

4-19-2022

Indoor Positioning Algorithm Based on XGBoost Prediction and Elastic Net Error Compensation

Xiaofei Kang

College of Communication and Information Engineering, Xi'an University of Science and Technology, Xi'an 710054, China; 949592499@qq.com

Xuan Zeng

College of Communication and Information Engineering, Xi'an University of Science and Technology, Xi'an 710054, China;

Wei Qiao

College of Communication and Information Engineering, Xi'an University of Science and Technology, Xi'an 710054, China;

Follow this and additional works at: <https://dc-china-simulation.researchcommons.org/journal>



Part of the [Artificial Intelligence and Robotics Commons](#), [Computer Engineering Commons](#), [Numerical Analysis and Scientific Computing Commons](#), [Operations Research](#), [Systems Engineering and Industrial Engineering Commons](#), and the [Systems Science Commons](#)

This Paper is brought to you for free and open access by Journal of System Simulation. It has been accepted for inclusion in Journal of System Simulation by an authorized editor of Journal of System Simulation.

Indoor Positioning Algorithm Based on XGBoost Prediction and Elastic Net Error Compensation

Abstract

Abstract: Aiming at the multi-depot half-open vehicle routing problem and *considering the soft time window constraints and vehicle speed changes*, an optimization model with the goal of maximizing average customer satisfaction, shortest distribution distance and minimum distribution cost is established and a two-stage solution algorithm is designed. The self-adaptive grid density method and neighborhood crowding density method are used to maintain the external archives and to select the global optimal particles, and the convergence of the multi-objective particle swarm optimization (MOPSO) and the diversity of the later population can be improved to obtain the initial feasible solution. The initial feasible solution is optimized by the variable neighborhood search algorithm (VNS) to reduce the delivery distance and cost. The rationality of the model and the effectiveness of the two-stage algorithm design are verified by the simulation experiments.

Keywords

time-varying speed, soft time window, multi-depot half-open, multi-objective particle swarm optimization, variable neighborhood search algorithm

Recommended Citation

Xiaofei Kang, Xuan Zeng, Wei Qiao. Indoor Positioning Algorithm Based on XGBoost Prediction and Elastic Net Error Compensation[J]. Journal of System Simulation, 2022, 34(4): 719-726.

基于 XGBoost 预测及弹性网误差补偿的室内定位算法

康晓非, 曾璇, 乔威

(西安科技大学 通信与信息工程学院, 陕西 西安 710054)

摘要: 为解决室内定位系统中因环境动态变化而导致定位精度下降的问题, 提出一种基于 XGBoost 并融合弹性网的误差补偿算法。采用 XGBoost 定位模型对目标位置进行初步预测, 当室内环境改变后, 再采用弹性网算法构建误差补偿模型, 修正 XGBoost 定位模型的定位误差, 并与基于 K 近邻、支持向量机、随机森林、梯度提升决策树等定位算法做对比。实验结果表明: 在更新 15% 指纹数据库样本的情况下, 该算法在 80% 分位处的定位精度控制在 0.73 m 以内, 明显优于其他定位算法, 且较基于 XGBoost 的定位算法精度提高了 25.5%。

关键词: 室内定位; WiFi 指纹; 极限梯度提升; 弹性网; 误差补偿

中图分类号: TP391; TN 92

文献标志码: A

文章编号: 1004-731X(2022)04-0719-08

DOI: 10.16182/j.issn1004731x.joss.20-0866

Indoor Positioning Algorithm Based on XGBoost Prediction and Elastic Net Error Compensation

Kang Xiaofei, Zeng Xuan, Qiao Wei

(College of Communication and Information Engineering, Xi'an University of Science and Technology, Xi'an 710054, China)

Abstract: Arming at the decreased positioning accuracy caused by the environment dynamic change of indoor positioning system, an error compensation algorithm based on XGBoost fusion elastic net is proposed. XGBoost positioning model is used to make a preliminary prediction on the target position. When the indoor environment changes, the elastic net algorithm is used to construct an error compensation model to correct the positioning error of XGBoost positioning model. The experimental results show that when only 15% of the fingerprint database samples need to be updated, the positioning accuracy of the proposed algorithm is controlled in 0.73m at the 80% percentile, which is significantly better than those of the K-nearest neighbor (KNN), support vector machine (SVM), random forest (RF) and gradient boosting decision tree (GBDT) positioning algorithms, and the accuracy increases 25.5% than XGBoost.

Keywords: indoor positioning; WiFi fingerprint; XGBoost; elastic net; error compensation

引言

基于位置的服务(location based services, LBS)正被广泛应用于医疗监控、救援服务、商业推广等众多领域, 而作为其核心支撑的定位技术也由室外延伸到室内^[1-2]。在室内定位领域, 定位技术

主要包括三边定位法和指纹定位法两大类。三边定位法是根据锚点(如 WiFi 接入点、蓝牙信标等)和定位点之间的测量值采用几何方法估计定位点的位置。这些测量值可以是接收信号强度(received signal strength, RSS)^[3]、到达时间(time of arrival, TOA)^[4]、到达时差(time difference of arrival,

收稿日期: 2020-11-09 修回日期: 2020-12-17

基金项目: 国家自然科学基金(61801372)

第一作者: 康晓非(1973-), 男, 博士, 副教授, 研究方向为室内无线定位、无线通信系统中信道估计及信号检测。

E-mail: 949592499@qq.com

TDOA)^[5]、往返传播时间(round-trip time of flight, RTOF)^[6]、到达角度(angle of arrival, AOA)^[7]、相机位姿^[8-9]等。此类方法一般定位精度较高,但基于RSS的方法依赖于精确的信道模型且在非视距(non line of sight, NLOS)情况下性能急剧下降;基于TOA, TDOA和RTOF的方法对时钟精度和同步要求较高;基于AOA和相机位姿的方法则受限于特定的应用环境和硬件设备^[2]。指纹定位法^[10]是通过指纹匹配估计位置信息。指纹是指某个定位点所能接收到的某种信号测量信息(RSS、信道状态信息、地磁等),例如WiFi指纹是一组来自可扫描到的接入点(access point, AP)的RSS测量值。指纹定位法一般分2个阶段实现:离线阶段通过数据测量构建RSS指纹向量与位置坐标之间的指纹库(radio map);在线阶段通过指纹匹配预测定位点位置坐标^[11]。此类方法实现较简单,特别是采用WiFi指纹定位无需额外布设专用设施,具有低成本和普适性的优势。但在实际环境下,因指纹的时变性以及指纹与物理位置的非确定性会直接影响定位精度,此外,指纹库的构建和更新也是很大的挑战。

考虑到指纹定位法本质上可以建模为监督学习问题,因此,一些诸如K近邻(K-nearest neighbor, KNN)^[12]、支持向量机(support vector machine, SVM)^[13]、随机森林(random forest, RF)^[14]等机器学习算法已被成功应用于指纹定位法中,可以有效提高定位的精度和稳定性。在此类研究中,一方面,定位精度与机器学习算法的性能有关,如文献[12]中的KNN算法易受AP数量及环境因素影响,其定位精度波动较大;文献[13]中的SVM算法当RSS数据为高维时变特征时,将会导致运算时间长且拟合的映射关系不适用;文献[14]中RF算法在RSS样本存在奇异点情况下会影响算法的分类精度和分类器的泛化能力。另一方面,指纹库也是影响此类方法定位精度的重要因素。指纹库通常是静态的,而室内环境可能是动态多变的,如房间布局、墙壁材料的变化和人体遮挡等因素都会引起无线传播环境的变化,使得原静

态指纹库与动态环境之间不能很好的适配,从而导致较大的定位误差^[15]。传统的解决方案是再次采集RSS样本以重建或更新指纹库,这在很大程度上增加了指纹库的维护成本。另一种解决方案是采用滤波器减少位置测量误差,其中最常用的是卡尔曼滤波器^[16-18]和粒子滤波器^[19-21]。卡尔曼滤波器对于统计数据为线性、高斯时性能最佳,而对于非高斯的实际环境,其性能损失较大;粒子滤波器适用于非线性、非高斯的情况,但其复杂度较高且会因存在粒子退化问题而使性能下降。

基于以上分析,针对动态环境下指纹库发生变化的应用场景,为有效提高室内定位精度,且对环境变化具有较好的鲁棒性,本文提出了一种采用极限梯度提升(extreme gradient boosting, XGBoost)进行初步位置预测并结合弹性网(elastic net)实现误差补偿(error compensation, EC)的室内定位算法(XGBoost-EC)。该算法仅需更新较少的指纹库样本,通过误差补偿方法便可有效解决因环境发生变化而导致定位精度下降的问题。

1 基于XGBoost预测及弹性网误差补偿的室内定位原理

XGBoost-EC室内定位方法如图1所示。该方法分为离线阶段和在线阶段:离线阶段在各参考点(reference point, RP)采集RSS值和其位置坐标信息,从而建立指纹数据库(radio map),该指纹库中的样本特征数据为RP位置的一组RSS值构成的向量,样本标签为RP位置坐标,基于该指纹数据库采用XGBoost算法训练并建立XGBoost预测模型。当环境动态变化后,针对新环境,采集并更新少量的RSS指纹库样本,将更新的部分样本用已建立好的XGBoost预测模型进行目标位置的初步预测,记录预测位置与真实位置的坐标误差并建立误差数据库,在该数据库中,样本特征数据为XGBoost预测位置坐标,样本标签为位置坐标误差,基于该误差数据库采用弹性网进行训练,

进而构建误差补偿模型; 在线阶段待定位点将扫描到的一组RSS值输入给XGBoost预测模型实现位置的初步定位, 因无线环境的变化, 该定位结果会存在较大的误差, 因此, 进一步将该定位结果输入给弹性网误差补偿模型完成误差预测, 在XGBoost初步定位的结果上进行误差补偿, 实现精确定位。在本文提出的算法中, XGBoost预测模型的优势在于通过二阶泰勒级数近似损失函数

不仅加快了拟合速度, 而且大大缩减了生成树的规模, 同时在目标函数中引入了正则项, 既控制了树的复杂度又防止了过拟合, 此外其并行处理方式则有助于加快训练速度, 可获得较高的定位精度和运行效率^[22-23]。弹性网误差补偿模型是在线性回归的基础上引入了L1和L2范式惩罚项^[24], 这既适合于稀疏模型的学习, 又具有良好的泛化能力, 能高效地估计误差, 完成误差补偿。

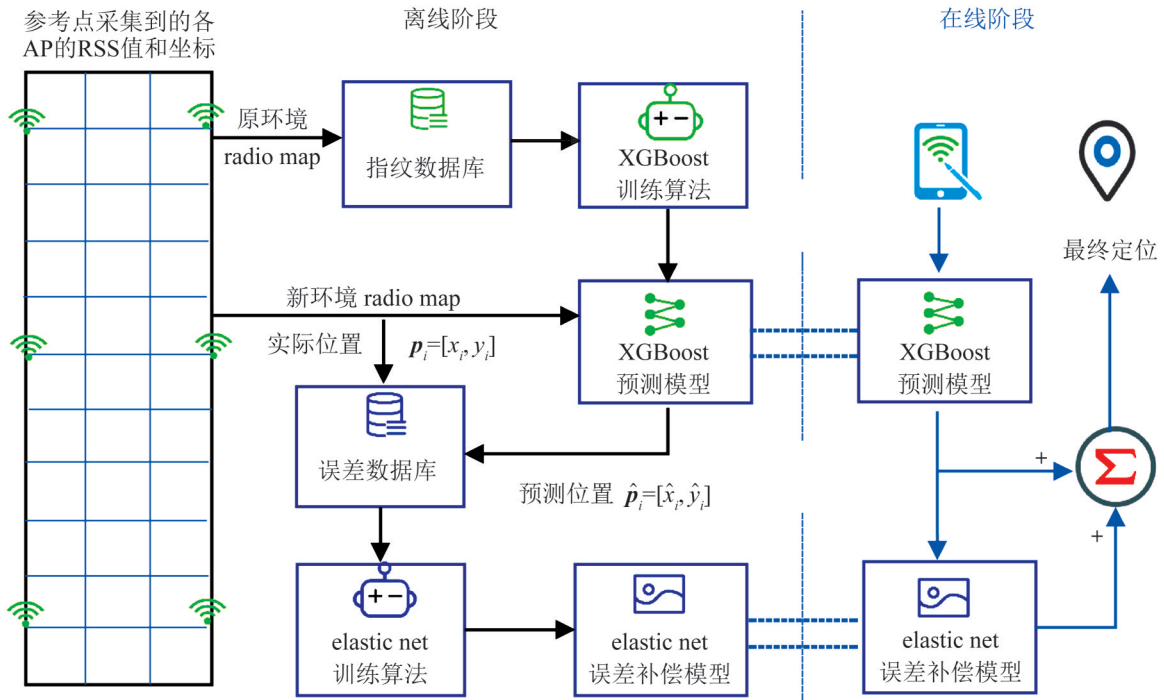


图1 基于XGBoost-EC的室内定位方法

Fig. 1 Indoor positioning method based on XGBoost-EC

2 基于XGBoost定位模型

对于包含 N 个样本的 WiFi 指纹数据集 $D = \{(\mathbf{r}_i, \mathbf{p}_i)\} (|D|=N, \mathbf{r}_i \in \mathbb{R}^M, \mathbf{p}_i \in \mathbb{R}^2)$, 其中 $\mathbf{r}_i = [r_i^1, r_i^2, \dots, r_i^M]$ 为 M 个 RSS 测量值组成的向量, 表示第 i 个 RP 的 RSS 指纹, $\mathbf{p}_i = [x_i, y_i]$ 为第 i 个 RP 的物理位置坐标。可将数据集 D 分解为 D_1 和 D_2 两个数据集并分别对其建模, 即 $D_1 = \{(\mathbf{r}_i, x_i)\} (|D_1|=N, \mathbf{r}_i \in \mathbb{R}^M, x_i \in \mathbb{R})$, $D_2 = \{(\mathbf{r}_i, y_i)\} (|D_2|=N, \mathbf{r}_i \in \mathbb{R}^M, y_i \in \mathbb{R})$ 。针对 D_1 和 D_2 两个数据集分别采用 XGBoost 算法对物理位置 $\mathbf{p}_i$ 的横坐标 x_i 和纵坐标 y_i 进行预测, 从而得到 $\mathbf{p}_i$ 的估计值

$\hat{\mathbf{p}}_i = [\hat{x}_i, \hat{y}_i]$ 。以数据集 D_2 为例, 若已经训练了 K 棵树, 则第 i 个样本的预测值为

$$\hat{y}_i^K = \sum_{k=1}^K f_k(\mathbf{r}_i), f_k \in F \quad (1)$$

式中: $f_k(\mathbf{r}_i)$ 为第 k 棵树对样本 $\mathbf{r}_i$ 的预测值。目标函数可以建模为

$$J = \sum_{i=1}^N L(y_i, \hat{y}_i^K) + \sum_{k=1}^K \Omega(f_k) \quad (2)$$

式中: $L(\cdot)$ 为损失函数; $L(y_i, \hat{y}_i^K)$ 为样本 $\mathbf{r}_i$ 的训练误差, 表征该模型对于训练集的匹配程度; $\Omega(f_k)$ 为第 k 棵树的正则项, 用于控制模型的复杂度,

防止过拟合。依据叠加式训练的特点,当训练第 k 棵树时,前 $k-1$ 棵树是已训练好的,因此,式(2)可转化为

$$J^k = \sum_{i=1}^N L(y_i, \hat{y}_i^{k-1} + f_k(\mathbf{r}_i)) + \Omega(f_k) \quad (3)$$

将式(3)通过二阶泰勒级数展开,并考虑到对于第 k 棵树, $L(y_i, \hat{y}_i^{k-1})$ 为常数,去除常数项后目标函数可简化为

$$J^k \approx \sum_{i=1}^N \left[g_i f_k(\mathbf{r}_i) + \frac{1}{2} h_i f_k^2(\mathbf{r}_i) \right] + \Omega(f_k) \quad (4)$$

式中: g_i 与 h_i 分别为损失函数的一阶偏导数和二阶偏导数,即 $g_i = \frac{\partial L(y_i, \hat{y}_i^{k-1})}{\partial \hat{y}_i^{k-1}}$, $h_i = \frac{\partial^2 L(y_i, \hat{y}_i^{k-1})}{\partial (\hat{y}_i^{k-1})^2}$ 。

为了求解,需将 $f_k(\mathbf{r}_i)$ 通过参数化表示,因此,定义第 k 棵树的第 t 个叶节点的权重为 ω_t ,全部的叶节点个数为 T ,样本 $\mathbf{r}_i$ 所在的叶节点位置为 $q(\mathbf{r}_i)$,第 k 棵树的第 t 个叶节点的全部样本集合为 $I_t = \{\mathbf{r}_i | q(\mathbf{r}_i) = t\}$ 。则表征一棵树复杂度的正则项 Ω 可以用叶节点的个数 T 和权重 ω_t 的平滑程度来表示:

$$\Omega(f_k) = \gamma T + \frac{1}{2} \lambda \sum_{t=1}^T \omega_t^2 \quad (5)$$

式中: γ 和 λ 为调整复杂度的超参数。式中前一项用叶节点个数乘以收缩系数 γ ,后一项用L2范数来表示叶节点的权重的平滑程度。将式(5)代入式(4)可得参数化后的新目标函数为

$$J^k = \sum_{i=1}^N (g_i \omega_{q(\mathbf{r}_i)} + \frac{1}{2} h_i \omega_{q(\mathbf{r}_i)}^2) + \gamma T + \frac{1}{2} \lambda \sum_{t=1}^T \omega_t^2 \quad (6)$$

将优化函数重新按照每个叶节点组合并舍弃常数项,则式(6)可进一步简化为

$$J^k = \sum_{t=1}^T \left[\left(\sum_{i \in I_t} g_i \right) \omega_t + \frac{1}{2} \left(\sum_{i \in I_t} h_i + \lambda \right) \omega_t^2 \right] + \gamma T \quad (7)$$

其中, G_t , H_t 为常值,因此,式(7)是一个关于二次函数求最优解问题,当树的结构固定时,叶节点位置 $q(\mathbf{r}_i)$ 已经固定,式(7)对 ω_t 求偏导并令其导函数为0,便可得当前树结构下的最优叶节点权重

和目标函数,可分别表示为

$$\omega_t^* = G_t / (H_t + \lambda) \quad (8)$$

$$J^* = -\frac{1}{2} \sum_{t=1}^T \frac{G_t^2}{H_t + \lambda} + \gamma T \quad (9)$$

式(9)可通过贪心算法寻找最好的切分点和树的结构,完成XGBoost最优模型的训练,得到 $\hat{y}_i$ 预测值,同理,针对 D_1 数据集建模可得到 $\hat{x}_i$ 的预测值,从而完成对物理位置坐标 $\mathbf{p}_i$ 的预测 $\hat{\mathbf{p}}_i = [\hat{x}_i, \hat{y}_i]$ 。

当室内环境发生变化时,原指纹数据库将失效,基于XGBoost的室内定位模型会产生较大的定位误差,为了提高定位精度,提出基于弹性网的误差补偿改进算法,该算法仅需更新少量指纹数据库,采用弹性网预测误差并通过误差补偿改善定位精度。

3 基于弹性网的误差补偿模型

对于变化的室内环境,采样少量指纹数据构建新的数据集 $S = \{(\hat{\mathbf{p}}_i, \mathbf{e}_i)\} (|S| = L, \hat{\mathbf{p}}_i \in \mathbb{R}^2, \mathbf{e}_i \in \mathbb{R}^2)$,其中, $\hat{\mathbf{p}}_i = [\hat{x}_i, \hat{y}_i]$ 定义为通过XGBoost算法对第 i 个RP的物理位置坐标的估计向量, $\mathbf{e}_i = [\Delta x_i, \Delta y_i]$ 定义为第 i 个RP分别在 x 轴和 y 轴方向上的估计值与真实值的误差向量,即

$$\Delta x_i = x_i - \hat{x}_i \quad (10)$$

$$\Delta y_i = y_i - \hat{y}_i \quad (11)$$

将数据集 S 分为 S_1 和 S_2 两个数据集并分别对其建模,即 $S_1 = \{(\hat{\mathbf{p}}_i, \Delta x_i)\} (|S| = L, \hat{\mathbf{p}}_i \in \mathbb{R}^2, \Delta x_i \in \mathbb{R})$, $S_2 = \{(\hat{\mathbf{p}}_i, \Delta y_i)\} (|S| = L, \hat{\mathbf{p}}_i \in \mathbb{R}^2, \Delta y_i \in \mathbb{R})$ 。针对 S_1 和 S_2 分别采用弹性网算法对误差值 Δx_i 和 Δy_i 进行预测,得到相应的预测值 $\Delta \hat{x}_i$ 和 $\Delta \hat{y}_i$,以 S_2 数据集为例,目标函数可以建模为

$$J(\theta) = \arg \min_{\theta} \left\{ \sum_{i=1}^L (\Delta y_i - P_{\theta}(\hat{x}_i, \hat{y}_i))^2 + \beta \sum_{j=1}^J [\alpha \theta_j^2 + (1 - \alpha) |\theta_j|] \right\} \quad (12)$$

式中: $P_{\theta}(\hat{x}_i, \hat{y}_i)$ 为 Δy_i 的预测值 $\Delta \hat{y}_i$,是关于 $\hat{x}_i$ 和 $\hat{y}_i$ 的多项式; $\theta = [\theta_1, \theta_2, \dots, \theta_J]$ 为多项式系数; α, β

为模型超参数。 $\Delta\hat{y}_i$ 可表示为

$$\Delta\hat{y}_i \approx P_\theta(\hat{x}_i, \hat{y}_i) \approx \theta_0 + \theta_1 \hat{x}_i + \theta_2 \hat{y}_i + \theta_3 \hat{x}_i \hat{y}_i + \theta_4 \hat{x}_i^2 + \theta_5 \hat{y}_i^2 + \dots + \theta_{j-1} \hat{x}_i^j + \theta_j \hat{y}_i^j \quad (13)$$

通过训练弹性网回归模型, 可求解出模型参数 θ , 从而得到预测值 $\Delta\hat{y}_i$ 。同理, 针对数据集 S_i 采用弹性网算法可得到预测值 $\Delta\hat{x}_i$ 。进一步通过误差补偿提高估计精度:

$$\hat{x}'_i = \hat{x}_i + \Delta\hat{x}_i \quad (14)$$

$$\hat{y}'_i = \hat{y}_i + \Delta\hat{y}_i \quad (15)$$

至此, 经过XGBoost算法的初步位置预测和弹性网算法的误差补偿后, 对室内平面内第 i 个物理位置坐标 p_i 的估计值为 $\hat{p}'_i = [\hat{x}'_i, \hat{y}'_i]$, XGBoost-EC算法的具体实现步骤如图2所示。

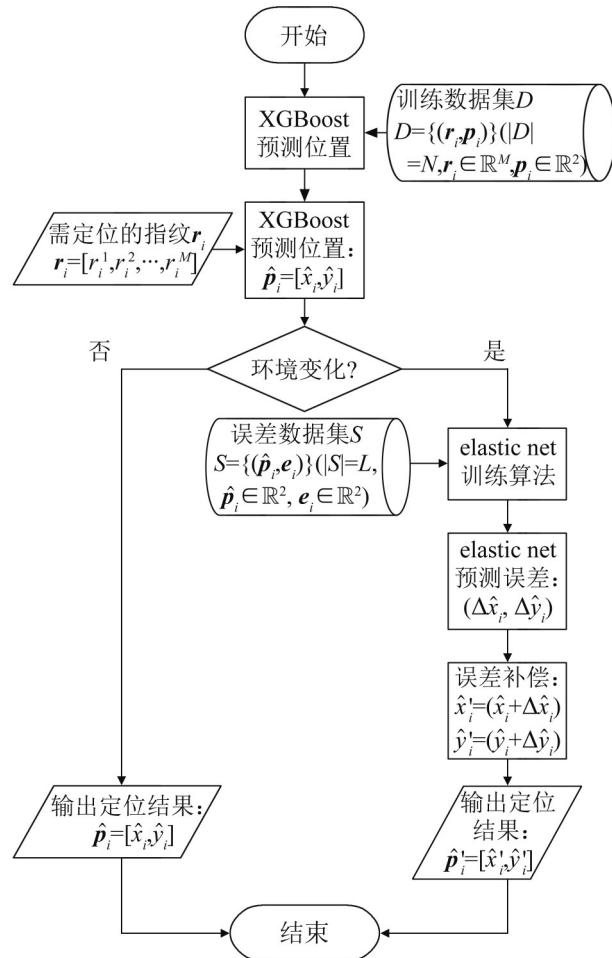


图2 XGBoost-EC的室内定位算法流程

Fig. 2 Flow chart of indoor positioning algorithm based on XGBoost-EC

4 实验结果及性能分析

如图3所示, 实验环境选取某大学研究生院地下停车场, 考虑一个长40 m(y轴), 宽5.5 m(x轴)的矩形区域, 共放置10个高度均为2 m的无线接入点(AP), 其平面坐标(x, y)分别为(0, 0), (5.5, 0), (0, 10), (5.5, 10), (0, 20), (5.5, 20), (0, 30), (5.5, 30), (0, 40), (5.5, 40); 将该矩形测试区域划分成880个0.5 m×0.5 m的网格, 每个网格的顶点设置为参考点(RP), 共得到960个RP点, 通过运行手机上的WiFi信号强度采集软件记录每个RP的位置坐标, 并测量该位置接收到的各AP的RSS值, 逐条存储每个RP的位置信息和RSS值, 实现离线位置指纹库的构建。



图3 测试实验环境

Fig. 3 Test experiment environment

定位精度是评估定位算法的重要指标^[25-26], 基于上述环境的位置指纹库, 图4比较了基于不同机器学习的定位算法误差累积分布函数(cumulative distribution function, CDF)性能^[26]。统计了实验数据的平均定位精度和80%分位处的定位精度, 如表1所示。从图4和表1可以看出, KNN算法的性能较差, SVM算法性能优于KNN, 但不及RF、GBDT和XGBoost这些集成学习算法。集成学习算法通过集合了多个弱学习器来提升机器学习的性能, 因此通常较单个的机器学习算法能获得更好的预测结果。

在这3种集成学习算法中, RF采用Bagging方式集成, 而GBDT采用Boosting方式集成, Boosting方式是基于梯度下降理论串行迭代地学习每个弱学习器, 迭代时通过设置前一级弱学习器错分样本较大的权重来提升性能, 因此GBDT获得了比RF更好的模型性能, XGBoost算法是GBDT算法的工程优化和创新, 较之GBDT, 其对损失函数进行了二阶泰勒级数展开, 利用二阶信息指导弱学习器的训练, 提高了训练的收敛速度, 同时, 在目标函数中引入了正则项, 防止了模型过拟合。因此这3种集成学习算法中XGBoost定位精度最高, GBDT次之。XGBoost-EC是在XGBoost算法基础上采用弹性网对定位精度进行了误差补偿, 进一步提高了定位精度, 其性能优于基于XGBoost的定位算法。

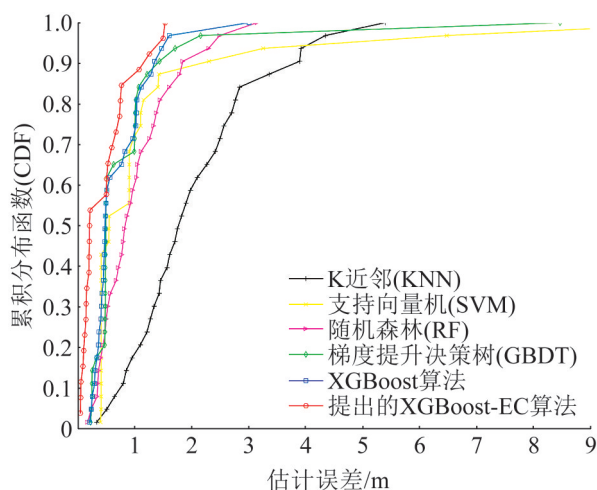


图4 不同机器学习定位算法的CDF性能比较
Fig. 4 Comparison of CDF performance of different machine learning-based indoor positioning algorithms

表1 几种室内定位算法的精度比较
Table 1 Comparison of accuracy of several indoor positioning algorithms

算法	平均定位精度/m	80%分位点精度/m	时间复杂度
KNN	2.02	2.75	$O(n \cdot m)$
SVM	1.21	1.33	$O(n^2 \cdot m)$
RF	1.02	1.48	$O(n \cdot \log n \cdot m \cdot k)$
GBDT	0.87	1.01	$O(n \cdot \log n \cdot m \cdot d)$
XGBoost	0.71	0.98	$O(p \cdot \log n + k \cdot d \cdot p)$
XGBoost-EC	0.47	0.73	$O(p \cdot \log n + k \cdot d \cdot p)$

注: n 为训练样本数量; m 为特征维度; k 为树的数量; d 为树的深度; p 为训练样本中特征列无缺失样本数

计算复杂度是评价定位算法的另一重要指标, XGBoost-EC算法在采用XGBoost进行预测中, 需要采用基于列块结构的预排序算法进行一次排序, 排序的时间复杂度为 $O(p \log n)$, 其中 n 为训练样本数, p 为训练样本中特征列无缺失样本数。在进一步构建树的过程中, 线性查找的时间复杂度为 $O(k \cdot d \cdot p)$, 其中 k 为树的数量, d 为树的深度, 故预测部分的时间复杂度为 $O(p \log n + kdp)$; 在采用弹性网进行误差补偿中, 弹性网络结合了岭回归和 LASSO(least absolute shrinkage and selection operator)回归, 两者具有相同的时间复杂度, 故误差补偿部分的时间复杂度为 $O(n_1 m_1^2)$, 其中 n_1 为训练样本数, m_1 为特征维度, 所以XGBoost-EC算法总时间复杂度为 $O(p \log n + kdp) + O(n_1 m_1^2)$ 。考虑到该算法在误差补偿时仅更新少量样本($n_1 \ll n$)且样本特征为坐标信息($m_1=2$), 因此, XGBoost-EC时间复杂度近似为 $O(p \log n + kdp)$, 与XGBoost算法的复杂度相当, 表1给出了基于机器学习的几种定位算法的时间复杂度比较。

考虑一个室内环境动态变化的场景, 通过人为加噪来模拟无线环境的改变, 图5分别给出了当2个和8个AP的RSS值受到标准方差为2的噪声污染后, 在仅更新原指纹数据库15%样本的情况下, XGBoost算法和XGBoost-EC算法的定位精度比较。从图5可以看出, 8个受到噪声污染AP的定位性能较2个受到噪声污染AP有所下降, 但每种情况下, XGBoost-EC算法均优于XGBoost定位算法, 这是因为采用弹性网构建误差补偿模型, 不仅可以有效学习稀疏模型, 而且提高了模型的泛化能力, 在较精确预测误差的基础上, 对原定精度进行了有效补偿。

选取一段运动轨迹来评估XGBoost-EC在动态场景下的跟踪性能。图6通过实验仿真了真实轨迹, 采用XGBoost算法的定位轨迹和XGBoost-EC定位轨迹的误差比较。图6表明, XGBoost-EC定位轨迹较XGBoost定位估计更接近于真实轨迹, 定位精度和跟踪性能较高。

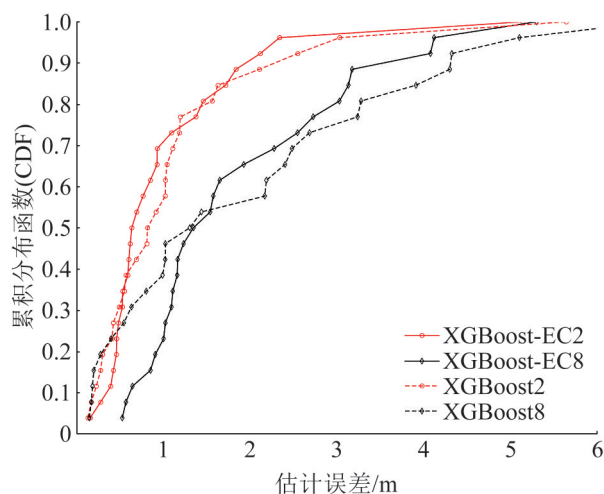


图5 不同AP数目加噪后的CDF性能比较

Fig. 5 Comparison of CDF performance of different noisy AP numbers

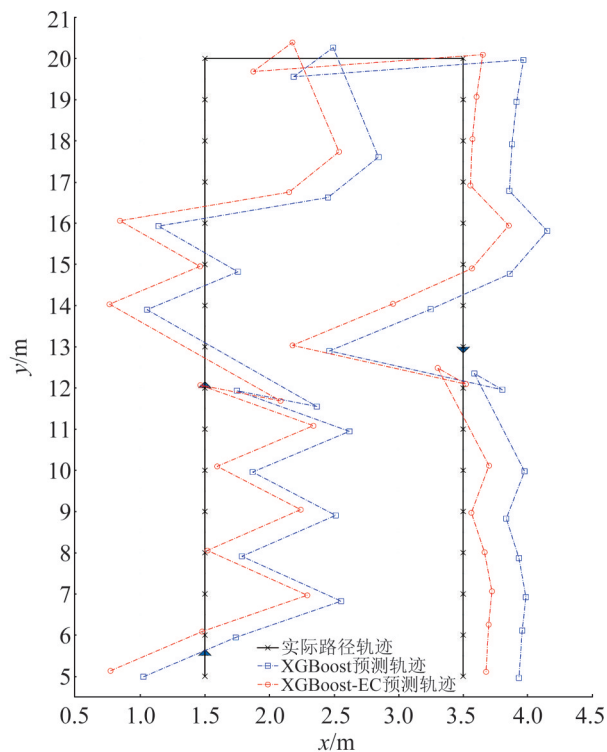


图6 路径轨迹定位精度比较

Fig. 6 Comparison of positioning accuracy of path trajectory

5 结论

在采用指纹法的室内定位系统中, 环境的动态变化会导致原指纹库失效而使定位精度急剧下降, 重建指纹库费时且耗力。为了解决这一问题,

提出了基于机器学习的误差补偿方法, 该方法通过XGBoost算法实现目标位置的初步预测并结合弹性网算法完成误差补偿。利用XGBoost集成学习的优势可有效提高初步定位的精度和稳健性。当室内环境发生改变时, 进一步采用弹性网构建误差补偿模型, 利用弹性网良好的泛化能力和快速的收敛特性可以高效、精确地估计由环境改变所引起的误差, 完成误差补偿。该方法仅需更新少量指纹库样本便可有效地提高定位精度且具有较强的稳健性, 很好地满足了在复杂、多变的室内环境下精确定位的需求。

进一步的研究工作可考虑在此框架中引入LightGBM算法实现位置的初步预测, LightGBM与XGBoost都是GBDT集成算法的工程优化和高效实现, LightGBM在训练速度和内存消耗方面具有较大的优势, 可提升定位算法的工程应用价值。在采用弹性网实现误差补偿时, 可考虑采用二阶中心有限差分法对数据进行预处理并重新定义损失函数, 同时引入地标信息进行校正, 以进一步提高定位精度和实现跟踪轨迹的平滑处理。

参考文献:

- [1] Zafari F, Gkelias A, Leung K K. A Survey of Indoor Localization Systems and Technologies[J]. IEEE Communications Surveys & Tutorials (S1553-877X), 2019, 21(3): 2568-2599.
- [2] Li Z, Xu K, Wang H, et al. Machine-Learning-Based Positioning: A Survey and Future Directions[J]. IEEE Network (S0890-8044), 2019, 33(3): 96-101.
- [3] Penada L, Azenha A, Carvalho A. Trilateration for Indoors Positioning within the Framework of Wireless Communications[C]//2009 35th Annual Conference of IEEE Industrial Electronics. Porto, Portugal: IEEE, 2009: 2732-2737.
- [4] Peterson B B, Kmiecik C, Hartnett R, et al. Spread Spectrum Indoor Geolocation[J]. Navigation (S2161-4296), 1998, 45 (2): 97-102.
- [5] Yamasaki R, Ogino A, Tamaki T, et al. TDOA Location System for IEEE 802.11 b WLAN [C]//IEEE Wireless Communications and Networking Conference. New Orleans, LA, USA: IEEE, 2005: 2338-2343.
- [6] Kim H. Double-sided Two-Way Ranging Algorithm to

- Reduce Ranging Time[J]. IEEE Communications Letters (S1558-2558), 2009, 13(7): 486-488.
- [7] Van Veen B D, Buckley K M. Beamforming: A Versatile Approach to Spatial Filtering[J]. IEEE Assp Magazine (S0740-7467), 1988, 5(2): 4-24.
- [8] Sadeghi H, Valaee S, Shirani S. Ocrapose. An Indoor Positioning System Using Smartphone/Tablet Cameras and OCR-aided Stereo Feature Matching[C]//2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). South Brisbane, Queensland, Australia: IEEE, 2015: 1473-1477.
- [9] Vedadi F, Valaee S. Automatic Visual Fingerprinting for Indoor Image-Based Localization Applications[J]. IEEE Transactions on Systems, Man, and Cybernetics: Systems (S2168-2216), 2017, 50(1): 305-317.
- [10] Davidson P, Piché R. A Survey of Selected Indoor Positioning Methods for Smartphones[J]. IEEE Communications Surveys & Tutorials (S1553-877X), 2016, 19(2): 1347-1370.
- [11] Sorour S, Lostanlen Y, Valaee S, et al. Joint Indoor Localization and Radio Map Construction with Limited Deployment Load[J]. IEEE Transactions on Mobile Computing (S1536-1233), 2014, 14(5): 1031-1043.
- [12] Ge X, Qu Z. Optimization WIFI Indoor Positioning KNN Algorithm Location-Based Fingerprint[C]//2016 7th IEEE International Conference on Software Engineering and Service Science. Beijing, China: IEEE, 2016: 135-137.
- [13] 桑楠, 袁兴中, 周瑞. 基于SVM分类和回归的WiFi室内定位方法[J]. 计算机应用研究, 2014, 31(6): 1820-1823.
- Sang Nan, Yuan Xingzhong, Zhou Rui. A WiFi Indoor Positioning Method Based on SVM Classification and Regression[J]. Application Research of Computers, 2014, 31(6): 1820-1823.
- [14] 郭妍, 陈晓, 任晓晔. 一种优化随机森林模型的室内定位方法[J]. 激光杂志, 2018, 39(10): 70-74.
- Guo Yan, Chen Xiao, Ren Xiaoye. An Indoor Positioning Method for Optimizing Random Forest Models[J]. Laser Journal, 2018, 39(10): 70-74.
- [15] Sou S I, Lin W H, Lan K C, et al. Indoor Location Learning over Wireless Fingerprinting System with Particle Markov Chain Model[J]. IEEE Access (S2169-3536), 2019, 7(4): 8713-8725.
- [16] Li M, Kang X, Qiao W. Performance Comparison and Evaluation of Indoor Positioning Technology Based on Machine Learning Algorithms[C]//2019 IEEE 19th International Conference on Communication Technology (ICCT). Xi'an, China: IEEE, 2019: 456-460.
- [17] Wang J, Park J G. A Novel Fingerprint Localization Algorithm Based on Modified Channel State Information Using Kalman Filter[J]. Journal of Electrical Engineering & Technology (S1975-0102), 2020, 15(6): 1811-1819.
- [18] Lipka M, Sippel E, Vossiek M. An Extended Kalman Filter for Direct, Real-Time, Phase-Based High Precision Indoor Localization[J]. IEEE Access (S2169-3536), 2019, 7(3): 25288-25297.
- [19] Gu F, Hu X, Ramezani M, et al. Indoor Localization Improved by Spatial Context-a Survey[J]. ACM Computing Surveys (CSUR) (S0360-0300), 2019, 52(3): 1-35.
- [20] Villacrés J L C, Zhao Z, Braun T, et al. A Particle Filter-Based Reinforcement Learning Approach for Reliable Wireless Indoor Positioning[J]. IEEE Journal on Selected Areas in Communications (S0733-8716), 2019, 37(11): 2457-2473.
- [21] 康晓非, 李梦梦, 乔威. 基于WiFi指纹的高精度室内定位融合算法[J]. 西安科技大学学报, 2020, 40(3): 470-476.
- Kang Xiaofei, Li Mengmeng, Qiao Wei. High-precision Indoor Localization Fusion Algorithm Based on WiFi Fingerprint[J]. Journal of Xi'an University of Science and Technology, 2020, 40(3): 470-476.
- [22] Chen T, Guestrin C. XGBoost: A Scalable Tree Boosting System[C]//22nd ACM Sigkdd International Conference on Knowledge Discovery and data mining. San Francisco, United States: Association for Computing Machinery, 2016: 785-794.
- [23] 张玄黎, 修春娣, 王延昭, 等. 基于CSI-XGBoost的高精度WiFi室内定位算法[J]. 北京航空航天大学学报, 2018, 44(12): 77-85.
- Zhang Xuanli, Xiu Chundi, Wang Yanzhao, et al. High-precision WiFi Indoor Localization Algorithm Based on CSI-XGBoost [J]. Journal of Beijing University of Aeronautics and Astronautics, 2018, 44(12): 77-85.
- [24] Zou H, Hastie T. Regularization and Variable Selection Via the Elastic Net[J]. Journal of the Royal Statistical Society: Series B (Statistical Methodology) (S1467-9868), 2005, 67(2): 301-320.
- [25] 李方敏, 张韬, 刘凯, 等. 基于距离测量和位置指纹的室内定位方法研究[J]. 计算机学报, 2019, 42(2): 109-120.
- Li Fangmin, Zhang Tao, Liu Kai, et al. An Indoor Localization Method Based on Location Fingerprint and Range Measurement[J]. Chinese Journal of Computers, 2019, 42(2): 109-120.
- [26] Sinha R S, Hwang S H. Comparison of CNN Applications for RSSI-based Fingerprint Indoor Localization[J]. Electronics (S2079-9292), 2019, 8(9): 989-1113.