

DOI:10.11918/202005081

一种面向室内环境变动的人员目标无源定位算法

刘嘉伟^{1,2},毛文字^{1,2},鲁华祥^{1,2,3,4}

(1. 中国科学院大学,北京 100089; 2. 中国科学院 半导体研究所,北京 100083; 3. 中国科学院 脑科学与智能技术卓越创新中心,上海 200031; 4. 半导体神经网络智能感知与计算技术北京市重点实验室(中国科学院),北京 100083)

摘要: 现有的基于接收信号强度(RSS)的人员目标无源室内定位算法在定位环境变动的情况下难以兼顾人工工作量、时间消耗和定位准确率。针对这个问题,本文提出了基于迁移聚类和坐标融合的变分自编码器(FusVAE)的室内环境变动下人员目标无源定位算法。在环境变动后,采集少量无标签RSS样本,然后使用本文提出的基于度量学习的半监督模糊C均值聚类(SFCMML)对其进行精确聚类和标签标注,对原有的定位模型进行重训练,只需很小的人工和时间代价就可以使原定位模型在新环境下也具有较高的定位准确率。同时,针对变动后环境下采集RSS样本较少的问题,本文提出了基于坐标融合的变分自编码器(FusVAE),对新环境下的RSS样本进行数据增强,丰富了RSS样本的数量和质量,提高了定位模型的泛化能力。实验结果表明,在环境变动的情况下,本文提出的算法的平均定位准确率可达88.6%,和同领域同类型算法相比,具有较高的定位精度和较好的环境变动适应性,更适用于变动环境下的人员目标无源室内定位问题。

关键词: 无源室内定位;RSS;机器学习;迁移学习;变分自编码器

中图分类号: TP391 **文献标志码:** A **文章编号:** 0367-6234(2021)08-0039-10

Device-free indoor localization algorithm for changing environment

LIU Jiawei^{1,2}, MAO Wenyu^{1,2}, LU Huaxiang^{1,2,3,4}

(1. University of Chinese Academy of Sciences, Beijing 100089, China; 2. Institute of Semiconductors, Chinese Academy of Sciences, Beijing 100083, China; 3. Center for Excellence in Brain Science and Intelligence Technology, Chinese Academy of Sciences, Shanghai 200031, China; 4. Semiconductor Neural Network Intelligent Perception and Computing Technology Beijing Key Lab (Chinese Academy of Sciences), Beijing 100083, China)

Abstract: The existing human target device-free indoor localization algorithm based on received signal strength (RSS) is difficult to give consideration to artificial workload, time consumption, and positioning accuracy under the circumstance of environment changes. In view of this problem, this paper proposes a device-free localization algorithm based on transfer clustering and fusion variational auto-encoder (FusVAE) in changing indoor environment. After environment changes, a small amount of RSS samples without labels were collected. Then, a semi-supervised fuzzy C-means clustering based on metric learning (SFCMML) was proposed to accurately cluster and label the samples, and the original model was retrained, where only a small amount of artificial work and process time was required to achieve a high localization accuracy for the original model in the new environment. In addition, aiming at the problem that the RSS samples collected in the new environment were in small quantity, the FusVAE was constructed based on coordinate fusion to generate RSS samples in the new environment for data enhancement, which could enrich the quantity and quality of the RSS samples, improve the generalization ability of the model, and enhance the positioning accuracy. Experimental results show that under the circumstance of environment changes, the average positioning accuracy of the proposed algorithm reached 88.6%. Compared with the algorithms of the same type in the same field, the proposed algorithm had higher positioning accuracy and better environmental adaptability, which is more applicable to device-free indoor localization in changing environment.

Keywords: device-free localization;RSS;machine learning;transfer learning;VAE

近些年来,基于接收信号强度(RSS)的无源室内定位技术因其良好的便利性和用户隐私性,得到

了广泛的研究。该方法一般分为两种,一种是基于损耗模型的方法,通过非线性函数拟合目标位置与相关链路RSS变化关系实现人员目标的定位。Wang等^[1]建立了目标位置的观察模型,通过粒子滤波来定位人员目标。Kaltioikallio等^[2]建立了基于信号衰落特征的RTI模型,提出了辨别并滤除非敏感链路的算法从而提高了定位精度。Yang等^[3]综

收稿日期: 2020-05-08
基金项目: 国家自然科学基金(61701473,U19A2080,U1936106)
作者简介: 刘嘉伟(1995—),男,硕士研究生;
鲁华祥(1965—),男,研究员,博士生导师
通信作者: 毛文字,maowenyu@semi.ac.cn

合了路径损耗、天线高度等参数对接收功率的影响,提出了适用于室内走廊环境的路径损耗定位模型。这些方法通过利用固定模型和固定特征来进行普通室内环境下的人员目标无源定位,并取得了良好的定位精度。但当室内环境发生变化时,链路属性与目标位置的对应关系也会发生变化,致使该方法的定位准确率下降。另一种是基于机器学习的方法,通过离线采集数据并训练模型进行室内定位。Youssef 等^[4]和 Zhang 等^[5]分别利用指纹匹配进行了无源室内定位的初步研究。Wagner 等^[6]将检测区域以不同坐标划分并利用人工神经网络模型对目标进行定位。Wang 等^[7]提出一种曲线拟合和位置搜索的室内定位方法,利用穷举搜索和梯度下降搜索来找出距离误差最小的位置。Mao 等^[8]通过双向小波滤波对 RSS 数据进行优化并利用集成学习模型实现目标在复杂环境下的定位。Fazelinia 等^[9]使用一种基于 ROC 曲线的特征选择技术将 RSS 数据降维,然后利用优化的贝叶斯决策模型进行定位。

在实际场景中,室内布局和物品摆放的位置会经常发生变化,造成 RSS 链路模型和样本特征的改变,从而导致原定位模型不再适用于变动后的环境。对此,许多研究人员也展开了研究。Chen 等^[10]提出一种动态调整室内定位算法,减少了环境变动和功率变化对定位系统的影响。Guo 等^[11]提出多指纹分类器模型,提高了指纹和分类器之间的互补性

从而增加了模型环境适应性。Manh 等^[12]提出一种多维核密度估计方法,该方法对信号噪声和动态环境都具有较强的鲁棒性。这些方法虽然取得了一定进展,但是很难同时兼顾人工工作量、时间消耗、计算复杂度和定位准确率之间的平衡。

本文基于迁移学习的思想和数据增强的手段,提出了基于迁移聚类 and FusVAE 的室内环境变动后的人员目标无源定位算法,研究主要解决的问题如图 1 所示。在环境变动后,采集少量变动后环境的无标签 RSS 样本,使用本文提出了基于度量学习的半监督模糊 C 均值聚类(SFCMML)对其进行精确聚类和标签标注,并对原定位模型进行重训练,得到能够适应变动后环境的定位模型。以较小的时间、工作量和计算代价,提高了变动后环境下的定位准确率、提升了定位精度。同时,为了解决变动后环境下 RSS 样本较少,模型训练不充分的问题,本文利用生成模型对 RSS 样本进行数据增强,改善重训练过程。为了使生成的 RSS 新样本更符合变动后环境下的样本特征和分布,定位过程更适应变动后的环境,对变分自编码器(VAE)进行了改进,提出了基于坐标融合的变分自编码器(FusVAE),增加了训练数据的数量和质量,提高了模型的泛化能力。实验结果表明,在环境变动的情况下,本文提出的算法的平均定位准确率达到了 88.6%,与同领域同类型算法相比具有较高的定位精度和较好的环境变动适应性。

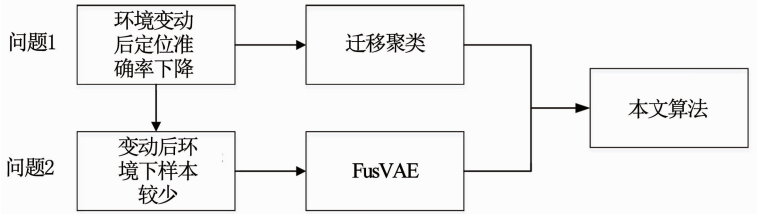


图 1 本文研究解决的问题

Fig. 1 The research problem in the study

1 基于迁移聚类和 FusVAE 的人员目标无源室内定位算法

本文提出的基于迁移聚类和 FusVAE 的室内环境变动下人员目标无源定位算法的整体架构见图 2。先在原环境中通过逐个坐标点站定采集的方法获得原环境下的已标签 RSS 数据,使用该样本集训练随机森林定位模型。在环境改变后,被定位目标在定位区域内快速遍历各个坐标点,并在被定位目标移动的过程中采集 RSS 数据获得新环境下的无标签 RSS 样本,这种采集方法消耗时间较短,且

不需要进行繁琐的标签标注工作。然后利用本文提出的基于度量学习的半监督模糊 C 均值聚类(SFCMML)算法对这些新样本进行聚类 and 标注标签,生成定位模型重训练所需的新环境下带标签的样本。用处理后的样本进行重训练,原定位模型只需进行微调,便可以很好地适应环境变动后人员目标的定位过程,且具有较高的定位精度。在重训练的过程中,使用基于坐标融合的变分自编码器(FusVAE)算法,利用原环境和新环境下的先验知识,生成新环境下的高质量 RSS 样本进行数据增强,进一步提高了定位准确率。

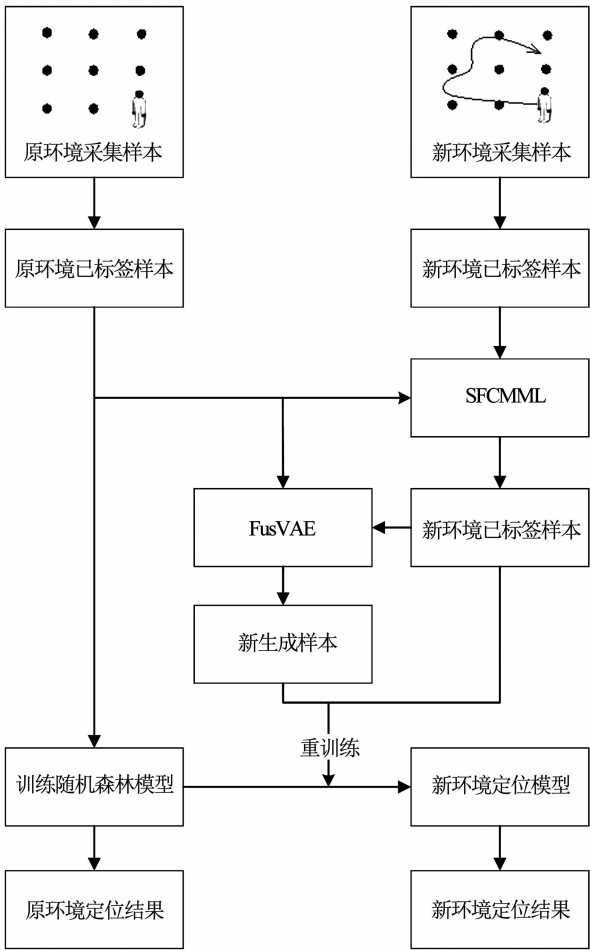


图 2 基于迁移聚类的人员目标无源室内定位算法的整体架构
Fig. 2 Overall architecture of device-free indoor localization based on transfer clustering

1.1 基于度量学习的半监督模糊 C 均值聚类 (SFCMML)

当环境改变后,在改变后的环境中采集少量的数据,和原环境每个坐标站定采集 RSS 数据不同,被定位目标只需在环境中随机走动,遍历每个坐标点即可完成 RSS 数据的采集。与原环境中繁琐的数据采集相比,可以节约大量样本采集时间,但是这样采集的数据为一个无标签样本集,无法直接用于定位模型的重训练。为获得新环境下的已标签样本集,需要对其进行聚类及标签标注。传统的聚类算法没有将原环境和变动后环境数据之间的联系运用到聚类的过程中,利用原环境的先验知识对变动后环境的聚类进行指导,导致 RSS 数据聚类效果较差。针对这个问题,本文提出了基于度量学习的半监督模糊 C 均值聚类算法 (semi-supervised fuzzy C -means clustering based on metric learning, SFCMML), 算法架构见图 3。该算法通过系数矩阵 A_i 在原环境中学习出不同坐标点样本每个维度不同的权重,使用 A_i 修改新环境坐标点在聚类过程中使用的距离函数,使新环境下每类样本在聚类过程中都能遵循

原环境坐标点数据的结构,实现了坐标点样本的精确聚类,相比于传统的聚类算法,该算法将原环境下的先验知识运用到了新环境的聚类中,提高了聚类和标签标注的准确率,从而使重训练后的定位模型能够更好地适用于变动后的环境,提高了定位准确率。接下来对 SFCMML 算法的原理进行详细介绍。

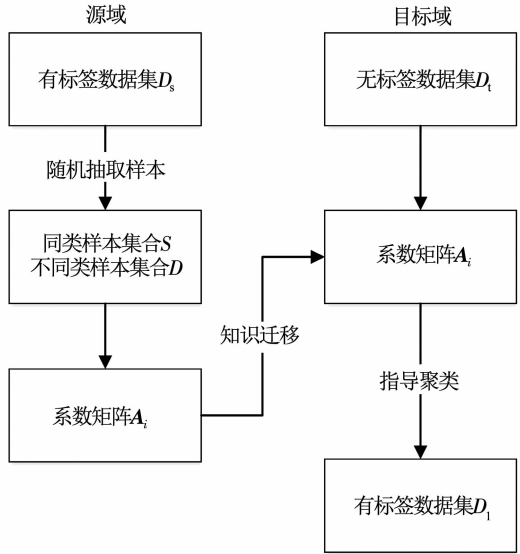


图 3 SFCMML 算法整体架构
Fig. 3 Overall architecture of SFCMML

D_s 是被定位目标在原环境中各坐标点上站定采集得到的样本集, D_s 中的数据为已标签样本, 样本的标签为目标所站的坐标点的标号。 D_t 是被定位目标在新环境中随机走动采集得到的样本集, D_t 中的数据为未标签样本。对于 D_s 和 D_t , 在数据分布上存在一个假设, 即在同一环境下, 不同坐标点的数据分布完全不同, 在不同环境下, 相同坐标点的分布虽然不同, 但是有一定的相似性。所以只要找到原环境和新环境数据分布的潜在关系, 就可以建立联系, 将原环境的知识迁移到新环境中, 对 D_t 中的样本进行指导聚类。

本文首先引入成对约束, 成对约束可以用于改变底层的距离度量, 对聚类过程进行引导, 达到更好的聚类效果。从 D_s 中的每类样本中随机选取若干样本组成样本对, $S \subset D_s \times D_s$ 是相似样本对集合, 集合中每个元素为一个样本对 (X_p, X_q) , 其中 X_p 和 X_q 都为同一类的样本。 $D \subset D_s \times D_s$ 是不相似样本对集合, 集合中元素为一个样本对 (X_p, X_q) , 其中 X_p 和 X_q 都为不同类的样本。这样就构建了成对约束, 对于同类样本, 我们希望样本间的距离越小越好, 对于不同类样本, 我们希望样本间的距离越大越好, 所以希望可以找到一个功能函数 f 满足

$$\begin{aligned} \min f(X_p, X_q) \quad (X_p, X_q) \in S \\ \max f(X_p, X_q) \quad (X_p, X_q) \in D \end{aligned} \quad (1)$$

对于每一类坐标点样本,每个维度代表了每条链路的 RSS 值,由于网络中存在易受影响、RSS 值变化较大的活跃链路和基本不会受到影响、RSS 值基本不变的非活跃链路,所以每个坐标点样本的维度的权重不同,都有自己的数据分布特征。本文使用对称正定矩阵 A_i 作为系数矩阵对欧氏距离进行修改,修改后的距离度量函数如式(2)所示,这样就可以在计算距离函数的时候使样本每个维度的权重都不同,表示出该类样本的数据结构特征。所以只要求解出原环境数据集 D_s 每类样本的 A_i ,就可以将其数据分布的先验知识通过系数矩阵 A_i 进行知识迁移,在 D_t 聚类过程中起到指导作用。

$$f_{A_i}(X_p, X_q) = \|X_p - X_q\|_{A_i} = \left[(X_p, X_q)^T A_i (X_p, X_q) \right]^{\frac{1}{2}} \quad (2)$$

使用修改后的距离度量函数可以将式(2)转化为下式

$$\min_{(X_p, X_q) \in S} \|X_p - X_q\|_{A_i}^2$$

$$\sum_{(X_p, X_q) \in D} \|X_p - X_q\|_{A_i} \geq t, A_i \geq 0 \quad (3)$$

如果 t 是一个正的对角矩阵,可以将式(3)中的最优化问题转换为式(4),通过求解式(4)的最小值就可以求得系数矩阵 A_i 。

$$g(A_i) = \sum_{(X_p, X_q) \in S} \|X_p - X_q\|_{A_i}^2 - \log \left[\sum_{(X_p, X_q) \in D} \|X_p - X_q\|_{A_i} \right] \quad (4)$$

求解出原环境下每类坐标点样本的系数矩阵 A_i 后,就可以使用能够反映出原环境各类样本分布的距离函数对 D_t 进行聚类,这样的距离度量对于每个坐标点都使用的不同度量,使得数据的聚类可以是不同的形状,同类样本间的距离更小,从而提高聚类的准确率, SFCMML 的目标函数

$$J = \sum_{i=1}^k \sum_{j=1}^n f_{A_i}(x_j, c_i) U_{ij}^m \quad (5)$$

式中: k 为坐标点的数目; n 为 D_t 中样本的数量; c_i 为第 i 类的类中心; U 为隶属度矩阵; u_{ij} 为第 j 个样本属于第 i 类的隶属度; m 为控制算法柔性的幂指数参数,取值范围为 $[1, \infty)$, 如果 m 过大,则聚类效果会很差,如果 m 设置为 1,则算法会接近硬聚类算法 HCM,在许多聚类问题中, m 的取值一般为 $[1, 3]$,而在本文的聚类问题中,经过充分的实验证明, m 取值为 2 效果最好。

c_i 和 u_{ij} 计算公式分别如式(6)和式(7)所示:

$$c_i = \frac{\sum_{j=1}^n u_{ij} x_j}{\sum_{j=1}^n u_{ij}}, i = 1, 2, \dots, k \quad (6)$$

$$u_{ij} = \frac{1}{\sum_{q=1}^k \left(\frac{\|c_i - x_j\|_{A_i}^2}{\|c_q - x_j\|_{A_i}^2} \right)^{\frac{1}{m-1}}}$$

$$i = 1, 2, \dots, k; j = 1, 2, \dots, n \quad (7)$$

1.2 基于坐标融合的变分自编码器 (FusVAE)

在迁移聚类过程中,为提高定位效率,在变动后环境下采集的 RSS 样本数量较少, D_t 的样本数量远小于 D_s 的样本数量,导致采集的数据并不能完全反映出每个坐标点数据的概率分布,如果直接对原定位置模型进行重训练,虽然可以提高定位准确率,但是提升不大,还有很大的改进空间。所以,可以使用生成模型生成变动后环境的样本进行数据增强。对于目前主流的生成模型,对抗生成网络 (generative adversarial networks, GAN) 调参过程复杂,生成的样本多样性差,本文引入数据增强方法的目的是提高研究和定位效率,如果使用 GAN 生成模型,会生成大量无效或重复样本,样本多样性不高,不适合本文的应用场景。流模型的生成函数需要精心设计,求解过程复杂,在本文研究的定位过程中,室内环境多变,每次环境变动后都需要重新设计生成函数,增加了定位过程的计算复杂度,降低了定位效率。而变分自编码器 (variational auto-encoder, VAE) 不需要调试过多参数,计算复杂度低,且生成样本质量较高,所以选用 VAE 作为生成模型生成新环境下的样本进行数据增强。但是普通的 VAE 不能很好地与变动后环境下的定位问题结合起来,变动后的数据既有新环境的数据分布特点,又保留了一部分原环境中的数据分布特点,普通的 VAE 只考虑了新环境对生成样本的影响,完全舍弃了原环境的知识,缺少了原环境中先验知识的指导和限制,从而导致了生成样本质量较低。所以本文提出了一种基于坐标融合的变分自编码器 (FusVAE), 算法架构见图 4。

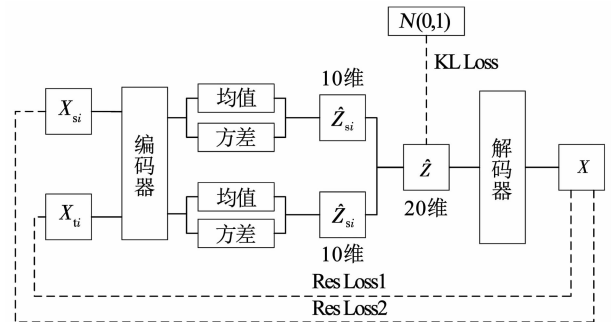


图 4 FusVAE 的算法架构

Fig. 4 Overall architecture of FusVAE

如图 4 所示, 本文将两个不同环境下的坐标 i 样本 X_{si} 和 X_{ti} 作为编码器的输入, 通过编码器将输入数据编码成一个高斯分布的隐变量, 得到两个分

布 $\hat{Z}_{si}$ 和 $\hat{Z}_{ti}$, 分别表示了 X_{si} 和 X_{ti} 的隐变量分布, 其中 $\hat{Z}_{si}$ 和 $\hat{Z}_{ti}$ 的维数都为 10 维。由于 $\hat{Z}_{si}$ 和 $\hat{Z}_{ti}$ 在隐空间中代表不同样本的数据信息, 所以可以将两个分布拼接起来组成新的隐变量分布, 记为 $\hat{Z}$, 维数为 20 维, 可以表示为

$$\hat{Z} = [\hat{Z}_{s1}, \hat{Z}_{s2}, \dots, \hat{Z}_{s10}, \hat{Z}_{t1}, \hat{Z}_{t2}, \dots, \hat{Z}_{t10}] \quad (8)$$

隐变量分布 $\hat{Z}$ 既保留了原环境下的数据分布信息, 又具有新环境下的数据分布信息, 使通过解码器解码后生成的新样本可以同时保留两个环境下的数据分布特征, 更接近环境变动后的真实 RSS 样本。

FusVAE 的损失函数分为 3 部分, 第 1 部分是 VAE 原有的损失函数, 由于篇幅限制, 对于 VAE 的原理本文不再赘述, 可详见参考文献 [13-14]。该损失函数是为了使后验分布的近似分布 $q_{\varphi}(z|x(i))$ 趋近于标准正态分布 $N(0, 1)$, 保证模型生成能力的 KL 散度, 可以表示为

$$D_{KL}(N(\mu, \sigma^2) \| N(0, 1)) = \frac{1}{2}(-\log \sigma^2 + \mu^2 + \sigma^2 - 1) \quad (9)$$

式中第 2 部分和第 3 部分是本文为了使生成的新样本更接近 X_{si} 和 X_{ti} , 对原损失函数进行的改进。

引入重构损失函数作为约束条件, 将其添加到

FusVAE 的损失函数中, 其中与 X_{ti} 的重构损失函数记为 R_1 , 计算公式为

$$R_1 = E_{X_{ti}}[\|X_{ti} - X\|] \quad (10)$$

与 X_{si} 的重构损失函数记为 R_2 , 计算公式如式(11)所示

$$R_2 = E_{X_{si}}[\|X_{si} - X\|] \quad (11)$$

所以 FusVAE 完整的损失函数如式(12)所示

$$L = D_{KL}(q_{\varphi}(Z|X) \| N(0, 1)) + \lambda R_1 + (1 - \lambda) R_2 \quad (12)$$

式中 λ 为一个超参数, 取值范围为 $[0, 1]$, 可以用来调节两个重构损失函数的权重。如果我们希望生成的新样本更接近 X_{ti} 中的样本, 可以将 λ 设置地比较大, 如果希望生成的新样本更接近 X_{si} 中的样本, 可以将 λ 设置地较小。

和传统的 VAE 相比, 本文提出的 FusVAE 可控性更高, 原因在于, 不仅在隐变量空间中利用了原环境和新环境中的先验知识, 还可以通过超参数 λ 控制新样本与两个环境中样本的相似程度, 这样生成的样本在分布上可以更接近我们希望接近的坐标点样本, 将其用于模型微调, 可以进一步提高定位准确率。

本文算法的具体步骤和流程见图 5。

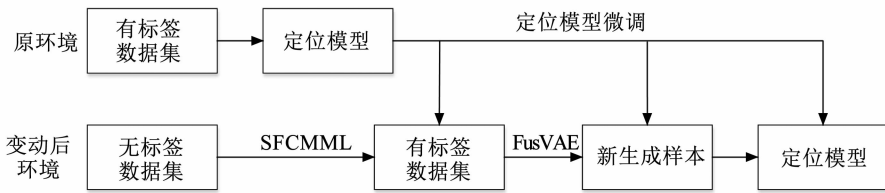


图 5 本文算法流程图

Fig. 5 Flow chart of the proposed algorithm

2 实验设计与结果分析

2.1 实验设置

为验证本文提出算法的有效性与优越性, 将办公室环境和实验室环境作为定位区域并分别设置实验。两个环境的面积分别为 30 m^2 和 4 相邻坐标点距离为 0.7 m 。将 6 个射频节点如图 5 所示布置在定位区域周围。环境实拍图和平面图分别见图 6、7, 被定位的单人员目标在每个坐标点上站定一段时间采集 RSS 数据, 每个坐标点的采集时间为 120 s 。射频节点使用的芯片是 CC2530, 工作频率在 2.45 GHz , 使用 IEEE 802.15.4 通信协议。

为了体现环境的动态性, 设置 5 组实验依次移动环境中的物品, 最开始移动椅子、书籍等对环境变

化影响较小的物品, 再逐渐移动显示器、实验器材等对环境变化影响较大的物品, 呈现出环境变动由小到大的效果。在每次环境变动后, 被定位的单人员目标在环境中随机走动 30 s , 完成 RSS 数据的采集。办公室环境下的环境变动图见图 8, 同时在实验室环境下做出类似的 5 组实验。

2.2 结果评价标准

为了验证本文算法的有效性, 使用定位准确率和 RMSE 作为评价标准。

定位准确率 C 计算公式为

$$C = \frac{x_r}{x_w} \quad (13)$$

式中, x_r 为坐标点正确分类的数目, x_w 为定位过程中所有定位结果的总数目, 这是一种直观体现分类准确率高低的方法。



图 6 环境实拍

Fig. 6 Picture of the environment

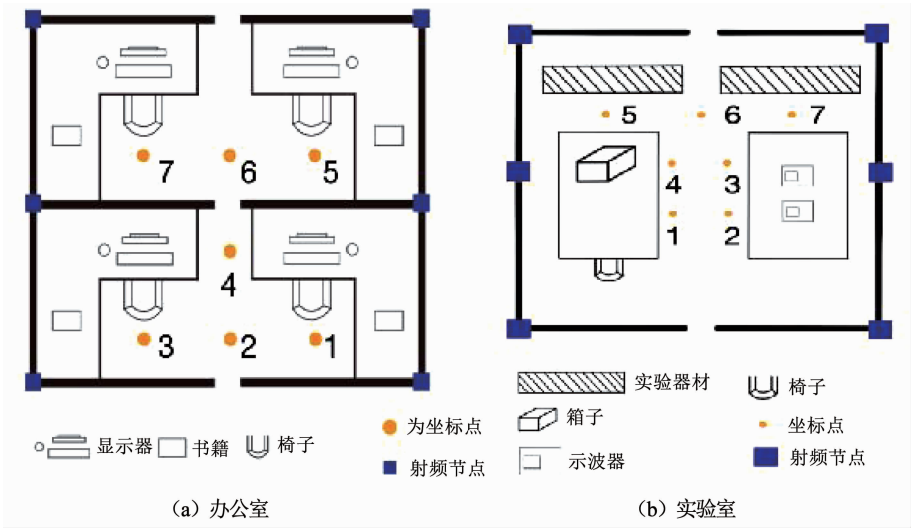


图 7 环境平面

Fig. 7 Plan of the environment

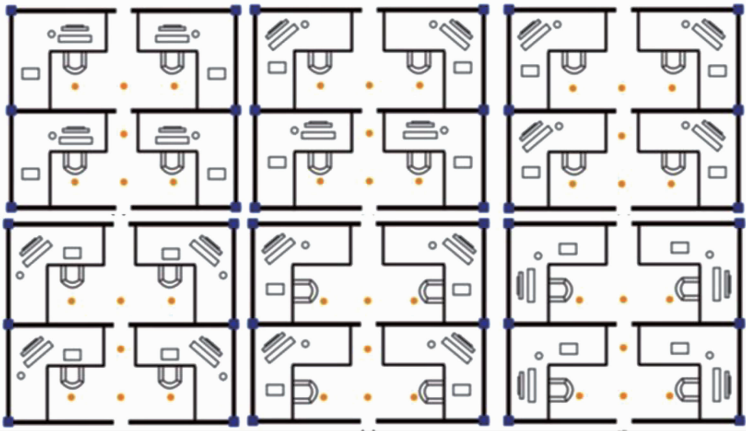


图 8 环境变动

Fig. 8 Change in the environment

RMSE 计算公式如式(14)所示

$$R_{\text{RMSE}} = \sqrt{\frac{\sum_{c=1}^N (P_c - R_c)^2}{N}} \quad (14)$$

式中: P_c 为预测的坐标点, R_c 为目标所处的真实的坐标点, $|P_c - R_c|$ 为两个坐标点间的物理距离。当一个坐标点错分时,错分为相邻坐标点和距离很远的

坐标点对定位精度的影响是不同的,所以 RMSE 可以在空间上体现出算法的定位精度。

2.3 实验结果分析

2.3.1 不同聚类算法的定位准确率对比

为了体现本文提出的 SFCMML 算法聚类的优越性,和目前主要使用的 DBSCAN^[15]及模糊 C 均值聚类(FCM)^[16]算法进行对比,将几种算法都应用到

本文的算法中。同时为了体现本文算法的必要性, 体现出迁移学习对定位准确率的提升, 将环境改变后不做任何处理的定位准确率和 RMSE 一起进行对比, 定位准确率对比见图 9, RMSE 对比结果见表 1。

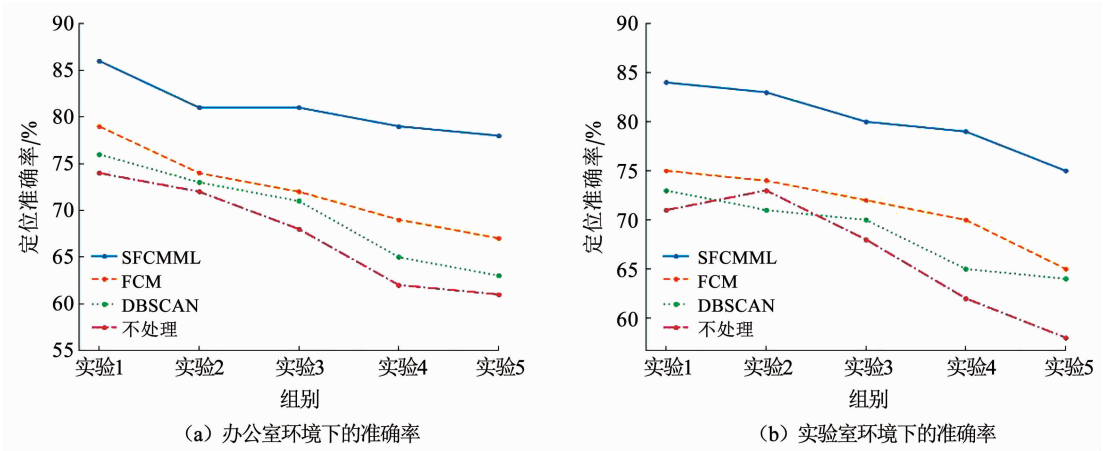


图 9 不同聚类算法定位准确率对比

Fig. 9 Comparison of positioning accuracy of different clustering algorithms

表 1 不同聚类算法的 RMSE

Tab. 1 RMSE for different clustering algorithms								m
组别	不处理		DBSCAN		FCM		SFCMML	
	办公室	实验室	办公室	实验室	办公室	实验室	办公室	实验室
实验 1	0.66	0.72	0.57	0.69	0.45	0.60	0.35	0.39
实验 2	0.73	0.65	0.67	0.67	0.61	0.61	0.42	0.40
实验 3	0.84	0.91	0.77	0.79	0.69	0.65	0.42	0.43
实验 4	1.22	1.30	1.02	1.05	0.85	0.86	0.45	0.44
实验 5	1.31	1.45	1.11	1.06	0.88	0.97	0.47	0.50

通过实验结果可以得出结论, 所提基于迁移聚类的人员目标无源室内定位算法可以提升变动环境后的定位准确率, 但是如果在聚类过程中选择效果较差的算法, 不仅提升效果不明显, 在有些情况下甚至会使准确率下降。所以在迁移学习过程中要选择本文提出的 SFCMML 聚类算法, 聚类准确率高, 定位准确率改善效果明显, 准确率的提升基本都在 10% 以上, 最高准确率提升高达 17%, 定位的误差都控制在 0.5 m 以内。

2.3.2 FusVAE 对定位准确率的影响

FusVAE 通过超参数 λ 控制生成的新样本与原环境样本或新环境样本的相似程度。为了选取最优的 λ , 选择任意坐标点在原环境下和新环境下的样本集作为 FusVAE 的输入, 调节 λ 的值生成不同的样本集, 使用 PCA 降维, 将 30 维的数据映射到三维空间中, 观察数据分布, 见图 10。

图中深色点为新环境下坐标点的样本集, 浅色点为新生成的样本, 通过对比可以发现当 $\lambda = 0.8$ 时, 生成的 RSS 样本在分布上与原样本更接近、更

紧凑、质量更高, 从后续的定位准确率也可以证明将 λ 设为 0.8 时生成样本质量最高, 数据增强效果最好。所以将 λ 的值设置为 0.8, 生成新环境下的 RSS 样本对定位模型进行重训练。

为了检验 FusVAE 生成的样本能够起到数据增强, 提升定位准确率的效果, 并且体现 FusVAE 在数据增强时相较于普通 VAE 有更好的效果。将 λ 的值设置为 0.8, 在 5 组实验中分别不使用数据增强、使用 VAE 进行数据增强、使用 FusVAE 进行数据增强, 然后对比定位准确率, 定位准确率对比结果见图 11。

实验结果表明通过 VAE 对基于迁移聚类的无源室内定位进行数据增强, 可以提高最终的定位准确率。但是本文提出的 FusVAE 因为充分利用了原环境和新环境中的先验知识对样本生成过程进行约束和控制, 生成的新样本更接近新环境下的样本数据分布, 对定位准确率的提升效果更明显, 对定位准确率的平均提升约 8%, 比 VAE 的提升率高约 4%。

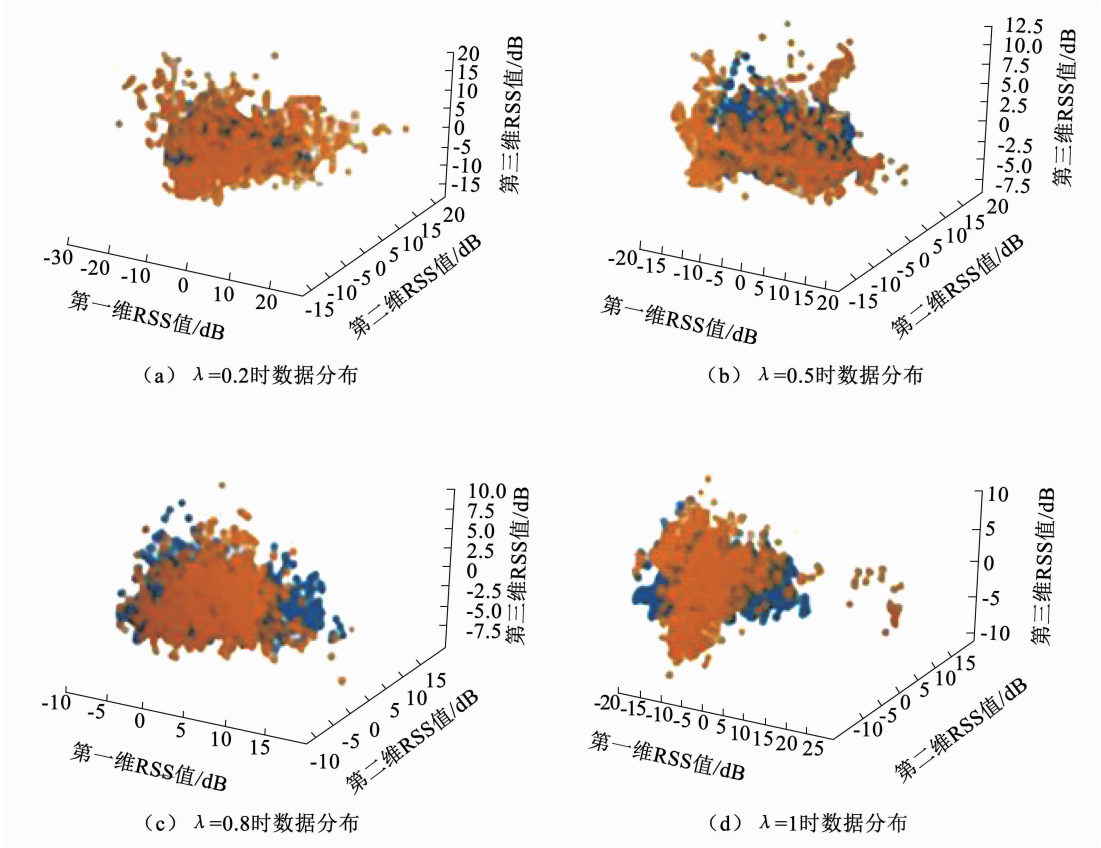


图 10 新生成样本集与原样本集分布散点图

Fig. 10 Distribution scatter diagram of the generated sample set and the original sample set

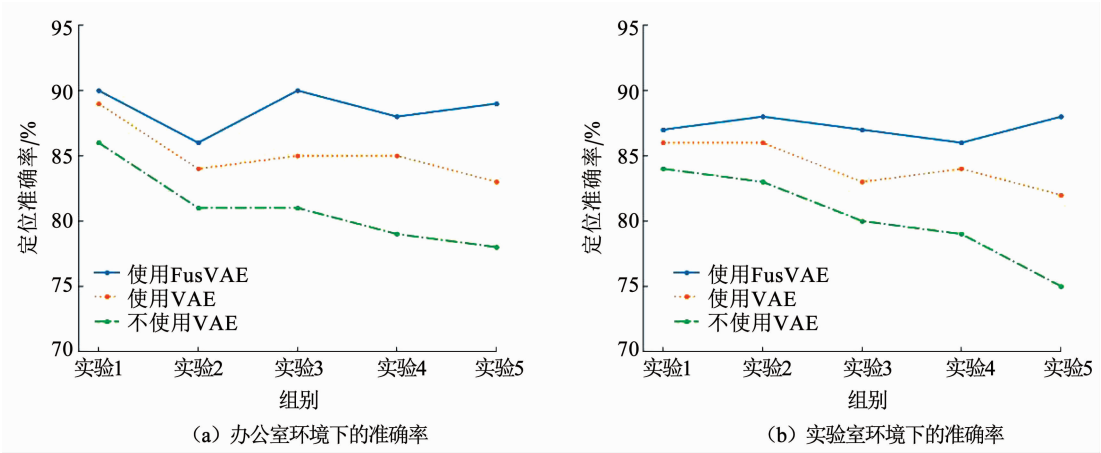


图 11 VAE 和 FusVAE 定位准确率对比

Fig. 11 Comparison of positioning accuracy of VAE and FusVAE

2.3.3 本文算法与其他变动后环境下的定位算法对比

为了体现本文提出的变动后环境下的基于迁移聚类 和 FusVAE 的人员目标无源室内定位算法的优越性。与其他变动环境下室内定位算法进行对比，包括 KAAL^[11]、ILSITL^[17]、RA^[12]。ILSITL 是一种基于实例迁移的室内定位算法，将新环境下采集的数据直接和原环境采集的数据对原定位模型进行调整。KAAL 的方法通过训练多个指纹分类器，利用指纹和多分类器之间的互补性提高动态环境下的定

位准确率。RA 是一种基于区域定位的定位方法，建立的 不是 RSS 值与坐标点之间的联系，而是通过 RSS 值与坐标区域 的联系来降低环境变化为定位准确率的影响。准确率对比见图 12，RMSE 结果见表 2。

通过对比可以发现本文算法在每次环境变动的定位准确率都很高，平均定位准确率为 88.6%，平均 RMSE 为 0.312 m，高于其他算法的定位准确率。而且环境变化时，定位准确率基本不发生变化，极大程度地降低了环境变动对于定位准确率的影响，说明了本文算法的有效性和相比于其他算法的优越性。

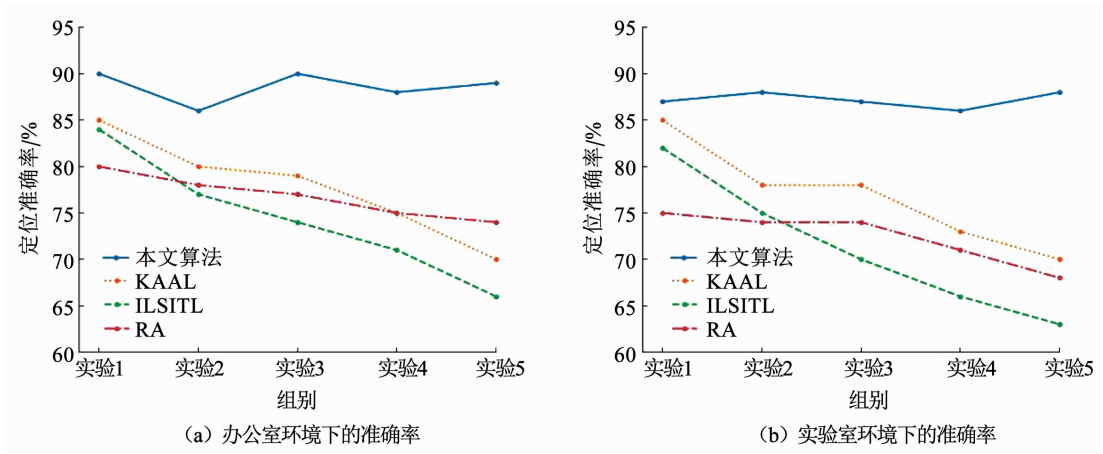


图 12 不同算法定位准确率的对比

Fig. 12 Comparison of positioning accuracy of different algorithms

表 2 不同定位算法的 RMSE

Tab. 2 RMSE for different localization algorithms

组别	RA		KAAL		ILSITL		本文算法	
	办公室	实验室	办公室	实验室	办公室	实验室	办公室	实验室
实验 1	0.52	0.66	0.40	0.37	0.46	0.41	0.31	0.34
实验 2	0.55	0.63	0.51	0.61	0.58	0.67	0.35	0.32
实验 3	0.61	0.65	0.47	0.63	0.61	0.72	0.30	0.32
实验 4	0.69	0.69	0.55	0.71	0.67	0.91	0.31	0.35
实验 5	0.69	0.75	0.71	0.72	0.88	0.93	0.29	0.33

2.3.4 整体定位结果展示

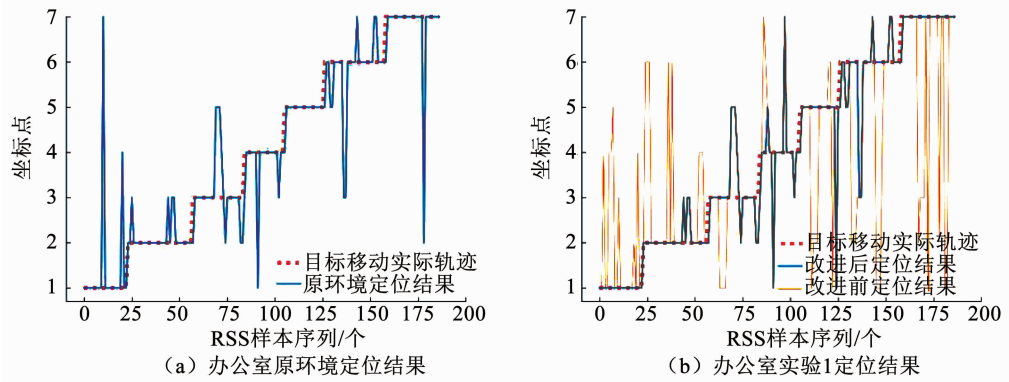
为了直观地体现本文算法对环境变动后定位准确率的改善,设计实验让被定位目标按照坐标点1→坐标点 2→坐标点 3→坐标点 4→坐标点 5→坐标点 6→坐标点 7 的顺序遍历坐标点,同时采集 RSS 数据作为测试集用于检验定位的效果。在办公室环境和实验室环境,对比本文算法的定位结果、未改进前的定位结果和目标实际移动路径,见图 13。

3 结 论

1)提出了一种将迁移聚类和 FusVAE 结合使用

的数据增强型人员目标无源室内定位方法,通过迁移聚类对新环境中的无标签数据进行指导聚类 and 标注标签,解决了环境变动后定位模型的重训练问题,然后使用 FusVAE 对新环境 RSS 样本进行补充,实现了模型重训练过程的数据增强,有效提高了定位算法的变动环境适用性和定位准确率,实验结果表明,该方法在办公室环境下的平均定位准确率达 88.6% ;

2)提出了基于度量学习的半监督模糊 C 均值聚类(SFCMML)算法,根据原环境中的先验知识修改聚类过程中的距离函数,赋予坐标点样本每个维



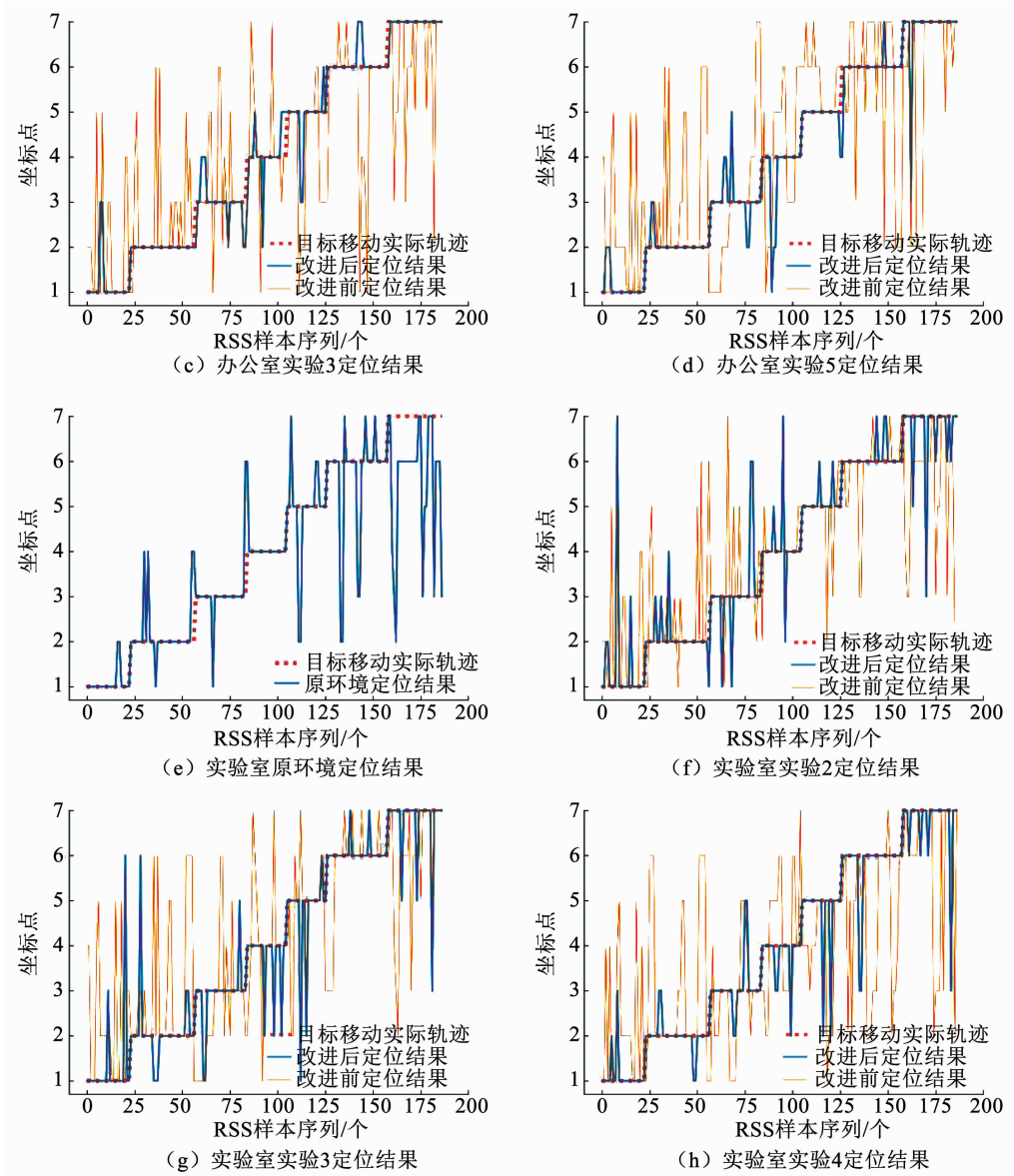


图 13 定位结果对比

Fig. 13 Comparison of positioning results

度不同的权重,有效提高了聚类准确率,实验结果表明,该算法对变动后环境的定位准确率提升明显,在办公室环境下准确率的提升最高可达 16% ;

3) 提出了基于坐标融合的变分自编码器 (FusVAE) 算法,利用原环境和新环境下的先验知识,改进了隐变量结构并在损失函数中添加了生成样本与原环境样本和新环境样本间的重构误差,增强了生成的 RSS 数据的样本质量,进一步提高了模型的泛化能力。综合实验结果表明,该算法可以进一步提升环境变动后的定位准确率,在办公室环境下准确率的提升最高可达 13% 。

参考文献

[1] WANG Jie, GAO Qinghua, YU Yan, et al. Robust device-free wireless localization based on differential RSS measurements[J]. IEEE Transactions on Industrial Electronics, 2013, 60(12): 5943.

DOI:10.1109/TIE.2012.2228145
[2] KALTIOKALLIO O, BOCCA M, PATWARI N. A fade level-based spatial model for radio tomographic imaging[J]. IEEE Transactions on Mobile Computing, 2014, 13(6): 1159. DOI: 10.1109/TMC.2013.158
[3] YANG Jing, ZHANG Heng, LI Zetao, et al. Propagation characteristics in corridor scenario about typical radio wave [C]//2017 9th International Conference on Advanced Infocomm Technology (ICAIT). Chengdu: IEEE, 2017: 5. DOI: 10.1109/ICAIT.2017.8388818
[4] SEIFELDIN M, SAEED A, KOSBA A, et al. Nuzzer: a large-scale device-free passive localization system for wireless environments[J]. IEEE Transactions on Mobile Computing, 2013, 12(7): 1321. DOI: 10.1109/TMC.2012.106
[5] ZHANG Dian, LIU Yunhuai, LIONEL M. Link-centric probabilistic coverage model for transceiver-free object detection in wireless networks [C]//2010 IEEE 30th International Conference on Distributed Computing Systems. Genova: IEEE, 2010: 116. DOI: 10.1109/ICDCS.2010.47 (下转第 124 页)