



(12) 发明专利

(10) 授权公告号 CN 112004203 B

(45) 授权公告日 2022. 04. 12

(21) 申请号 202010961908.4

H04W 64/00 (2009.01)

(22) 申请日 2020.09.14

G01S 11/06 (2006.01)

(65) 同一申请的已公布的文献号  
申请公布号 CN 112004203 A

(56) 对比文件

CN 109951798 A, 2019.06.28

CN 109121081 A, 2019.01.01

(43) 申请公布日 2020.11.27

CN 106162868 A, 2016.11.23

(73) 专利权人 西安科技大学  
地址 710054 陕西省西安市雁塔区雁塔中路58号

审查员 韩祎

(72) 发明人 康晓非 姚萌 柳子惠 曾璇 乔威

(74) 专利代理机构 西安维赛恩专利代理事务所  
(普通合伙) 61257

代理人 李明全

(51) Int. Cl.

H04W 4/33 (2018.01)

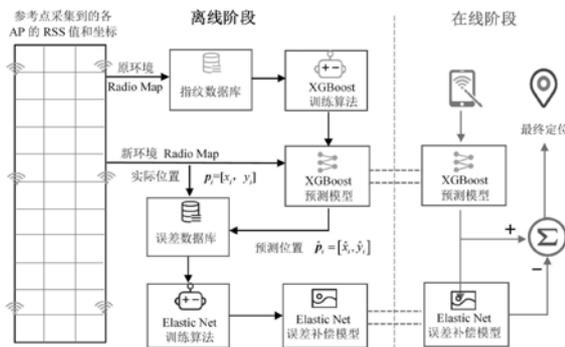
权利要求书3页 说明书10页 附图5页

(54) 发明名称

一种基于位置预测及误差补偿的室内定位方法与装置

(57) 摘要

本发明公开了一种基于位置预测及误差补偿的室内定位方法及装置,包括获取待预测位置的信号强度信息;以信号强度信息作为输入信息、结合位置预测模型对待预测位置进行位置预测,得到第一位置信息;以第一位置信息作为输入信息、结合误差修正模型对待预测位置进行位置预测,得到最终位置预测信息;本发明通过室内环境变化前的指纹库和位置估计模型进行用户的初步定位,再通过新的指纹库建立误差修正模型,对初步定位信息进行误差修正,达到了用户室内位置的精确定位,不再需要重新采集大量的指纹信息重新建立指纹库,也不需要再重新建立位置估计模型,降低了对重新采集的指纹信息的数据量,进而降低了当室内环境发生改变后定位装置的维护成本。



1. 一种基于位置预测及误差补偿的室内定位方法,其特征在于,包括以下步骤:

获取待预测位置的信号强度信息;

以所述信号强度信息作为输入信息、结合位置预测模型对所述待预测位置进行位置预测,得到第一位置信息;其中,所述位置预测模型通过室内环境发生变化前已有的第一指纹库训练得出;

以所述第一位置信息作为输入信息、结合误差修正模型对所述待预测位置进行位置预测,得到第二位置信息;其中,所述误差修正模型通过室内环境发生变化后重新采集的第二指纹库训练得到,所述第二位置信息为所述待预测位置的最终位置预测信息;

构建所述第二指纹库包括:

获取多组第二指纹数据集;其中,每组所述第二指纹数据集包括室内物理位置坐标及其对应的信号强度信息,且所述第二指纹数据集的数量小于所述第一指纹数据集的数量;

基于所述位置预测模型对每组所述第二指纹数据集的室内物理位置坐标进行预测,得到每组所述第二指纹数据集对应的预测室内位置坐标;

根据所述室内物理位置坐标和所述预测室内位置坐标生成室内物理位置坐标的误差值;

将每组所述第二指纹数据集中的室内物理位置坐标及其对应的误差值组合,得到第二指纹库。

2. 如权利要求1所述的一种基于位置预测及误差补偿的室内定位方法,其特征在于,获取待预测位置的信号强度信息包括:

重复获取待预测位置相对于室内每个AP的多个信号强度值;

计算所述待预测位置相对于室内每个AP的多个信号强度值的均值;

将所有所述均值组合成信号强度向量;其中,所述信号强度向量为所述信号强度信息。

3. 如权利要求2所述的一种基于位置预测及误差补偿的室内定位方法,其特征在于,所述位置预测模型的训练方法为:

构建所述第一指纹库;其中,所述第一指纹库包括多组第一指纹数据集,每组所述第一指纹数据集均包括室内物理位置坐标以及其对应的信号强度信息;

将所述第一指纹数据集分为第一指纹数据子集和第二指纹数据子集;其中,所述第一指纹数据子集包括所述室内物理位置坐标的横坐标及其对应的所述信号强度信息,所述第二指纹数据子集包括所述室内物理位置坐标的纵坐标及其对应的所述信号强度信息;

采用XGBoost算法的初始目标函数分别对所述横坐标和纵坐标进行预测,得到预测横坐标和预测纵坐标;

以每组所述第一指纹数据集中的室内物理位置坐标和对应的预测横坐标和预测纵坐标,结合通过贪心算法迭代优化所述初始目标函数,得到最优目标函数;其中,所述最优目标函数为位置预测模型。

4. 如权利要求3所述的一种基于位置预测及误差补偿的室内定位方法,其特征在于,所述最优目标函数为:

$$\mathcal{J}^* = -\frac{1}{2} \sum_{t=1}^T \frac{G_t^2}{H_t + \lambda} + \gamma T,$$

其中,  $t$ 为XGBoost算法中第 $k$ 棵树的叶子节点序数,  $T$ 为第 $k$ 棵树的叶子节点的总数,  $G_t$ 、 $H_t$ 均为常值,  $\gamma$ 和 $\lambda$ 为调整每棵树复杂度的超参数。

5. 如权利要求3或4所述的一种基于位置预测及误差补偿的室内定位方法, 其特征在于, 构建所述第一指纹库包括:

将室内环境发生变化前的室内区域划分为等尺寸的若干个网格;

在每个所述网格内采集室内物理位置坐标以及其对应的信号强度信息;

将采集的室内物理位置坐标以及对应的信号强度信息组合形成第一指纹库。

6. 如权利要求3所述的一种基于位置预测及误差补偿的室内定位方法, 其特征在于, 所述误差修正模型的训练方法为:

构建所述第二指纹库; 其中, 所述第二指纹库包括多组室内物理位置坐标及其对应的位置预测坐标;

采用弹性网算法对所述第二指纹库进行训练构建误差修正模型。

7. 如权利要求6所述的一种基于位置预测及误差补偿的室内定位方法, 其特征在于, 所述获取多组第二指纹数据集包括:

将室内环境发生变化后的室内区域划分为等尺寸的若干个网格;

在每个所述网格内采集室内物理位置坐标以及其对应的信号强度信息;

将所有网格内采集的室内物理位置坐标以及对应的信号强度信息组合得到多组第二指纹数据集。

8. 一种基于位置预测及误差补偿的室内定位装置, 其特征在于, 包括:

获取模块, 用于获取待预测位置的信号强度信息;

位置预测模块, 用于以所述信号强度信息作为输入信息、结合位置预测模型对所述待预测位置进行位置预测, 得到第一位置信息; 其中, 所述位置预测模型通过室内环境发生变化前已有的第一指纹库训练得出的位置预测模型;

误差修正模块, 用于以所述第一位置信息作为输入信息、结合误差修正模型对所述待预测位置进行位置预测, 得到第二位置信息; 其中, 所述误差修正模型通过室内环境发生变化后重新采集的第二指纹库训练得到, 所述第二位置信息为所述待预测位置的最终位置预测信息;

构建所述第二指纹库包括:

获取多组第二指纹数据集; 其中, 每组所述第二指纹数据集包括室内物理位置坐标及其对应的信号强度信息, 且所述第二指纹数据集的数量小于所述第一指纹数据集的数量;

基于所述位置预测模型对每组所述第二指纹数据集的室内物理位置坐标进行预测, 得到每组所述第二指纹数据集对应的预测室内位置坐标;

根据所述室内物理位置坐标和所述预测室内位置坐标生成室内物理位置坐标的误差值;

将每组所述第二指纹数据集中的室内物理位置坐标及其对应的误差值组合, 得到第二指纹库。

9. 一种基于位置预测及误差补偿的室内定位设备, 包括存储器、处理器以及存储在所述存储器中并可在所述处理器上运行的计算机程序, 其特征在于, 所述处理器执行所述计算机程序时实现如权利要求1至7任一项所述的一种基于位置预测及误差补偿的室内定位

方法。

## 一种基于位置预测及误差补偿的室内定位方法与装置

### 技术领域

[0001] 本发明属于室内定位技术领域,尤其涉及一种基于位置预测及误差补偿的室内定位方法与装置。

### 背景技术

[0002] 室内定位技术是基于位置的服务(Location Based Services,LBS),是重要商业应用的核心支撑技术。基于指纹的室内定位方法因能很好地解决非视距(Non Line of Sight,NLOS)信道条件下定位精度下降问题而备受关注。

[0003] 指纹定位方法的定位精度较大程度会受到位置指纹库(Radio Map)的影响。当室内环境的建筑结构、房间布局等发生变化时,室内无线传播环境也随之发生变化,从而导致其与已经建立的位置指纹库之间存在较大的差距,从而产生较大的定位误差。

[0004] 这时,需要再次采集RSS样本重建或更新指纹库。但是,重建或更新指纹库是一件非常耗时和费力的工作,很大程度上增加了指纹库的维护成本。

### 发明内容

[0005] 本发明的目的是提供一种基于位置预测及误差补偿的室内定位方法及装置,在室内环境发生变化时,仅需重新采集少量数据,且不会降低定位精度。

[0006] 本发明采用以下技术方案:一种基于位置预测及误差补偿的室内定位方法,包括以下步骤:

[0007] 获取待预测位置的信号强度信息;

[0008] 以信号强度信息作为输入信息、结合位置预测模型对待预测位置进行位置预测,得到第一位置信息;其中,位置预测模型通过室内环境发生变化前已有的第一指纹库训练得出;

[0009] 以第一位置信息作为输入信息、结合误差修正模型对待预测位置进行位置预测,得到第二位置信息;其中,误差修正模型通过室内环境发生变化后重新采集的第二指纹库训练得到,第二位置信息为待预测位置的最终位置预测信息。

[0010] 进一步地,获取待预测位置的信号强度信息包括:

[0011] 重复获取待预测位置相对于室内每个AP的多个信号强度值;

[0012] 计算待预测位置相对于室内每个AP的多个信号强度值的均值;

[0013] 将所有均值组合成信号强度向量;其中,信号强度向量为信号强度信息。

[0014] 进一步地,位置预测模型的训练方法为:

[0015] 构建第一指纹库;其中,第一指纹库包括多组第一指纹数据集,每组第一指纹数据集均包括室内物理位置坐标及其对应的信号强度信息;

[0016] 将第一指纹数据集分为第一指纹数据子集和第二指纹数据子集;其中,第一指纹数据子集包括室内物理位置坐标的横坐标及其对应的信号强度信息,第二指纹数据子集包括室内物理位置坐标的纵坐标及其对应的信号强度信息;

[0017] 采用XGBoost算法的初始目标函数分别对横坐标和纵坐标进行预测,得到预测横坐标和预测纵坐标;

[0018] 以每组第一指纹数据集中的室内物理位置坐标和对应的预测横坐标和预测纵坐标,结合通过贪心算法迭代优化初始目标函数,得到最优目标函数;其中,最优目标函数为位置预测模型。

[0019] 进一步地,最优目标函数为:

$$[0020] \quad \mathcal{J}^* = -\frac{1}{2} \sum_{t=1}^T \frac{G_t^2}{H_t + \lambda} + \gamma T,$$

[0021] 其中,t为XGBoost算法中第k棵树的叶子节点序数,T为第k棵树的叶子节点的总数, $G_t$ 、 $H_t$ 均为常值, $\gamma$ 和 $\lambda$ 为调整每棵树复杂度的超参数。

[0022] 进一步地,构建第一指纹库包括:

[0023] 将室内环境发生变化前的室内区域划分为等尺寸的若干个网格;

[0024] 在每个网格内采集室内物理位置坐标以及其对应的信号强度信息;

[0025] 将采集的室内物理位置坐标以及对应的信号强度信息组合形成第一指纹库。

[0026] 进一步地,误差修正模型的训练方法为:

[0027] 构建第二指纹库;其中,第二指纹库包括多组室内物理位置坐标及其对应的位置预测坐标;

[0028] 采用弹性网算法对第二指纹库进行训练构建误差修正模型。

[0029] 进一步地,构建第二指纹库包括:

[0030] 获取多组第二指纹数据集;其中,每组第二指纹数据集包括室内物理位置坐标及其对应的信号强度信息,且第二指纹数据集的数量小于第一指纹数据集的数量;

[0031] 基于位置预测模型对每组第二指纹数据集的室内物理位置坐标进行预测,得到每组第二指纹数据集对应的预测室内位置坐标;

[0032] 根据室内物理位置坐标和预测室内位置坐标生成室内物理位置坐标的误差值;

[0033] 将每组第二指纹数据集中的室内物理位置坐标及其对应的误差值组合,得到第二指纹库。

[0034] 进一步地,获取多组第二指纹数据集包括:

[0035] 将室内环境发生变化后的室内区域划分为等尺寸的若干个网格;

[0036] 在每个网格内采集室内物理位置坐标以及其对应的信号强度信息;

[0037] 将所有网格内采集的室内物理位置坐标以及对应的信号强度信息组合得到多组第二指纹数据集。

[0038] 本发明的另一种技术方案:一种基于XGBoost位置预测及误差补偿的室内定位装置,包括:

[0039] 获取模块,用于获取待预测位置的信号强度信息;

[0040] 位置预测模块,用于以信号强度信息作为输入信息、结合位置预测模型对待定位位置进行位置预测,得到第一位置信息;其中,位置预测模型通过室内环境发生变化前已有的第一指纹库训练得出的位置预测模型;

[0041] 误差修正模块,用于以第一位置信息作为输入信息、结合误差修正模型对待定位

位置进行位置预测,得到第二位置信息;其中,误差修正模型通过室内环境发生变化后重新采集的第二指纹库训练得到,第二位置信息为待预测位置的最终位置预测信息。

[0042] 本发明的另一种技术方案:一种基于XGBoost位置预测及误差补偿的室内定位设备,包括存储器、处理器以及存储在存储器中并可在处理器上运行的计算机程序,处理器执行计算机程序时实现上述任一项的一种基于XGBoost位置预测及误差补偿的室内定位方法。

[0043] 本发明的有益效果是:本发明通过室内环境变化前的指纹库和位置估计模型进行用户的初步定位,再通过新的指纹库建立误差修正模型,对初步定位信息进行误差修正,达到了用户室内位置的精确定位,不再需要重新采集大量的指纹信息重新建立指纹库,也不需要再重新建立位置估计模型,降低了对重新采集的指纹信息的数据量,进而降低了当室内环境发生改变后定位装置的维护成本。

## 附图说明

[0044] 图1为现有技术中机器学习结合滤波的室内定位方法的流程图;

[0045] 图2为本发明实施例中的室内定位方法流程图;

[0046] 图3为验证实施例中本发明方法与现有方法的CDF性能比较图;

[0047] 图4为验证实施例中当不同数量的AP受到相同噪声污染时本发明方法与现有方法的CDF性能比较图;

[0048] 图5为验证实施例中当相同数量的AP收到不同噪声污染时本发明方法与现有方法的CDF性能比较图;

[0049] 图6为验证实施例中本发明方法与现有方法的运动轨迹比较图。

## 具体实施方式

[0050] 下面结合附图和具体实施方式对本发明进行详细说明。

[0051] 在现有技术中,当室内环境发生变化后,需要再次采集RSS样本重建或更新指纹库。但是,重建或更新指纹库是一件非常耗时和费力的工作,很大程度上增加了指纹库的维护成本。

[0052] 为了克服这个问题,出现了基于梯度提升决策树(Gradient Boosting Decision Tree,GBDT)机器学习算法结合滤波的精确室内定位算法,如图1所示。该算法包括离线和在线两个阶段,实现步骤描述如下:离线阶段,1、先将室内环境划分成等尺寸的若干个网格;2、针对每个网格点采集并记录物理位置坐标以及WIFI信号的接收信号强度(Received Signal Strength,RSS)值,从而建立位置指纹库(Radio Map);3、采用GBDT集成学习算法对位置指纹库数据进行训练从而建立GBDT定位模型。在线阶段:4、手机用户在该室内环境中扫描所接收到的WiFi信号强度(RSS)值;5、利用训练好的GBDT模型预测手机用户的位置,该预测位置会受指纹库的变化而引起定位精度下降,在图1中称为位置粗估计;6、为了提高定位精度,进一步采用滤波算法进行跟踪补偿,如图1中的粒子滤波方法进行处理,从而得到位置的精细估计。

[0053] 上述方法主要是通过滤波,解决位置指纹库变化及提高定位精度问题,最常用滤波方法是卡尔曼滤波和粒子滤波。卡尔曼滤波产生的线性滤波器在基础数据统计数据为高

斯时是最佳的,但是,当统计数据偏离高斯性时,最优性就消失了,粒子滤波则会出现粒子退化而使定位精度下降。

[0054] 本发明提出了一种基于XGBoost初步定位并结合弹性网(Elastic Net)进行误差补偿(Error Compensation,EC)的室内定位方法(XGBoost-EC)。该方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的定位精度下降问题,

[0055] 该方法将初步定位和误差补偿分别建模为非线性机器学习问题,利用XGBoost的并行处理优势和良好泛化能力以及弹性网的稀疏模型学习能力和稳定性优点,可实现高精度定位,同时克服了滤波器适用条件受限的缺陷,本发明的方法在室内环境发生变化的条件下具有较高的定位精度和很好的稳健性。

[0056] 本发明公开了一种基于位置预测及误差补偿的室内定位方法,如图2所示,包括以下步骤:

[0057] 获取待预测位置的信号强度信息;以信号强度信息作为输入信息、结合位置预测模型对待预测位置进行位置预测,得到第一位置信息;其中,位置预测模型通过室内环境发生变化前已有的第一指纹库训练得出;以第一位置信息作为输入信息、结合误差修正模型对待预测位置进行位置预测,得到第二位置信息;其中,误差修正模型通过室内环境发生变化后重新采集的第二指纹库训练得到,第二位置信息为待预测位置的最终位置预测信息。

[0058] 本发明通过室内环境变化前的指纹库和位置估计模型进行用户的初步定位,再通过新的指纹库建立误差修正模型,对初步定位信息进行误差修正,达到了用户室内位置的精确定位,不再需要重新采集大量的指纹信息重新建立指纹库,也不需要再重新建立位置估计模型,降低了对重新采集的指纹信息的数据量,进而降低了当室内环境发生改变后定位装置的维护成本。

[0059] 在本发明的实施例中,获取待预测位置的信号强度信息包括:

[0060] 重复获取待预测位置相对于室内每个AP的多个信号强度值;计算待预测位置相对于室内每个AP的多个信号强度值的均值;将所有均值组合成信号强度向量;其中,信号强度向量为信号强度信息。

[0061] 在实际应用中,假设该方法应用在手机上,手机在某个位置点会收到多个AP的RSS值(信号强度值),在本实施例中是将多个RSS形成一个行向量,每个AP的RSS值对应着这个行向量中的一个元素。当然,每个AP的RSS值一般具有时变性,为了减少这种随机性,本实施例中设置在连续扫描AP五次,然后取平均值作为该AP的RSS值。

[0062] 在本发明实施例中,作为一种具体的实现方式,位置预测模型的训练方法为:

[0063] 构建第一指纹库;其中,第一指纹库包括多组第一指纹数据集,每组第一指纹数据集均包括室内物理位置坐标以及其对应的信号强度信息;将第一指纹数据集分为第一指纹数据子集和第二指纹数据子集;其中,第一指纹数据子集包括室内物理位置坐标的横坐标及其对应的信号强度信息,第二指纹数据子集包括室内物理位置坐标的纵坐标及其对应的信号强度信息;采用XGBoost算法的初始目标函数分别对横坐标和纵坐标进行预测,得到预测横坐标和预测纵坐标;以每组第一指纹数据集中的室内物理位置坐标和对应的预测横坐标和预测纵坐标,结合通过贪心算法迭代优化初始目标函数,得到最优目标函数;其中,最优目标函数为位置预测模型。

[0064] 具体地,最优目标函数为:

$$[0065] \quad \mathcal{J}^* = -\frac{1}{2} \sum_{t=1}^T \frac{G_t^2}{H_t + \lambda} + \gamma T,$$

[0066] 其中,  $t$  为 XGBoost 算法中第  $k$  棵树的叶子节点序数,  $T$  为第  $k$  棵树的叶子节点的总数,  $G_t$ 、 $H_t$  均为常值,  $\gamma$  和  $\lambda$  为调整每棵树复杂度的超参数。

[0067] 本发明实施例中, 构建第一指纹库包括:

[0068] 将室内环境发生变化前的室内区域划分为等尺寸的若干个网格; 在每个网格内采集室内物理位置坐标以及其对应的信号强度信息; 将采集的室内物理位置坐标以及对应的信号强度信息组合形成第一指纹库。

[0069] 具体的, 在本实施例中, 以信号强度信息作为输入信息、结合位置预测模型对待定位位置进行位置预测时采用基于 XGBoost 的定位模型。

[0070] 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  $M$  维接收信号强度矢量,  $\mathbf{p}_i = [x_i, y_i]$  为第  $i$  个参考点的物理位置坐标。

[0071] 可将数据集  $D$  分解为  $D_1$  和  $D_2$  两个数据集分别进行建模, 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  $\hat{\mathbf{p}}_i = [\hat{x}_i, \hat{y}_i]$ 。

[0072] 以数据集  $D_2$  为例, 假设已经训练了  $K$  棵树, 则对于第  $i$  个样本的预测值为:

$$[0073] \quad \hat{y}_i^{(K)} = \sum_{k=1}^K f_k(\mathbf{r}_i), \quad f_k \in F \quad (1)$$

[0074] 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:

$$[0075] \quad \mathcal{J} = \sum_{i=1}^N \mathcal{L}(y_i, \hat{y}_i^{(K)}) + \sum_{k=1}^K W(f_k) \quad (2)$$

[0076] 式中,  $\mathcal{L}(\bullet)$  表示损失函数,  $\Omega(f_k)$  表示第  $k$  棵树的复杂度,  $N$  为样本数。根据叠加式训练 (Additive Training) 的特点, 当训练第  $k$  棵树时, 前  $k-1$  棵树是已知的, 因此, 式 (2) 的最优化问题可以表示为:

$$[0077] \quad \mathcal{J}^{(k)} = \sum_{i=1}^N \mathcal{L}(y_i, \hat{y}_i^{(k-1)} + f_k(\mathbf{r}_i)) + W(f_k) \quad (3)$$

[0078] 进一步, 将目标函数通过二阶泰勒级数近似简化:

$$[0079] \quad \mathcal{J}^{(k)} = \sum_{i=1}^N g_i f_k(\mathbf{r}_i) + \frac{1}{2} h_i f_k^2(\mathbf{r}_i) + W(f_k) \quad (4)$$

[0080] 其中,  $g_i = \frac{\partial \mathcal{L}(y_i, \hat{y}_i^{(k-1)})}{\partial \hat{y}_i^{(k-1)}}$ ,  $h_i = \frac{\partial^2 \mathcal{L}(y_i, \hat{y}_i^{(k-1)})}{\partial^2 \hat{y}_i^{(k-1)}}$ ,  $g_i, h_i$  均为常数。

[0081] 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  $t$  个叶节点的值为  $\omega_t$ , 全部的叶子节点个数为  $T$ , 样本  $\mathbf{r}_i$  所在的叶节点位置为  $q(\mathbf{r}_i)$ , 第  $k$  棵树的第  $t$  个节点的全部样本集合为  $I_t = \{\mathbf{r}_i | q(\mathbf{r}_i) = t\}$ 。则一棵树的复杂度  $\Omega$  可以由两部分表示:

$$[0082] \quad W(f_k) = \gamma T + \frac{1}{2} \lambda \sum_{t=1}^T \omega_t^2 \quad (5)$$

[0083] 其中,  $\gamma$  和  $\lambda$  为调整复杂度的超参数。式 (4) 参数化后的新目标函数可表示为:

$$[0084] \quad \mathcal{J}^{(k)} = \sum_{i=1}^N \left( g_i \omega_{q(r_i)} + \frac{1}{2} h_i \omega_{q(r_i)}^2 \right) + \gamma T + \frac{1}{2} \lambda \sum_{i=1}^T \omega_i^2 \quad (6)$$

[0085] 进一步可化简为:

$$[0086] \quad \mathcal{J}^{(k)} = \sum_{i=1}^T \left[ \underbrace{\sum_{i \in I_t} g_i}_{\text{constant } G_t} \omega_t + \frac{1}{2} \left( \underbrace{\sum_{i \in I_t} h_i}_{\text{constant } H_t} + \lambda \right) \omega_t^2 \right] + \gamma T \quad (7)$$

[0087] 其中,  $G_t, H_t$  是常值, 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, 当前树结构下的最优目标函数值为:

$$[0088] \quad \mathcal{J}^* = -\frac{1}{2} \sum_{i=1}^T \frac{G_t^2}{H_t + \lambda} + \gamma T \quad (8)$$

[0089] 上式可通过贪心算法寻找最好的切分点和树的结构, 从而完成 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  $p_i$  的估计  $\hat{p}_i = [\hat{x}_i, \hat{y}_i]$ 。

[0090] 当环境发生变化时, 基于 XGBoost 的室内定位模型精度将下降, 为了提高定位精度, 提出了一种基于多项式回归误差补偿算法, 该算法仅需更新少量数据库, 便可实现定位精度的较大提升。

[0091] 具体的, 误差修正模型的训练方法为:

[0092] 构建第二指纹库; 其中, 第二指纹库包括多组室内物理位置坐标及其对应的位置预测坐标; 采用弹性网算法对第二指纹库进行训练构建误差修正模型。在该实施例中, 构建第二指纹库包括:

[0093] 获取多组第二指纹数据集; 其中, 每组第二指纹数据集包括室内物理位置坐标及其对应的信号强度信息, 且第二指纹数据集的数量小于第一指纹数据集的数量; 基于位置预测模型对每组第二指纹数据集的室内物理位置坐标进行预测, 得到每组第二指纹数据集对应的预测室内位置坐标; 根据室内物理位置坐标和预测室内位置坐标生成室内物理位置坐标的误差值; 将每组第二指纹数据集中的室内物理位置坐标及其对应的误差值组合, 得到第二指纹库。

[0094] 更为具体的, 获取多组第二指纹数据集包括:

[0095] 将室内环境发生变化后的室内区域划分为等尺寸的若干个网格; 在每个网格内采集室内物理位置坐标以及其对应的信号强度信息; 将所有网格内采集的室内物理位置坐标以及对应的信号强度信息组合得到多组第二指纹数据集。

[0096] 在本实施例中, 误差修正模型的训练方法如下:

[0097] 对于变化的环境, 采样少量指纹数据构建新的数据集  $S = \{(\hat{p}_i, e_i)\} (|S| = L, \hat{p}_i \in \mathbb{R}^2, e_i \in \mathbb{R}^2)$ ,

其中,  $\hat{p}_i = [\hat{x}_i, \hat{y}_i]$  定义为通过 XGBoost 算法对第  $i$  个参考点的物理位置坐标的估计向量。  $e_i = [\Delta x_i, \Delta y_i]$  定义为第  $i$  个参考点位置的  $x$  轴与  $y$  轴方向上, 估计值与真实的误差向量, 即:

$$[0098] \quad \Delta x_i = x_i - \hat{x}_i \quad (9)$$

$$[0099] \quad \Delta y_i = y_i - \hat{y}_i \quad (10)$$

[0100] 将数据集  $S$  分解为  $S_1$  和  $S_2$  两个数据集分别进行建模, 即  $S_1 = \{(\hat{p}_i, D x_i)\} (|S_1|=L, \hat{p}_i \in \mathbb{R}^2, D x_i \in \mathbb{R})$ ,  $S_2 = \{(\hat{p}_i, D y_i)\} (|S_2|=L, \hat{p}_i \in \mathbb{R}^2, D y_i \in \mathbb{R})$ 。针对  $S_1$  和  $S_2$  两个数据集分别采用弹性网算法实现对误差值  $\Delta x_i$  和  $\Delta y_i$  进行预测, 得到相应的预测值  $D \hat{x}_i$  和  $D \hat{y}_i$ , 以  $S_2$  数据集为例, 目标函数可以建模为:

$$[0101] \quad \mathcal{J}(\theta) = \arg \min_{\theta} \left\{ \sum_{i=1}^L (D y_i - P_{\theta}(\hat{x}_i, \hat{y}_i))^2 + \beta \sum_{j=1}^J [\alpha \theta_j^2 + (1-\alpha) |\theta_j|] \right\} \quad (11)$$

[0102] 其中,  $P_{\theta}(\hat{x}_i, \hat{y}_i)$  表示  $\Delta y_i$  的预测值  $D \hat{y}_i$ , 是关于  $\hat{x}_i$  和  $\hat{y}_i$  的多项式,  $\theta = [\theta_1, \theta_2, \dots, \theta_J]$  表示该多项式系数,  $\alpha, \beta$  为模型超参数。  $D \hat{y}_i$  可表示为:

$$[0103] \quad D \hat{y}_i \approx P_{\theta}(\hat{x}_i, \hat{y}_i) = \theta_0 + \theta_1 \hat{x}_i + \theta_2 \hat{y}_i + \theta_3 \hat{x}_i \hat{y}_i + \theta_4 \hat{x}_i^2 + \theta_5 \hat{y}_i^2 + \dots + \theta_{J-1} \hat{x}_i^J + \theta_J \hat{y}_i^J \quad (12)$$

[0104] 通过训练弹性网回归模型, 可求解出模型参数  $\theta$ , 从而得到预测值  $D \hat{y}_i$ 。同理, 针对数据集  $S_1$  采用弹性网算法可得到预测值  $D \hat{x}_i$ 。进一步通过误差补偿提高估计精度, 即:

$$[0105] \quad \hat{x}'_i = \hat{x}_i + D \hat{x}_i \quad (13)$$

$$[0106] \quad \hat{y}'_i = \hat{y}_i + D \hat{y}_i \quad (14)$$

[0107] 至此, 经过 XGBoost 算法初步位置预测和弹性网 (Elastic Net) 算法误差补偿 (EC) 后, 对室内平面内第  $i$  个物理位置坐标  $p_i$  的估计值为  $\hat{p}'_i = [\hat{x}'_i, \hat{y}'_i]$ 。

[0108] 本发明的方法与原有的方法的区别在于: 当室内环境发生变化时, 原有的方法是重新采集指纹重建指纹数据库, 因为, 虽然环境变化了但指纹库也重新更新了, 即便采用原有的算法定位精度也可以保证。这种方法缺陷在于重新采集指纹数据是一个费时耗力的艰巨工作, 大幅度增加指纹数据库的维护成本。

[0109] 本发明的方法是保留原有的指纹数据库和原有的预测算法, 在此基础上, 新建误差补偿指纹库, 该误差补偿指纹库是通过采集少量指纹而建立的, 并引入新的弹性网算法进行训练, 进而利用训练好的弹性网模型进行误差补偿估计, 最好利用估计出的误差值对原有算法估计信息进行修正。总之, 当室内环境发生变化时, 仅需更新少量指纹样本, 建立误差补偿指纹库, 利用弹性网算法进行训练进而预测误差, 利用该误差对原有 XGBoost 算法定位估计进行修正, 提高定位精度和定位的稳健性。

[0110] 对于变化的环境, 原有的指纹库会导致较大的定位估计误差。如果重新采集建立新的指纹库是耗时耗力的繁重工作, 避免这样做的一个思路是, 在原有指纹库基础上进行采集更新少量的指纹数据, 利用更新的指纹数据构建一个误差补偿数据库, 通过误差补偿数据库训练算法实现误差补偿, 进而对通过原有指纹库预测出的较大定位误差进行补偿, 提高定位精度。针对误差补偿数据库, 考虑到可以建模为一个非线性监督学习问题, 因此, 可以采用多种机器学习算法进行处理, 例如, 多项式回归算法、支持向量机算法等, 本发明采用的是弹性网算法, 较之其它算法, 弹性网适用于稀疏模型, 且具有更好的泛化能力, 这对提高定位精度以及定位模型的稳健性都有帮助。

[0111] 本发明的目标是更新少量指纹便可提高定位精度, 为了这个目标, 采集少量指纹建立误差指纹库, 通过弹性网算法训练误差指纹库。当然误差指纹库中的样本也可以很多, 这样更有利于训练弹性网模型, 但采集是一个艰巨的工作, 这背离了本发明的目标。

[0112] 总之本发明方法与原有的方法的区别在于:当环境发生变化时,原有的方法是重新采集指纹重建指纹数据库,因为虽然环境变化了但指纹库也重新更新了,即便采用原有的算法定位精度也可以保证。本发明方法是保留原有的指纹数据库和原有的预测算法,在此基础上,新建误差补偿指纹库,误差补偿指纹库是通过采集少量指纹而建立的,并引入新的弹性网算法,进行误差补偿估计,进一步利用估计出的误差值对原有算法估计信息进行修正。

[0113] 由于本发明依次通过位置预测模型和误差修正模型对采集信息进行处理,由于使用位置预测模型进行位置预测时本身具有一定的精度,只是由于室内环境变化,使得位置预测精度降低,因此,只需要在通过误差修正模型对位置预测信息进行修整。由于进行误差修正的数据源本身具有一定的精度,所以误差修正模型不需要具有非常高的精度,进而可以采集较少的数据指纹进行训练,避免了采集大量数据指纹进行训练的过程,

[0114] 该方法也可分为离线阶段和在线阶段。在离线阶段,首先在各参考点(RP)采集RSS值和坐标信息,从而建立指纹数据库。基于该指纹数据库采用XGBoost算法训练并建立XGBoost定位模型,利用该模型实现位置的初步预测。

[0115] 当环境发生变化后,仅需更新较少的RSS指纹库样本,并利用更新样本建立误差指纹库,基于误差数据库利用弹性网(Elastic Net)算法构建误差补偿模型(弹性网在线性回归的基础上加上了L1范式和L2范式共同作为惩罚项。这种组合既可以学习稀疏模型,同时可以保持岭回归的正则化特性,能高效地估计误差,完成误差补偿)。

[0116] 在线阶段,利用离线阶段已训练好的XGBoost预测模型和弹性网(Elastic Net)误差补偿模型实现精确定位。

[0117] 本发明的另一种技术方案:一种基于XGBoost位置预测及误差补偿的室内定位装置,包括:

[0118] 获取模块,用于获取待预测位置的信号强度信息;

[0119] 位置预测模块,用于以信号强度信息作为输入信息、结合位置预测模型对待定位位置进行位置预测,得到第一位置信息;其中,位置预测模型通过室内环境发生变化前已有的第一指纹库训练得出的位置预测模型;

[0120] 误差修正模块,用于以第一位置信息作为输入信息、结合误差修正模型对待定位位置进行位置预测,得到第二位置信息;其中,误差修正模型通过室内环境发生变化后重新采集的第二指纹库训练得到,第二位置信息为待预测位置的最终位置预测信息。

[0121] 本发明的另一种技术方案:一种基于XGBoost位置预测及误差补偿的室内定位设备,包括存储器、处理器以及存储在存储器中并可在处理器上运行的计算机程序,处理器执行计算机程序时实现上述任一项的一种基于XGBoost位置预测及误差补偿的室内定位方法。

[0122] 需要说明的是,上述模块之间的信息交互、执行过程等内容,由于与本发明方法实施例基于同一构思,其具体功能及带来的技术效果,具体可参见方法实施例部分,此处不再赘述。

[0123] 所属领域的技术人员可以清楚地了解到,为了描述的方便和简洁,仅以上述各功能模块的划分进行举例说明,实际应用中,可以根据需要而将上述功能分配由不同的功能模块完成,即将所述装置的内部结构划分成不同的功能模块,以完成以上描述的全部或者

部分功能。实施例中的各功能模块可以集成在一个处理单元中,也可以是各个单元单独物理存在,也可以两个或两个以上单元集成在一个单元中,上述集成的单元既可以采用硬件的形式实现,也可以采用软件功能单元的形式实现。另外,各功能模块的具体名称也只是为了便于相互区分,并不用于限制本申请的保护范围。上述装置中模块的具体工作过程,可以参考前述方法实施例中的对应过程,在此不再赘述。

[0124] 验证实施例:

[0125] 实验环境选取某地下停车场,考虑一个长40m(y轴),宽5.5m(x轴)的矩形区域,共放置10个高度均为2米无线接入点(Access Point,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0.5m×0.5m的网格,每个网格的顶点设置为参考点(Reference Point,RP),共得到960个RP点;采用本实施例中开发的手机APP软件测试并记录每个RP点所接收到的各AP的信号强度RSS值用于构建离线位置指纹库(Radio Map)。

[0126] 定位精度是评价定位算法的重要指标,基于上述环境位置指纹库,图3比较了包含本发明实施例的方法以及其他基于不同机器学习的定位算法的估计误差累积分布函数(Cumulative Distribution Function,CDF)性能,为了进一步更直观的说明该实验结果,基于该图中的实验数据统计结果(平均定位精度和80%的百分位处定位精度)如表1所示。

[0127] 从图3和表1可以看出:KNN算法实现简单但其CDF性能较差,SVM算法性能优于KNN,但不及RF、GBDT、XGBoost这些集成机器学习算法,在这三种集成学习算法中,XGBoost性能最好,本发明提出的XGBoost-EC是在XGBoost算法基础上采用弹性网对定位精度进行误差补偿,进一步提高了定位精度,其性能优于基于XGBoost的定位算法。

[0128] 表1几种定位算法的精度比较

算法	平均定位精度	80%分位点精度	算法运行时间
KNN	2.02m	2.75m	0.992 ms
SVM	1.21m	1.33m	14.96ms
RF	1.02 m	1.48 m	108ms
GBDT	0.87m	1.01m	223ms
XGBoost	0.71 m	0.98m	49.8ms
XGBoost-EC	0.47 m	0.73m	50.7ms

[0130] 本发明提出的算法的一个重要的应用场景是当环境发生变化后,仅需更新少量离线位置指纹库,便可有效补偿定位精度的损失。

[0131] 图4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,采用XGBoost算法和提出的XGBoost-EC算法进行定位的CDF性能比较,图中表明,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,当每种情况下,提出的算法均优于XGBoost定位算法。

[0132] 图5为4个受到噪声污染AP的环境下,噪声的标准方差分别取1、3、5、7时定位精度

CDF性能比较,图中表明,提出算法的定位精度受噪声的标准方差影响,随着标准方差值的增大,定位精度CDF性能下降。

[0133] 进一步,选取一段运动轨迹,图6通过实验仿真了真实轨迹、采用XGBoost算法的定位轨迹以及提出的XGBoost-EC定位轨迹的定位误差比较,图中表明,整个定XGBoost-EC定位轨迹较XGBoost定位估计更接近于真实轨迹,定位精度较高。

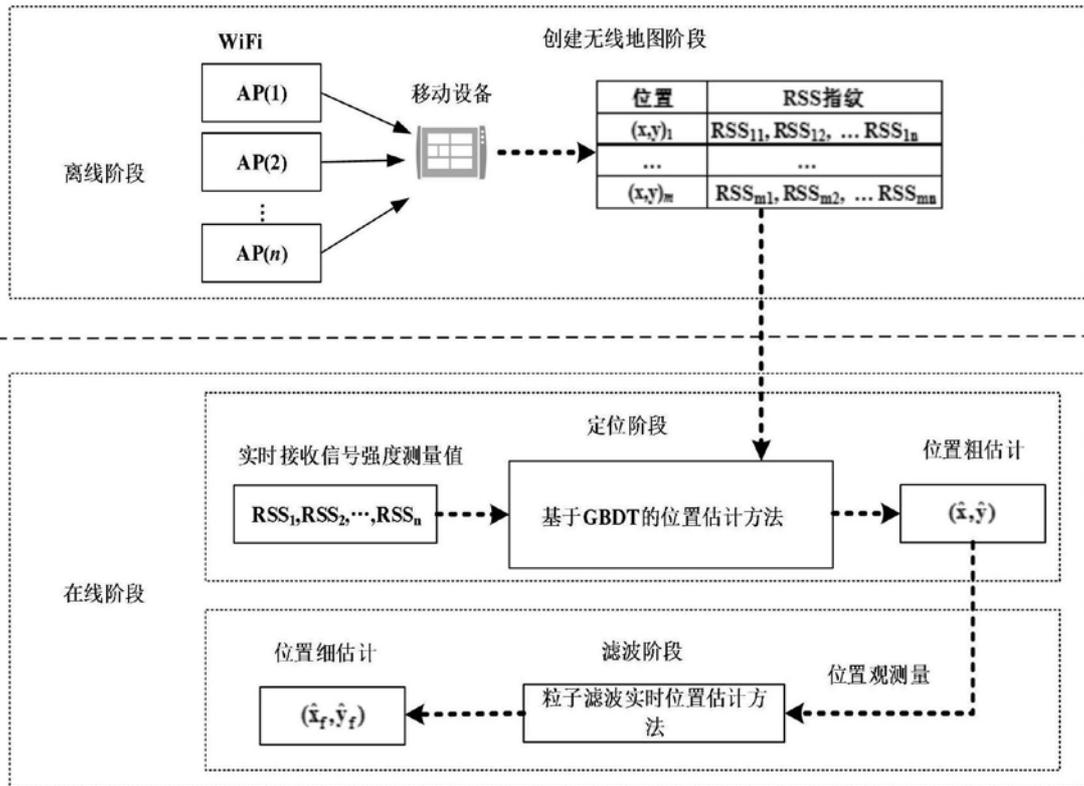


图1

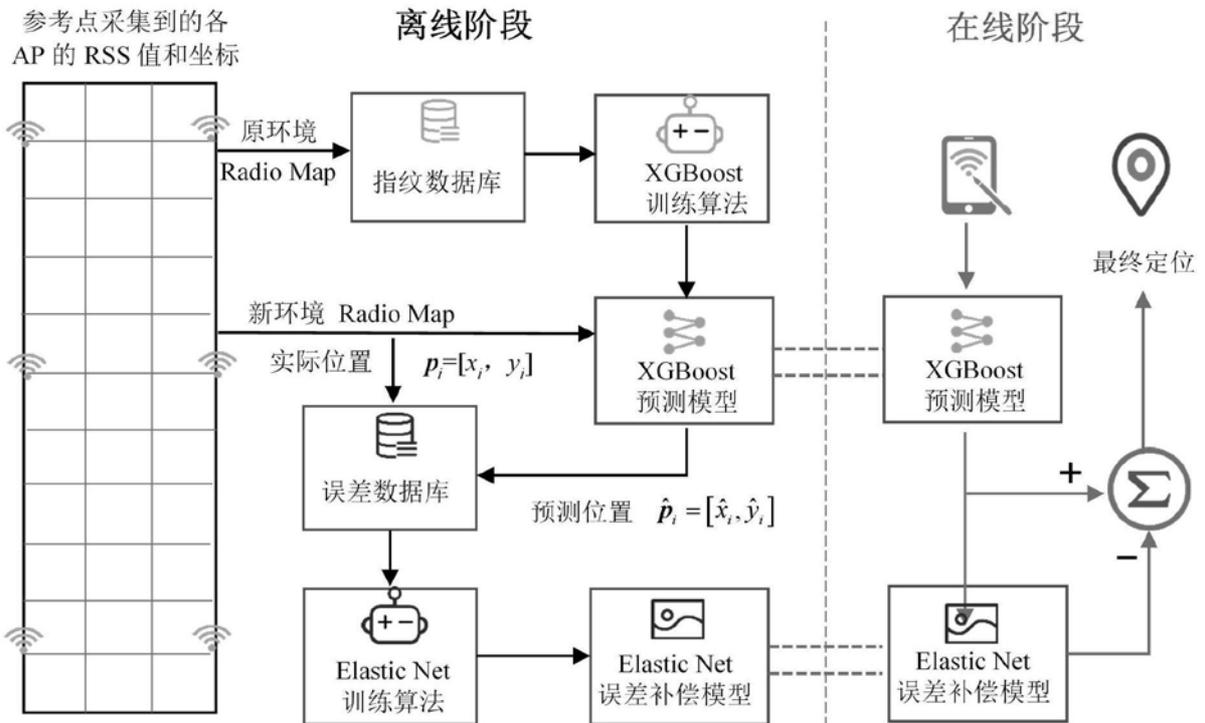


图2

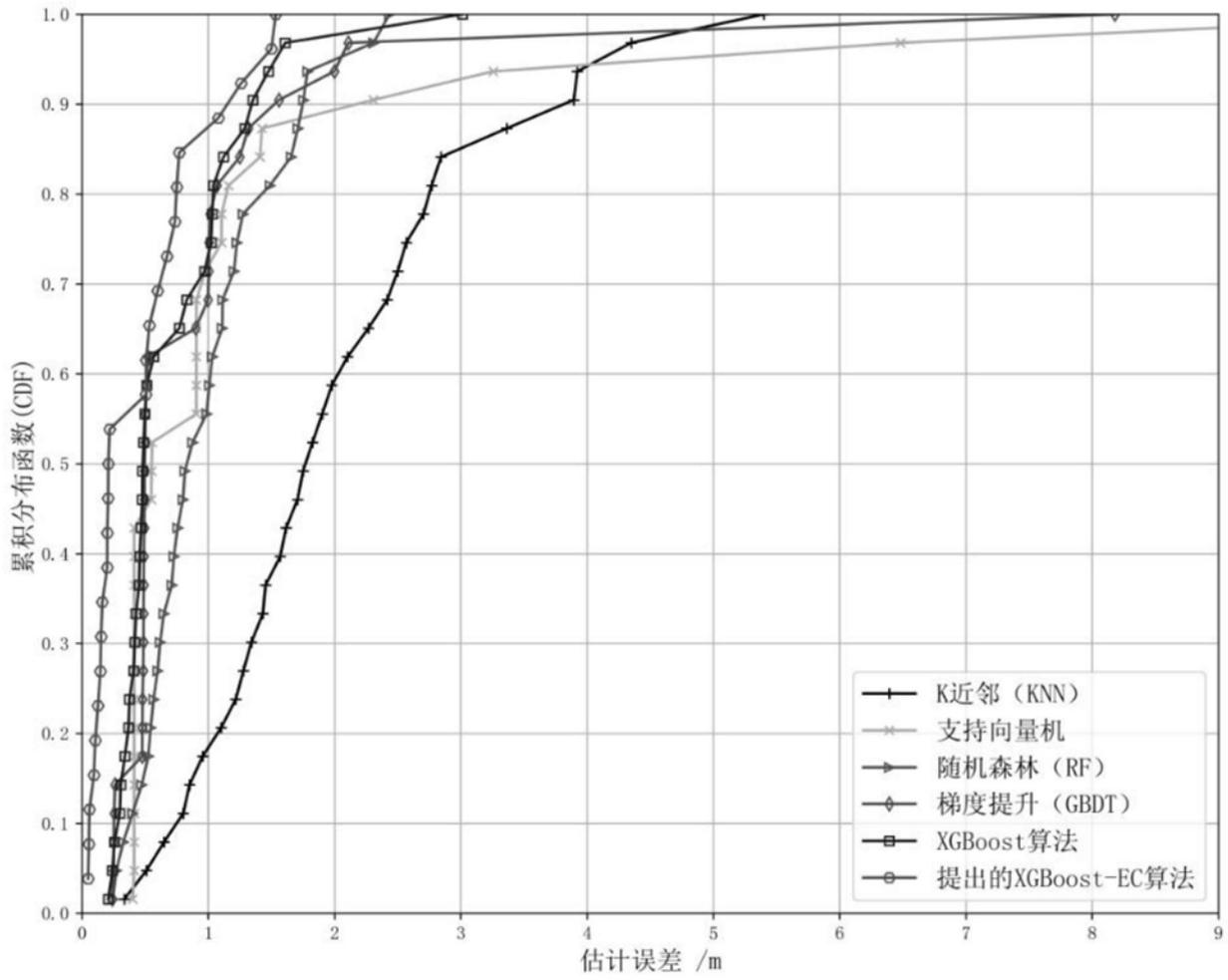


图3

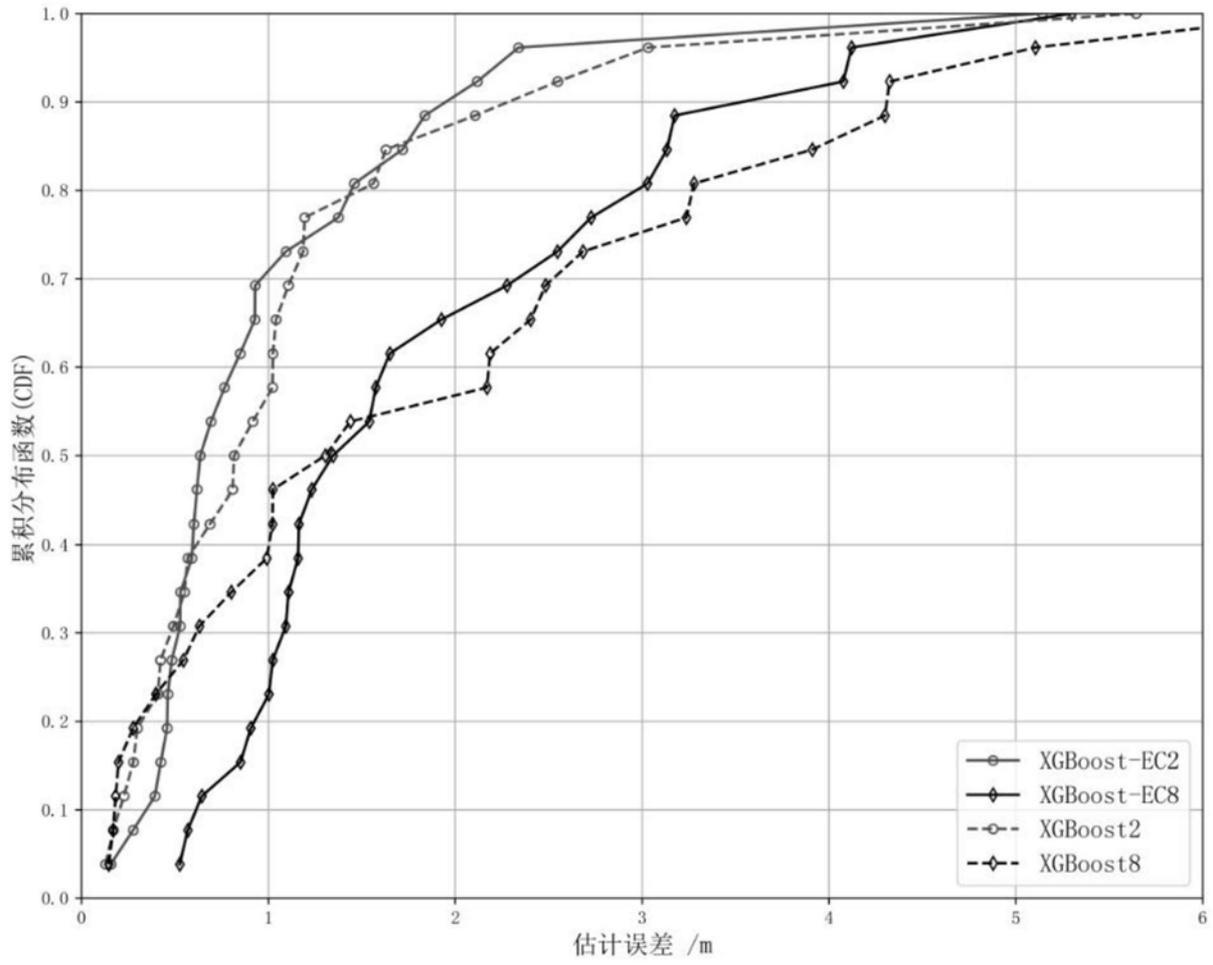


图4

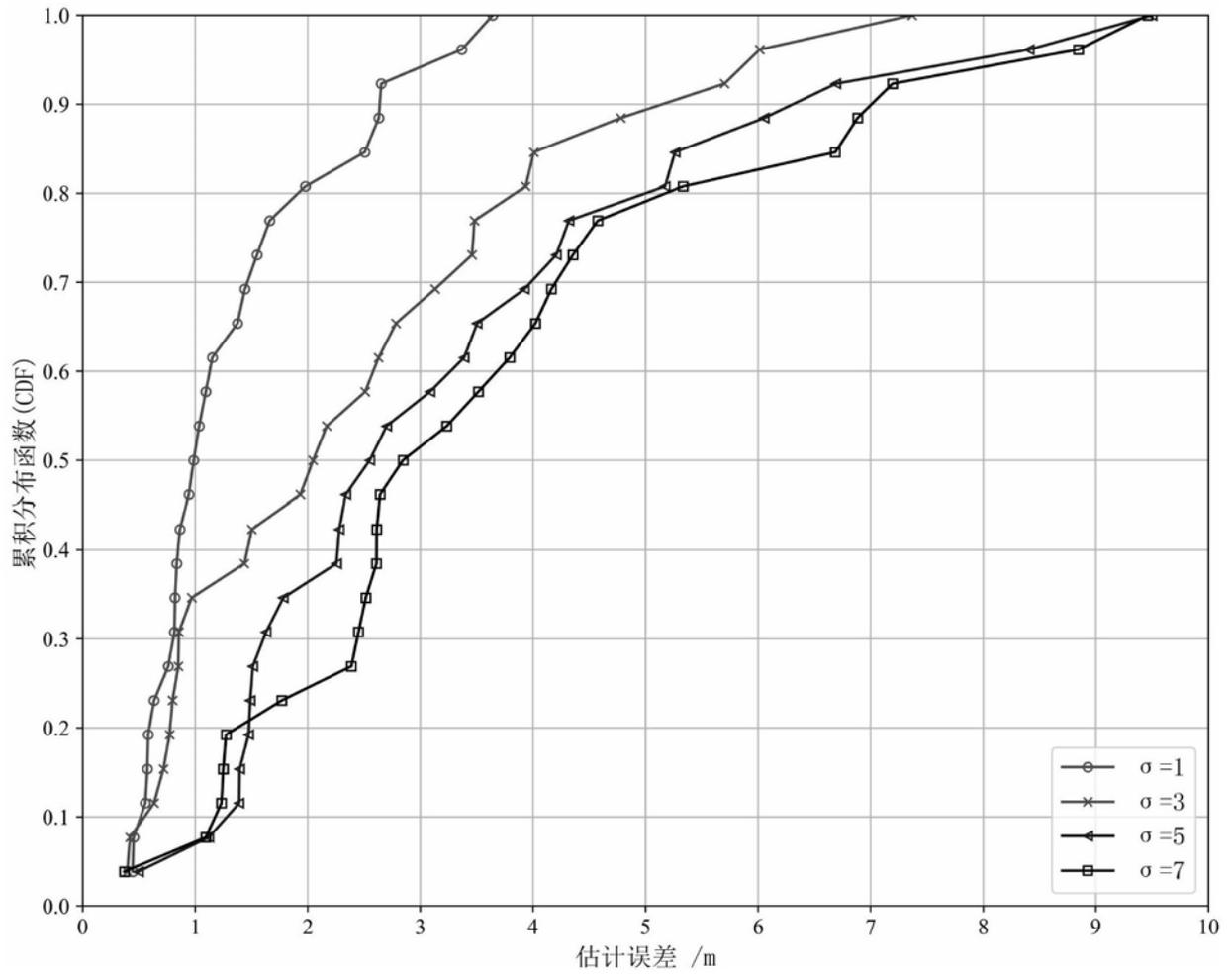


图5

