

基于深度学习的无线电作弊信号发现与定位系统

李竞铭¹ 赵明杰² 翟 会¹ 丁国如³ 张小飞¹

(1. 南京航空航天大学电子信息工程学院, 南京, 210016; 2. 浙江理工大学信息学院, 杭州, 310018; 3. 陆军工程大学通信工程学院, 南京, 210007)

摘 要: 随着无线电技术的飞速发展和无线电作弊手段的日新月异, 考试中利用无线电设备进行作弊的情况也逐年增加。为保证考试的公平性, 如何有效地发现和定位无线电作弊信号、建设智慧考场已成为当前迫切需要解决的新课题。鉴于此, 在室内定位和频谱监测技术的基础上, 本文设计了基于深度学习的无线电作弊信号发现与定位系统, 系统实现了无线电作弊设备判决、定位、告警、以及移动终端实时显示等功能, 为监考人员提供了直观、远程和实时的考场环境的安全情况, 为广大考生创造了一个公平的竞争环境。

关键词: 无线电作弊信号; 智慧考场; 室内定位; 频谱监测; 深度学习

中图分类号: TP391 **文献标志码:** A

Discovery and Location System of Radio Cheating Signal Based on Deep Learning

Li Jingming¹, Zhao Mingjie², Zhai Hui¹, Ding Guoru³, Zhang Xiaofei¹

(1. College of Electronic and Information Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing, 210016, China; 2. School of Information Science and Technology, Zhejiang Sci-Tech University, Hangzhou, 310018, China; 3. College of Communications Engineering, PLA Army Engineering University, Nanjing, 210007, China)

Abstract: With the rapid development of radio technology and the rapid change of radio cheating means, the use of radio equipment in the examination of cheating has also increased. In order to ensure the fairness of the examination, how to find and locate the radio cheating signal effectively and construct the intelligent examination room has become a new hot topic that needs to be solved urgently. In view of this, on the basis of indoor location and spectrum monitoring technology, a discovery and location system of radio cheating signal based on deep learning is designed in this paper, which realizes the functions of judgment, location, alarm for the radio cheating signal, and real-time displaying on the mobile terminals. The system provides an intuitive, remote and real-time examination environment for invigilators, and creates a fair competition environment for the examinees.

Key words: radio cheating signal; intelligent examination environment; indoor location; spectrum monitoring; deep learning

引 言

考试是教学活动的重要环节,是选拔优秀人才、检验教学质量和考察学生理解程度的主要手段。近年来,为彰显国家对考试公平性的重视,考试作弊首次纳入刑法。考试作弊所造成的危害极大,它不仅破坏了考试公平,侵犯了其他学生的利益,损害了学校的声誉,更降低了高等教育的公信力,因此成为高校教学管理研究的热点问题。随着无线电通信的广泛应用,考试作弊的手段也趋于隐性化、智能化,作弊者通常使用磁共振耳机、骨感耳机和针孔摄像头等先进的作弊器材与外界进行试题及答案的传递。由于便于隐藏、不需要人为取出且具有强大的收发功能,无线电作弊设备已成为大型考试中的作弊利器。目前常见的用来考试防作弊的措施大致有以下几种:(1)人脸识别指纹验证;(2)使用金属探测仪进行安检;(3)使用屏蔽设备;(4)视频监控等。

室内定位是一种对室内环境的位置估计技术^[1,2]。由于受到复杂的室内电波传播环境、室内建筑布局 and 人员移动性等因素的影响,室内信号衰落模型无法准确建立,因此其发展远滞后于室外定位技术。通常室内定位技术通过采集收发机之间无线链路的到达时间(Time of arrival, TOA)、到达时间差(Time difference of arrival, TDOA)、到达角(Angle of arrival, AOA)^[2]和接收信号强度(Receive signal strength indicator, RSSI)^[3]来估计收发机之间的距离,以此来构建位置估计模型,从而获取终端的位置信息。但是基于 TOA、TDOA 和 AOA 的定位技术需要进行相对时间和角度的测量,这大大增加了硬件成本,并不适用于广泛普及,因此本文采用 RSSI 来构建位置估计模型。

指纹识别^[4,5]方法专注于有效地比较实时的频谱扫描信息与室内预先记录的扫描数据库,因此对于本地信号干扰更为稳健。文献[6]采用最近邻插值法(K-nearest neighbor, KNN)方法,从指纹数据库中查询与其最相似的 K 个位置数据,对其位置进行平均。文献[7]对其进行了改进,使用了加权 KNN 方法。文献[8]中,随机森林算法也得到了很好的定位精度。然而,传统的机器学习方法对数据敏感度较高,往往需要对数据进行归一化、滤波等预处理,且容易过拟合导致泛化能力较弱,随时间推移网络的预测性能会受到影响。但随着定位精度的需求越来越高,数据量也随之增大,深度学习是一个很有前途的解决方案^[9]。深度学习^[10]可以直接学习提取数据隐含的结构特征而不需要预处理,并且可以有效避免过拟合等问题,从而提高网络泛化能力。因此,使用深度学习与自动编码器,将数据进行特征提取和降噪处理。文献[11]中,就利用深度学习算法通过 WiFi 实现了室内定位;文献[12]提出了一种对信道状态信息(Channel state information, CSI)进行深度学习的室内指纹系统,验证了深度学习在室内定位领域的广阔前景。

为了构建智慧考场,本文在室内定位和频谱监测技术的基础上,设计了基于深度学习的无线电作弊信号发现与定位系统,系统实现了无线电作弊设备判决、定位、告警以及移动终端实时显示等功能。本文的主要贡献为:(1)构建了基于通用软件无线电外设(Universal software radio peripheral, USRP)平台的无线电作弊信号发现与定位系统,实现了告警、移动终端实时显示等功能;(2)设计了基于深度学习的定位算法,实现了对电磁频谱数据的训练和学习,使平均定位精度达到 0.178 8 m。

1 系统总体设计

为使考场工作人员方便、直观、远程和实时地了解所监测环境的安全情况,无线电作弊信号发现与定位系统利用物联网的思想,在互联网基础上,将网络延伸到现实环境中,进行频谱和视频信息的采集与共享,实现对频谱监测信息的呈现,并且在网页中为用户提供了无线电作弊信号的定位与告警服务,并结合视频信息混合判决。如图 1 所示,整个系统架构包含电磁频谱数据采集、数据处理中心和网页端实时显示 3 个部分,分别完成对应的工作。

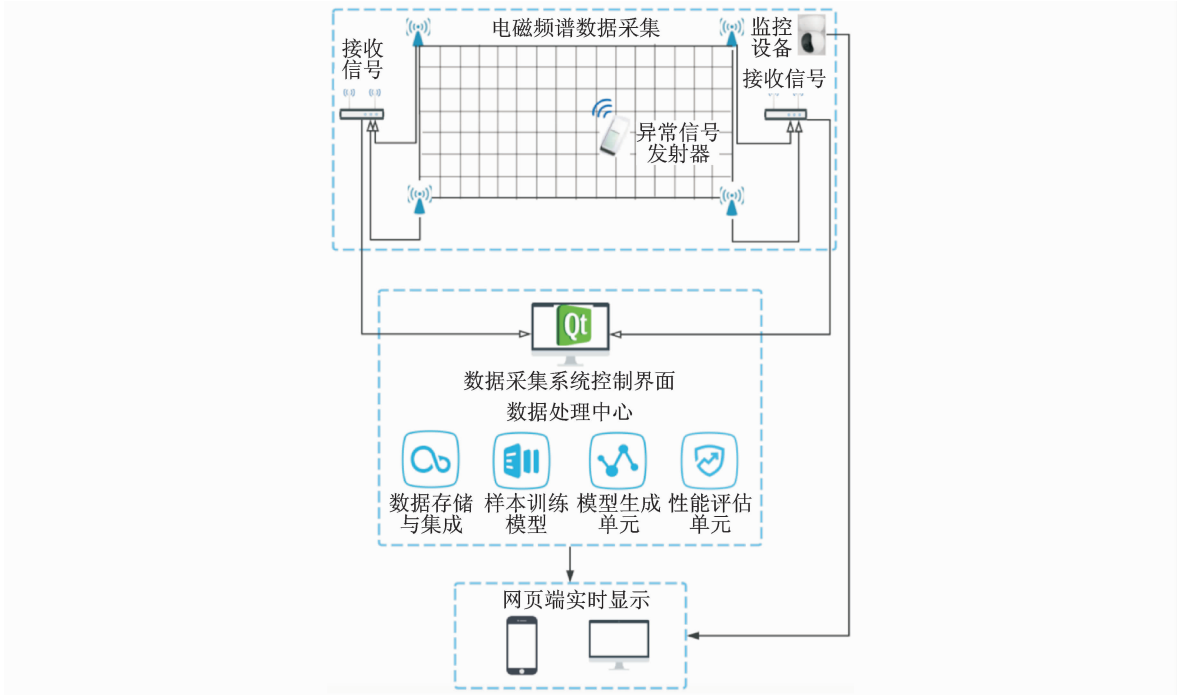


图 1 系统构架图

Fig. 1 System architecture diagram

1.1 电磁频谱数据采集

为了解当前检测的室内环境的电磁辐射情况,利用频谱监测设备 USRP 和 4 根天线测量并收集当前环境的电磁能量值。考虑到无线电作弊信号工作周期及工作频段,将采样率设为 1 MHz,采样点数 5 000,监测中心频率设为 2.09 MHz。为了有助于数据的训练和学习,将室内环境进行网格化分割,网格大小为 20 cm×20 cm,取 7 行 20 列共计 140 个测试点,开启发射源,反复在这 140 个测试点上进行测试,采集 4 根天线测得的频谱数据。

系统采样频谱图如图 2 所示,图中横坐标表示频率,纵坐标表示幅度。4 幅图分别对应 4 根天线采集到的数据。当频谱图出现如图 3 所示的情况时,说明当前检测环境有信号发出,即有存在无线电作弊信号的可能。

同时在监测环境中布设视频监控设备,用于发现无线电作弊信号时,可以实时传输视频数据到电脑上,便于监考人员查看该考生有无作弊行为并留存证据。

1.2 数据处理与评估

将采集到的数据进行存储和集成,为使测得的样本数据更加精确,对样本数据利用窄带滤波器的方法,有效提高了信噪比。为了对无线电作弊设备进行精确定位,应用基于深度学习的室内定位算法,继而利用样本数据对模型进行训练,得到模型后便可通过数值匹配的方法确定小范围的精确位置。最后,对定位性能进行评估。本文将精确度和精密度作为评估标准,即精确度越高,定位效果越好,同等精确度时,精密度越高越好。

(1) 精确度

精确度是定位系统的关键评估指标,通常用平均定位精度来衡量,即估计位置 and 实际位置的均方根

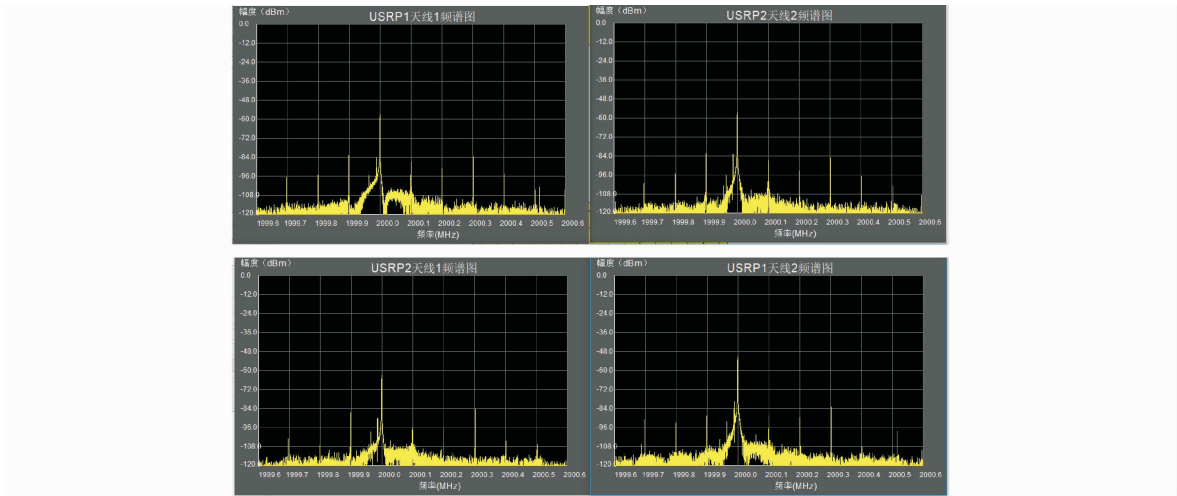


图 2 系统时域采样结果

Fig. 2 System sampling results on time domain

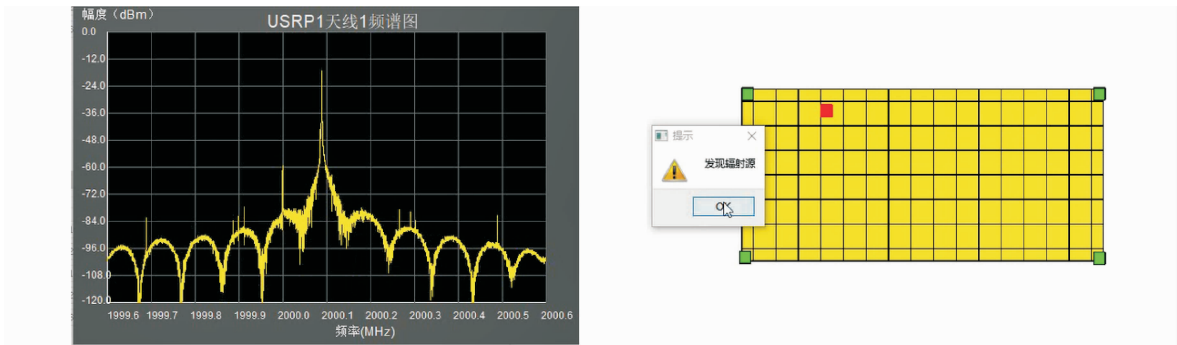


图 3 作弊信号的发现

Fig. 3 Discovery of cheating signal

误差(Root mean squared error, RMSE),即

$$\text{RMSE}=\sqrt{\frac{\sum_{i=1}^n\left\|\hat{\theta}_i-\theta\right\|^2}{n}}$$

(1)

式中: $\hat{\theta}_i$ 表示终端 n 次估计中的第 i 次的估计位置, θ 表示实际位置。

(2)精密度

与精确度不同,精密度考虑的是系统所获取的精确度的分布概率,是对精确度鲁棒性的度量。一般情况下,精密度可由几何精度因子(Geometric dilution of precision, GDOP)或者累计概率分布(Cumulative probability functions, CPF)表示。本文使用累计概率分布表示精密度,在同等精确度时,当定位精确度在 0.2 m 范围内,精密度越高越好。

1.3 异构数据综合判决

当发现异常信号信息时,软件界面和网页终端适时发出告警,上报给终端用户,终端用户可立刻通过查看考场监控视频,来综合判决当前环境和告警位置上有无作弊的隐患,并且通过视频回放功能留存

作弊证据。

2 系统模型

深度学习基本模型类似于传统神经网络的 3 层结构^[13]，是由输入层、隐藏层和输出层组成的多层网络。深度学习通过逐层训练方法使得梯度能够有效传播，因此深度神经网络通常含有多隐藏层，提高了网络表达能力。如图 4 所示，本系统所采用模型主要分为以下 3 个部分：自编码器、多层感知机、深度学习回归预测模型。

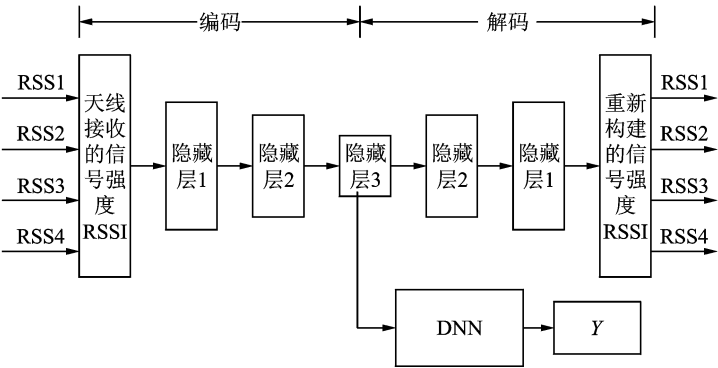


图 4 系统模型图
Fig. 4 System model diagram

2.1 自编码器

自编码器 (AutoEncoder)^[14] 可以用于压缩数据，提取特征或者降噪。利用深度学习的基本思想和网络的特性直接构建一个能够表达原始数据的网络结构，通过设定输入输出近似相等，学习网络的参数权值，进而构建自编码模型。文献[15]中提出了一种新颖的基于去噪自编码器的室内定位方法，验证了自编码器可以提高定位精度。

2.2 多层感知机

多层感知机 (Multilayer perceptron, MLP)，又称深度前反馈网络，是典型的深度学习模型。MLP 具有多层节点，其中每层与下一层完全连接，隐藏层的每个节点都使用非线性激活函数进行操作，并且利用反向传播模型训练网络。

(1) 激活函数

Sigmoid 函数：即 S 型函数，在神经网络中被广泛采用。Sigmoid 函数可以将实数压缩到 $[0, 1]$ 区间，具有强大的解释力，但在神经元趋于 0 或者 1 的时候会产生饱和导致梯度弥散，因此权值需要小心初始化，其公式为

$$f(x) = \frac{1}{1 + e^x}$$

(2)

Tanh 函数：该函数拥有良好的数据控制力，其将实数映射到 $[-1, 1]$ 区间，但依旧存在饱和问题。Tanh 函数公式为

$$f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$$

(3)

Relu 函数：是线性修正单元，当 $x < 0$ 时为 0，当 $x > 0$ 时斜率为 1。Relu 收敛速度更快，但是 Relu 也更加脆弱，大的梯度流可能导致神经元永久失效，选择合适的学习率或者层间批正则化可以避免这种情况的出现。其公式为

$$f(x)=\max(0,x)$$

(4)

(2)反向传播算法

训练样本集为 $\{(x^{(1)},y^{(1)}),\cdots,(x^{(m)},y^{(m)})\}$,其中 m 为样本的个数,用该样本集训练神经网络。实验中的损失函数为

$$J(W,b;x,y)=\frac{1}{2}\parallel h_{w,b}(x)-y\parallel^2$$

(5)

梯度下降法的关键步骤是计算偏导,对于网络的权值 W 和偏置项 b ,其迭代公式分别为

$$W_{ij}^{(1)}=W_{ij}^{(1)}-\alpha\frac{\partial}{\partial W_{ij}^{(1)}}J(W,b)$$

(6)

$$b_i^{(1)}=b_i^{(1)}-\alpha\frac{\partial}{\partial b_i^{(1)}}J(W,b)$$

(7)

其中 α 为学习速率。

2.3 深度学习回归预测模型

采用室内指纹定位方案只能对离散指纹点进行预测估计,为了实现更高精度的连续预测定位,利用深度学习搭建一个回归预测模型。该线性回归模型可表示为

$$f(x)=w^Tx+b$$

(8)

式中: x 表示输入, w 表示权重, b 表示偏差。 w 和 b 训练成最小化的目标函数,该模型首先处理输入数据,然后执行预训练;当到达输出层时,模型会反向传播;算法达到收敛时停止。文献[16]采用了回归预测模型来对基因数据进行学习。

3 仿真结果与分析

3.1 深度学习分类任务

室内定位场景通常采取基于指纹的定位方式。对室内环境的140个指纹点进行采集学习,从而估计出新的定位点。基于指纹的定位是一个多分类问题,传统机器学习对于二分类问题可以取得较好的效果,但是随着多分类问题数据维数的增高,机器学习问题变得相当困难。为了更好的泛化,机器学习方法往往会提出很强的针对特定任务的假设。同时,S函数的饱和特性使隐含层数受到限制,因此导致神经网络表达能力的不足。而深度学习的核心思想是假设数据由因素或特征组合,这些通用假设提高了深度学习算法。深度学习旨在克服机器学习中的‘维数灾难’‘梯度弥散’等一系列困难,从而在多分类问题中取得更好的效果。

针对室内定位场景的多分类任务,本文采用深度学习框架搭建一个具有五隐含层的深度神经网络分类器。为了防止梯度弥散,采用Relu作为隐含层激活函数,并利用Xavier初始化权重。对测试集分类效果进行仿真验证的结果如图5所示,其中实心圆为成功定位点,方框为实际测试点,三角为错误预测定位点。测试集样本数量为100,仿真结果显示定位精度可以达到90%的准确率。

3.2 深度学习回归预测模型

(1)单隐含层和三隐含层对比

两个网络均不进行AutoEncoder且隐含层均采用Relu作为激活函数,对测试集进行回归拟合来预测坐标

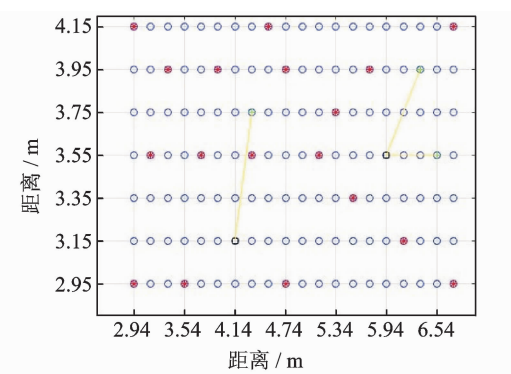


图5 深度学习分类仿真结果
Fig. 5 Simulation results of deep learning classification

点的结果如表 1 所示。从表中可以看出多隐含层预测效果明显优于单层,但是在定位误差上依旧较大。

(2) AutoEncoder

从数据分布可以看出,为了提高定位精度有必要将数据进行特征提取并降噪处理,为此在将数据输入深度学习回归网络之前先进行自编码处理,用于提取数据的高维特征并减小噪声影响。表 2 列出了采用 Tanh 激活函数的三隐含层深度回归预测网络是否进行 AutoEncoder 的定位精度的对比结果。

(3) 不同隐含层激活函数对比

隐含层激活函数的选择也会在很大程度上响应定位的精度。在分类任务中表现良好的 Relu 函数由于其对数据的控制能力较弱,在数据分布不均匀的情况下效果并不如 Sigmoid 等函数好。由于 Tanh 相当于 Sigmoid 函数的扩展形式,其中心是关于零点对称的,所以 Tanh 往往能取得比 Sigmoid 更好的效果,具体结果如表 3 所示。

根据实验结果,采用三隐藏层的深度学习回归网络对室内定位场景进行预测。将训练参数保存后利用实测数据对实验环境定位效果进行仿真验证,结果如图 6 所示。由图 6 可知,通过本系统,利用三隐藏层的深度学习回归网络对室内无线电作弊信号进行定位,其定位的精确度可达到 0.170 07 m,符合考场实际场景对定位精度的需求。

4 结束语

本系统以无线电作弊信号的发现与定位为出发点,通过部署 USRP 频谱监测设备对室内环境中电磁信号进行检测,了解目标区域的电磁安全情况。通过部署视频监控设备对室内环境进行实时监控,并将采集的实时数据收集起来,联网进行处理,运用深度学习算法将平均定位精度提高到 0.170 07 m,系统融合频谱监测设备结果与视频监控设备结果,混合判决是否存在作弊现象。最后结果以网页及 APP 形式直观地展现出来,可使监考和巡考人员可以方便、直观、远程和实时地了解所监测环境的安全情况。

参考文献:

[1] Liu H, Darabi H, Banerjee P, et al. Survey of wireless indoor positioning techniques and systems[J]. *Systems Man and Cybernetics*, 2007, 37(6): 1067-1080.

[2] 张令文, 杨刚. 超宽带室内定位关键技术[J]. *数据采集与处理*, 2013, 28(6): 706-713.

表 1 单隐含层和三隐含层对比结果

Tab. 1 Comparison results between single hidden layer and three hidden layers

隐含层数	平均定位误差/m	误差小于 0.20 m 定位点/%
单隐含层	0.331	19.0
三隐含层	0.263	51.2

表 2 AutoEncoder 定位结果对比

Tab. 2 Comparison positioning results of Autoencoder

AutoEncoder	平均定位精度/m	误差小于 0.20 m 定位点/%
否	0.206 8	58
是	0.178 8	74

表 3 不同隐含层激活函数定位精度对比

Tab. 3 Comparison of positioning accuracy with different hidden layer activation functions

激活函数	平均定位精度/m	误差小于 0.20 m 定位点/%
Relu	0.268 9	36
Sigmoid	0.206 0	57
Tanh	0.178 8	74

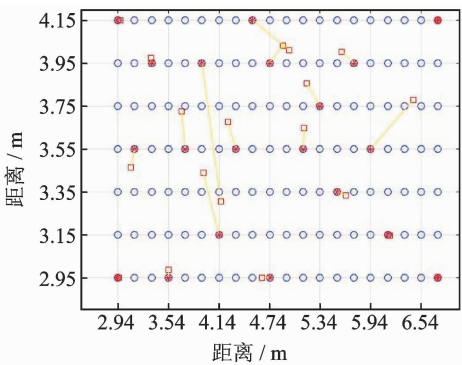


图 6 仿真结果
Fig. 6 Simulation result

Zhang Lingwen, Yang Gang. Ultra-wide-band based indoor positioning technologies[J]. Journal of Data Acquisition and Processing, 2013, 28(6): 706-713.

[3] Chen C, Chen Y, Lai H, et al. High accuracy indoor localization: A WiFi-based approach[C]//IEEE International Conference on Acoustics, Speech and Signal Processing. [S. l.]: IEEE, 2016: 6245-6249.

[4] Sugano M, Kawazoe T, Ohta Y, et al. Indoor localization system using RSSI measurement of wireless sensor network based on ZigBee standard[C]//The IASTED International Multi-Conference on Wireless and Optical Communication. Banff, Canada: [s. n.], 2006: 503-508.

[5] Vasisht D, Kumar S, Katabi D. Decimeter-level localization with a single WiFi access point[C]//Usenix Conference on Networked Systems Design and Implementation. Santa Clara, CA: USENIX Association, 2016: 165-178.

[6] Torrossospedra J, Montoliu R, Martinezuso A, et al. UJIIndoorLoc: A new multi-building and multi-floor database for WLAN fingerprint-based indoor localization problems[C]//International Conference on Indoor Positioning and Indoor Navigation. [S. l.]: IEEE, 2014: 261-270.

[7] Bahl P, Padmanabhan V N. RADAR: An in-building RF-based user location and tracking system[J]//Proc IEEE Infocom, 2000, 2: 775-784.

[8] Khodayari S, Maleki M, Hamed E, et al. A RSS-based fingerprinting method for positioning based on historical data[C]//International Symposium on Performance Evaluation of Computer and Telecommunication Systems. [S. l.]: IEEE, 2010: 306-310.

[9] Elbasiony R, Gomaa W. WiFi localization for mobile robots based on random forests and GPLVM[C]// International Conference on Machine Learning and Applications. [S. l.]: IEEE, 2014: 225-230.

[10] Wang X, Gao L, Mao S. PhaseFi: Phase fingerprinting for indoor localization with a deep learning approach[C]//Global Communications Conference. [S. l.]: IEEE, 2014: 1-6.

[11] Lecun Y, Bengio Y, Hinton G. Deep learning[J]. Nature, 2015, 521(7553): 436-444.

[12] Xie R, Quitadamo A, Cheng J, et al. A predictive model of gene expression using a deep learning framework[C]// IEEE International Conference on Bioinformatics and Biomedicine. [S. l.]: IEEE, 2017: 676-681.

[13] Wang X, Gao L, Mao S, et al. DeepFi: Deep learning for indoor fingerprinting using channel state information[C]//Wireless Communications and Networking Conference. [S. l.]: IEEE, 2015: 1666-1671.

[14] Vincent P, Larochelle H, Bengio Y, et al. Extracting and composing robust features with denoising autoencoders[C]// International Conference on Machine Learning. [S. l.]: ACM, 2008: 1096-1103.

[15] Xiao C, Yang D, Chen Z, et al. 3d ble indoor localization based on denoising autoencoder[J]. IEEE Access, 2017(99): 1.

[16] Nowicki M, Wietrzykowski J. Low-effort place recognition with WiFi fingerprints using deep learning[J]. arXiv: Robotics, 2016: 575-584.

作者简介:



李竞铭 (1993-), 女, 硕士研究生, 研究方向: 认知无线电、频谱大数据, E-mail: lijingmingjlu@163.com.



赵明杰 (1992-), 男, 硕士研究生, 研究方向: 无线网络信息处理技术。



翟会 (1987-), 男, 博士, 研究方向: 阵列信号处理。



丁国如 (1986-), 男, 讲师, 研究方向: 认知无线电、频谱大数据。



张小飞 (1977-), 男, 教授, 研究方向: 移动通信技术、通信信号处理、阵列信号处理。