, 合计约 14 人天/项目, 累计部署成本近似线性增长。LLM 方法采用“基座模型统一部署+LoRA Adapter 场景化适配”的架构, 需一次性投入基座模型选型评估、微调框架搭建与 Prompt 工程设计等基础设施成本约为 30 人天, 其后单项目成本随领域数据积累逐步下降: 冷启动阶段单项目约 8~10 人天, 积累阶段降至 4~6 人天, 成熟阶段稳定在约 4 人天 (含数据采集约 2 人天、LoRA 微调约 1 人天、部署测试约 1 人天), 传统 ML 方法中的特征工程环节在 LLM 架构中由信号文本化编码模板设计所替代, 该模板可跨项目复用, 因此新增项目无需重复投入特征工程成本:

由图 4 可知, 当部署项目数量达到 10 个时, 传统 ML 方法的累计成本约为 LLM 辅助方法的 1.6 倍; 当项目数量达到 100 个时, 成本差距进一步扩大。在多场景重复部署条件下, “基座模型+轻量化适配”的方式具有较低的边际适配成本。从成本构成的演化过程来看, LLM 辅助方法的累计成本曲线呈现边际递减特征, 可分为三个阶段: 一是冷启动阶段 (第 1-5 个项目), 需分摊基

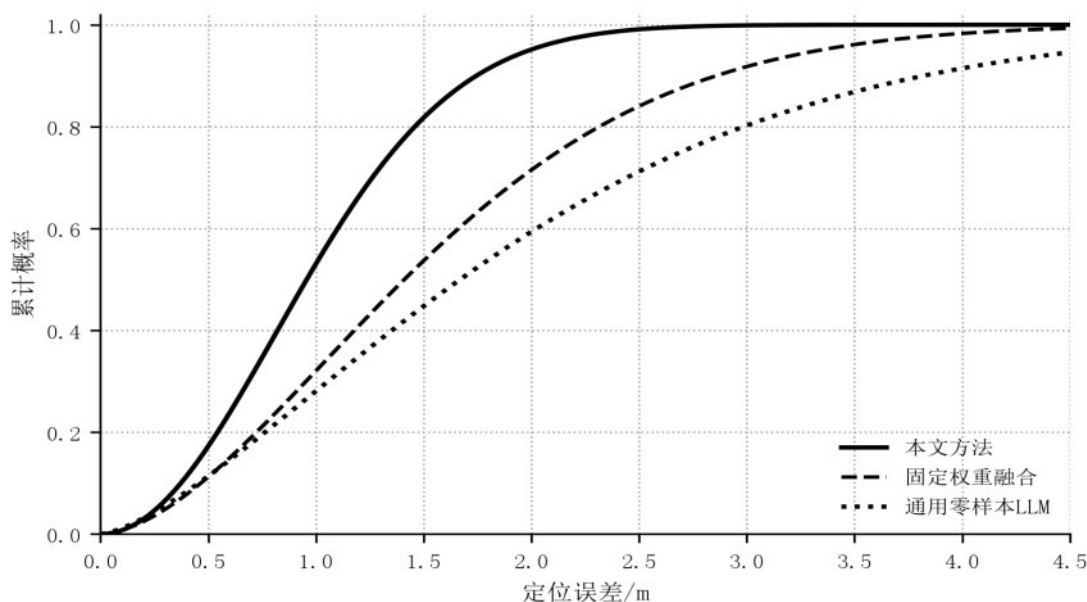


图3 定位误差累积分布函数(Cumulative Distribution Function, CDF)对比



表 10 数据采集的场景识别分类性能汇总

场景	样本数	准确率/%	精确率/%	召回率/%	F1
采集区 1-商场入口 LOS 蜂窝+Wi-Fi 双密集开阔 LOS 场景	320	96.3	95.8	96.9	0.963
采集区 2-店铺通道 LOS 蜂窝密集-Wi-Fi 稀疏开阔 LOS 场景	230	93.8	92.8	94.1	0.934
采集区 3-中央广场 LOS Wi-Fi 密集-蜂窝稀疏开阔 LOS 场景	260	95.8	96.1	95.4	0.957
采集区 4-商场入口品牌活动 隔断区 NLOS 双密集半遮挡 NLOS 场景	210	94.6	93.2	95.0	0.941
采集区 5-走廊外侧店铺通道 NLOS 蜂窝密集-Wi-Fi 稀疏半遮挡场景	200	91.5	90.2	92.3	0.912
采集区 6-中央广场品牌活动 隔断区 NLOS Wi-Fi 密集-蜂窝稀疏半遮挡场景	220	92.7	91.5	93.2	0.923
采集区 7-独立密闭区间 双密集重度遮挡 NLOS 场景	120	95.0	94.2	95.5	0.948
采集区 8-临街密闭商铺/配电房 蜂窝密集-Wi-Fi 稀疏重度遮挡 场景	80	97.5	96.8	97.8	0.973
采集区 9-地下车库 Wi-Fi 密集-蜂窝稀疏重度遮挡 场景	60	96.7	95.9	97.0	0.964
采集区 10-走廊开阔 双稀疏开阔 LOS 场景	160	94.4	93.6	94.7	0.941
采集区 11-建筑边角区域 双稀疏半遮挡 NLOS 场景	100	90.0	88.5	90.8	0.896
采集区 12-楼层衔接偏远角 双稀疏重度密闭 NLOS 场景	40	92.5	91.0	93.2	0.921
加权平均	2000	94.3	93.5	94.7	0.941

座模型选型评估、微调框架搭建与 Prompt 工程设计等基础设施成本，且 LoRA 权重从零开始训练，标注数据需求较多，单项目适配成本约 8~10 人天，累计投入高于传统 ML 方法；二是积累阶段（第 6-20 个项目），随着领域标注数据的积累和 LoRA 权重的逐步成熟，新增项目仅需增量微调即可完成适配，标注工作量因领域知识复用而减少约 50%，单项目成本降至 4~6 人天，累计成本曲线斜率明显放缓，并在此阶段与传统 ML 方法的成本曲线发生交叉；三是成熟阶段（第 21 个项

目及以后），领域 LoRA 模型已充分收敛，新增场景仅需极少量目标数据的轻量微调，部署测试流程亦实现标准化，单项目边际成本稳定在约 4 人天，累计成本曲线趋于平缓，与传统 ML 方法的成本差距持续扩大。

同时，LLM 方法当前仍面临一些技术局限：一是 LoRA 微调依赖领域标注数据的获取，在全新信号环境（如尚未建设的 6G 频段）下缺乏先验知识的冷启动问题仍有待研究；二是 LLM 推理的确定性保证问题，在安全关键定位应用中需结合冗余验证机制以规避模型幻觉风险；三是大规模并发推理的工程部署成本，当接入终端数量达到十万级以上时，LLM 推理服务器的算力投入仍需进一步优化。

4 结束语

本文参考 3GPP 定位框架并进行平台级扩展，明确了 LLM 辅助决策与传统定位解算的边界，将协议字段、信号统计特征和场景语义统一到平台侧调度流程中，通过多源数据接入、场景识别、约束化调度和融合解算提升复杂室内场景下的定位连续性。

实验结果表明，信号文本化编码与 LoRA 轻量化微调结合后可提升领域知识适配能力，场景识别加权准确率达到 94.3%；自适应融合策略可结合定位源质量、场景置信度和切换约束动态调整融合权重，开阔无遮挡 LOS 中庭场景定位均方根误差为 0.75 m，三类场景平均定位误差较固定权重融合方法降低 49.5%，跨场景定位连续性达到 88%。基于脱敏会话标识、MAC 哈希值等匿名标识构建动态映射关系（框架可扩展至蓝牙异构标识关联），实现 5G 与非 3GPP 短距定位平台之间的平滑接续。

总体来看，该方法在复杂室内场景下具有显著的工程应用价值，但仍需在更大规模并发、更多建筑类型和安全关键场景中进一步验证。后续

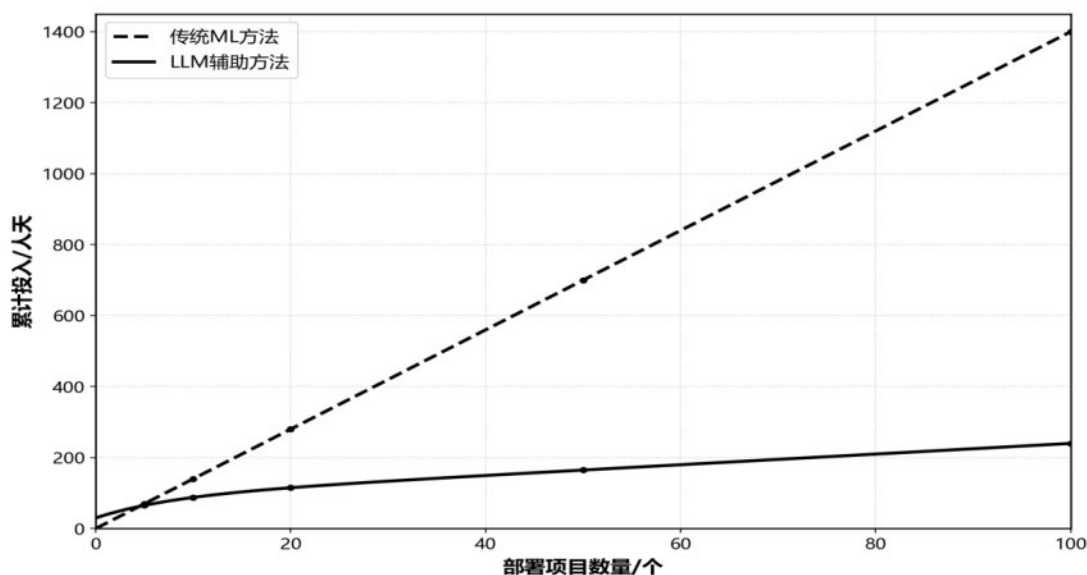


图4 ML与LLM方法规模化部署累计成本对比

研究可围绕多模态感知融合、端边协同轻量化部署和众包数据持续更新等方向开展，以进一步提升定位系统的鲁棒性、实时性和长期适配能力。

参考文献:

- [1] LIU H, DARABI H, BANERJEE P, et al. Survey of wireless indoor positioning techniques and systems[J]. IEEE Transactions on Systems, Man, and Cybernetics, Part C, 2007, 37(6): 1067-1080.
- [2] DARDARI D, CLOSAS P, DJURIC P M. Indoor tracking: theory, methods, and technologies[J]. IEEE Transactions on Vehicular Technology, 2015, 64(4): 1263-1278.
- [3] ZAFARI F, GKELIAS A, LEUNG K K. A survey of indoor localization systems and technologies[J]. IEEE Communications Surveys & Tutorials, 2019, 21(3): 2568-2599.
- [4] 3GPP. 3GPP TS 23.273 V19.0.0 5G System (5GS) location services (LCS); stage 2[S]. 2025.
- [5] 3GPP. 3GPP TS 38.305 V19.0.0 NG Radio Access Network (NG-RAN); stage 2 functional specification of user equipment positioning in NG-RAN[S]. 2025.
- [6] 3GPP. 3GPP TS 37.355 V19.0.0 LTE positioning protocol (LPP)[S]. 2025.
- [7] 3GPP. 3GPP TS 38.215 V19.0.0 NR; physical layer measurements[S]. 2025.
- [8] BREIMAN L. Random forests[J]. Machine Learning, 2001, 45(1): 5-32.
- [9] CHEN T, GUESTRIN C. XGBoost: a scalable tree boosting system[C]//Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. New York: ACM, 2016: 785-794.
- [10] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[C]//Advances in Neural Information Processing Systems 30. Red Hook: Curran Associates, 2017: 5998-6008.
- [11] DEVLIN J, CHANG M W, LEE K, et al. BERT: pre-training of deep bidirectional transformers for language understanding[C]//Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics. Minneapolis: ACL, 2019: 4171-4186.
- [12] WEI J, WANG X, SCHUURMANS D, et al. Chain-of-thought prompting elicits reasoning in large language models[C]//Advances in Neural Information Processing Systems 35. Red Hook: Curran Associates, 2022: 24824-24837.
- [13] HU E J, SHEN Y, WALLIS P, et al. LoRA: low-rank adaptation of large language models[C]//International Conference on Learning Representations. 2022.
- [14] DETTMERS T, PAGNONI A, HOLTZMAN A, et al. QLoRA: efficient finetuning of quantized LLMs[C]//Advances in Neural Information Processing Systems 36. Red Hook: Curran Associates, 2023: 10088-10115.

李成（1992年-），男，博士研究生，工程师，主要研究方向为5G-A/6G、卫星通信、融合定位。

牛锐（1991年-）男，硕士研究生，首席研究员，主要研究方向为低空互联网、定位网络、移动网络。

邵震（1977年-），男，硕士研究生，高级工程师，主要研究



方向为5G-A/6G感知、融合定位、WiFi。