

Wi-Pest: 一种基于 CSI 的储粮害虫检测方法

单少伟^{1,2}, 杨卫东^{1,2}, 肖乐^{1,2}, 王珂^{1,2}

(1. 河南工业大学信息科学与工程学院, 河南 郑州 450066;

2. 粮食信息处理与控制教育部重点实验室(河南工业大学), 河南 郑州 450066)

摘要: 在粮食储藏过程中, 影响粮食安全的环境和生物因素, 如粮食温度、环境湿度、水分、气体、霉变、害虫等, 均会对粮食储藏安全构成威胁, 其中害虫是威胁粮食储藏安全的一个重要因素。因此, 需要研究一种快速且有效的检测方法, 用于检测储粮害虫。现有的一些方法耗时、设备昂贵、具有潜在的健康危害并且检测效率较低, 提出了一种基于信道状态信息(CSI, channel state information)振幅的非接触、快速、低成本的储粮害虫检测方法, 即无线测害虫(Wi-Pest, wireless-pest)。通过使用CSI振幅数据验证储粮害虫检测的可行性, 在此基础上, 设计了Wi-Pest检测方法。首先对CSI振幅数据进行异常值检测、数据归一化和噪声消除预处理, 然后通过主成分分析(PCA, principal component analysis)方法压缩数据并提取主特征成分, 最后采用随机森林(RF, random forest)分类方法检测储粮害虫。实验结果表明, 所提方法在视距(LOS, line of sight)场景下, 能够检测粮堆活体害虫密度的异常情况, 检测精度平均可以达到97%。

关键词: CSI; 储粮害虫检测; 振幅; 随机森林分类

中图分类号: TN92

文献标识码: A

doi: 10.11959/j.issn.2096-3750.2020.00186

Wi-Pest: a method for detecting stored grain pests based on CSI

SHAN Shaowei^{1,2}, YANG Weidong^{1,2}, XIAO Le^{1,2}, WANG Ke^{1,2}

1. College of Information Science and Engineering, Henan University of Technology, Zhengzhou 450066, China

2. Key Laboratory of Grain Information Processing and Control (Henan University of Technology), Ministry of Education, Zhengzhou 450066, China

Abstract: The environmental and biological factors that affect the food security during the food storage, such as the food temperature, environment humidity, moisture, gas, mildew, pests and others pose a threat to the food storage security, among which the pest is an important factor threatening food storage security. Therefore, a fast and effective detection method is needed to detect stored grain pests. Some of the existing methods are time consuming, using expensive equipment, potentially harmful to health and inefficient. A non-contact, fast and low-cost detection method for stored grain pests based on the amplitude of the channel state information (CSI) was proposed, namely, wireless-pest (Wi-Pest). The feasibility of the pest detection in the stored grain was verified by using CSI amplitude data. On this basis, a Wi-Pest detection method was designed. Firstly, the amplitude data of CSI was preprocessed by outliers removal, data normalization and noise elimination. Then the principal component analysis (PCA) was used to compress the data and extract the main feature components. Finally, random forest (RF) classification method was used to detect stored grain pests. Experiments show that the abnormal density of live pests in grain heaps can be detected under the line of sight (LOS) scenario, and the detection accuracy of the proposed method can reach 97% on average.

Key words: CSI, stored grain pest detection, amplitude, random forest classification

收稿日期: 2019-10-21; 修回日期: 2020-03-03

通信作者: 杨卫东, Yangweidong@haut.edu.cn

基金项目: 国家重点研发计划(No.2017YFD0401001); 国家自然科学基金资助项目(No.61741107); 河南省省属高校基本科研业务费专项资金资助项目(No.2018RCJH12); 河南工业大学科研平台开放课题(No.KFJJ-2016-202)

Foundation Items: The National Key R&D Program of China (No.2017YFD0401001), The National Natural Science Foundation of China (No.61741107), The Fundamental Research Funds for the Henan Provincial Colleges and Universities (No.2018RCJH12), The Open Fund for Henan University of Technology (No.KFJJ-2016-202)

1 引言

粮食储藏安全关乎国计民生。全球粮食安全是世界人口面临的重要问题,世界人口数量正在迅速增长,到2050年将超过91亿人^[1-2]。随着人口数量和社会流动性的增长,到2050年,人类对粮食的需求量将达到目前需求量的两倍^[3]。目前,全球每年粮食产量超20亿吨^[4],因此,需要安全地存放这些粮食以应对未来人们对粮食的需求,特别是应对自然灾害和饥荒等紧急情况。在粮食储藏期间,储粮害虫的危害十分严重。全世界每年在储藏期间的粮食至少有10%被储粮害虫蛀食^[5],我国的国库储粮损失率为0.2%^[6],如何确保向不断增长的人口提供安全的粮食将是未来的主要挑战。为了减少储粮过程中粮食的损失,必须把储粮害虫数量控制在经济损害水平以下,这样既不会因储粮害虫造成损失,也不会因盲目地防治造成浪费。随着科学技术的发展,仍需不断研究和探讨新的、有效的储粮害虫检测方法,为储粮害虫的综合防治提供科学的决策依据,及时采取合理的防治措施,达到粮食保质、保量、保鲜的目的。在粮食储藏过程中,检测害虫的发生情况是当前粮食行业的研究热点之一,也是粮库日常管理人员的重要任务之一。因此,对储粮害虫进行有效检测具有重大的研究意义。

在粮食储藏过程中,为了防止害虫对粮食数量和品质造成损害,我国每年都要消耗巨大的人力、财力用于保证粮库储粮安全。储粮害虫检测技术方法大致分为两类,一是传统的人工检查法^[7];二是通过现代信息技术,如声学^[8]、图像^[9-10]、光学^[11]、电磁学^[12-14]等新方法,对害虫进行检测分类。现有的传统人工检测技术包含扦样法与诱集检查法。扦样法是现有粮食行业标准中国家规定的检查方法,主要通过粮库保管员定期扦取粮食样品进行检查。诱集法是利用害虫的某些习性,将害虫诱捕到诱捕设备中,然后由人工进行识别与计数。为了掌握粮库中害虫的发生情况,一般采用人工检查法对害虫进行检测,但工作量大、效率低,尤其是实现高效的自动检测难度大。而且由于保管员的个人技术等因素,检查结果的准确性也无法保证,难以适应现代化粮食储藏的需要。为了解决这些问题,许多国内外研究学者提出了各种新的研究方法。声音识别技术通过害虫在取食、运动、通信时发出的声音检测害虫。秦昕等^[8]提出了基于害虫活动声音的检测

方法,但受环境噪声的影响较大,装置复杂且成本高,在大范围推广应用有一定的困难。Liu等^[9]和赵彬宇等^[10]相继提出了基于害虫图像识别的研究方法,但只能识别粮粒外部的害虫,对于粮面下的害虫需要通过诱捕的方式获得。近红外识别法是基于粮粒对近红外光吸收与反射的差异,把藏有储粮害虫的粮粒与完好无损的粮粒进行区分。Elizabeth等^[11]提出使用近红外光谱法检测害虫,但不能在线测量,且工作效率低。软X射线法是根据谷粒受害虫蛀食以后,谷粒密度下降而使成像有所变化,是一种非破坏性检测方法。邵小龙等^[12]提出了一种基于软X射线与低场核磁检测小麦隐蔽性害虫玉米象的检测方法,但该技术不能检测包含虫卵的粮食样本。Mankin等^[13]首次提出基于微波雷达系统的储粮害虫检测方法,检测谷物中隐藏的不同大小的害虫。微波雷达系统基于多普勒效应,微波信号经移动物体反射后会产生多普勒效应。当害虫向接收器方向运动时,雷达反射频率略高于发射频率;反之,反射频率低于发射频率,即经反射后的微波信号与发射波信号的频率会产生微小的偏移,储粮害虫的运动可由发射和反射雷达频率区分。但微波雷达设备价格高,普及有一定的困难。另一方面,需要对检测环境做屏蔽背景噪声处理,操作也有难度。Pearson等^[14]通过测定小麦电导率检测麦粒中的隐蔽害虫,但需要单粒检测,检测效率低。上述方法在某些场合对害虫的检测具有一定的效果,但仍然难以满足目前检测的需要,因为这些技术中的一些方法耗时、设备价格高、有潜在的健康危害并且检测效率较低。

普通商用Wi-Fi设备逐渐在人类生活的各个方面普及,越来越受到研究者的关注,出现了许多基于Wi-Fi设备的无线信号进行的科研工作。基于CSI的传感系统已经应用于室内定位和无线设备传感领域,近年来使用CSI数据和指纹识别技术的室内定位系统已经成为主流的方法。FiFS^[15]和DeepFi^[16]系统都是使用CSI振幅值数据进行室内定位的。PhaseFi^[17]系统是利用已经校准的CSI相位数据和BiLoc^[18]系统将双模CSI数据与指纹结合,用于具有深度自动编码器网络功能的室内定位。CiFi^[19]系统考虑室内定位的相位差数据,结合深度卷积网络来学习CSI的图像数据,以提升定位精度并降低数据存储需求。此外,Com Loc^[20]系统通过分析CSI相位信息对室内环境的敏感性,提出信载波链路的

思想,通过相位差选取可靠、稳定的链路信号,减少对位置的误判,实现室内复杂环境下定位结果的高效性和有效性。

基于 CSI 的无线设备传感系统的功能主要包括活动识别、跌倒检测和生命体征检测。在活动识别方面, E-eyes^[21]系统使用 CSI 振幅实现了家庭活动识别的无设备定位系统。对于口语识别, Wi-Hear^[22]系统利用专门的定向天线测量由嘴唇活动引起的 CSI 数据变化。CARM^[23]系统提供了基于 CSI 的速度模型和活动模型,用于识别人类活动与 CSI 动态变化之间的关系。在跌倒检测方面, Wi-Fall^[24]和 RT-Fall^[25]使用 CSI 振幅和相位差检测对象的下降活动。在生命体征检测方面, PhaseBeat^[26]和 TensorBeat^[27]使用 CSI 相位差检测单人或多人的呼吸速率。在最近的 Wi-Fire^[28]和 Wi-Metal^[29]研究中, CSI 数据用于检测火灾事件和金属物体。

随着 Wi-Fi 的发展,可以利用 Wi-Fi 信号非接触地感知人类等移动物体的变化^[21-26]。鉴于以上基于 CSI 传感技术的成功应用和商用 Wi-Fi 易于访问、快速、低成本等因素,提出的检测方法基于 Wi-Fi CSI 信号感知活体害虫和害虫密度,由于 Wi-Fi 的特点,所以本文提出的方法是一种非接触、快速、低成本的检测方法。本文提出的基于 CSI 振幅的储粮害虫检测设计方法基于 Wi-Fi 技术的发展获取灵感,通过获取商业网卡物理层的 Wi-Fi CSI 感知活体害虫和害虫密度的异常变化。首先,通过使用 CSI 振幅数据验证储粮害虫检测的可行性,在此基础上,设计了 Wi-Pest 系统模型和 Wi-Pest 检测方法。Wi-Pest 系统模型包括数据预处理、特征提取和 RF 分类。首先,对 CSI 振幅数据进行异常值检测、数据归一化和噪声消除预处理;然后,通过 PCA 方法压缩数据并提取主特征成分;最后,采用 RF 分类方法检测害虫的异常情况。本文新方法的提出和系统的设计,一方面论证了通过 Wi-Fi 信号进行储粮害虫检测的可行性,提高了储粮害虫密度分类精度。另一方面,与现有的储粮害虫检测方法相比,人力与物力成本大大减少。

2 背景知识

CSI 表示通信链路的信道特性信息,是一种颗粒度的信道信息,反映发送器和接收器之间通信链路的信道特性。它不仅提供有关信号传输过程的基

本信息,还揭示了接收信号的信道特性,如多径效应、阴影衰落和失真。

近年来,正交频分复用(OFDM, orthogonal frequency division multiplexing)技术与多进多出(MIMO, multiple-input multiple-output)技术结合成为了新一代无线局域网和广域网的核心技术,并且实现了从商用无线信号提取 CSI^[30]。OFDM 技术^[31]广泛应用于物理层和现代无线通信系统,如 LTE^[32]和 Wi-Fi^[33]。OFDM 技术将总频谱分为多个正交子载波。此外,在子载波上传输的数据用来处理室内环境中频率选择的衰减。OFDM 技术应用于室内定位^[16]、活动识别^[23]等无线传感领域,它可以从某些 Wi-Fi 芯片组的开源设备中提取丰富的 CSI 数据。

对于在 2.4 GHz 或 5 GHz 频带中的 Wi-Fi OFDM 物理层,子载波可以被视为衰落信道上的窄带,其对射频的感应是非常稳定的。第 i 个子载波的信道频率响应计算式可以写为

$$h_i = |h_i| \exp\{j\angle h_i\} \quad (1)$$

其中, $|h_i|$ 和 $\angle h_i$ 是第 i 个子载波的振幅和相位信息。

Wi-Pest 系统利用 CSI 振幅数据进行储粮害虫检测,支持 IEEE 802.11n 协议的 Atheros AR9580 NIC 网卡的设备驱动程序,可以在物理层提取 56 个子载波的 CSI 数据。

3 储粮害虫检测方法

3.1 Wi-Pest 系统模型

Wi-Pest 系统模型如图 1 所示,该系统包含两个主要部分,即 CSI 数据提取部分和 CSI 数据处理部分。CSI 数据提取部分利用两台配备了 Atheros AR9580 NIC 网卡的笔记本电脑,其中配备一根天线的戴尔 PP181 型号笔记本电脑作为发射器,另一台配备 3 根天线的戴尔 Latitude 5480 型号笔记本电脑作为接收器,操作系统均是内核为 4.1.10+ 的 32 位 Ubuntu Linux 14.04 操作系统。本系统的核心部分是 CSI 数据处理部分,从 CSI 数据中提取主要特征,然后通过 RF 分类算法确定粮堆中是否有虫子。CSI 数据处理部分包含 3 个功能模块,即数据预处理、特征提取和 RF 分类。数据预处理包括对 CSI 振幅数据进行异常值检测、数据归一化和消除噪声。特征提取利用 PCA 方法提取 CSI 序列的特征,不仅保留了主要特征,还减小了 CSI 数据的空间维度。

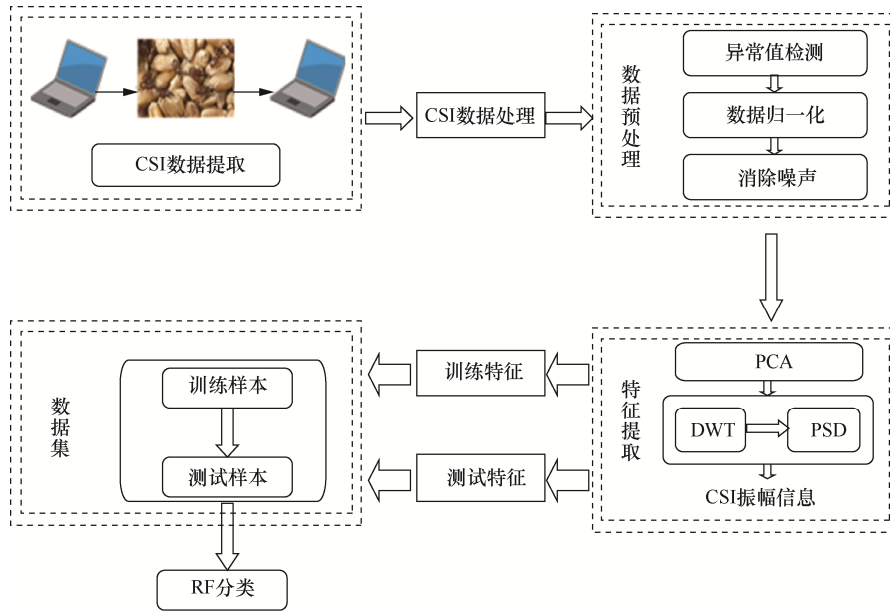


图1 Wi-Pest 系统模型

最后，应用 RF 分类算法检测害虫是否存在。

3.2 Wi-Pest 检测方法

Wi-Pest 检测方法包含数据预处理、特征提取、RF 分类^[34]3 部分。

3.2.1 数据预处理

数据预处理包含异常值检测、数据归一化和消除噪声，对采集的 CSI 数据进行校准。

1) 异常值检测

由于 Wi-Fi 信号对环境变化非常敏感，经常会因不确定的扰动因素产生不确定的异常值。在储粮害虫异常检测系统中，采用 Pauta 准则检测和去除 CSI 振幅数据中的异常值。详细过程如下。

步骤 1 令 $X_i (i=1,2,\dots,n)$ 表示 CSI 样本的振幅，计算其算术平均值 \bar{X} 为

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i \quad (2)$$

步骤 2 使用式(3)计算 V_i ，式(4)计算振幅的标准偏差。

$$V_i = X_i - \bar{X}, i=1,2,\dots,n \quad (3)$$

$$\sigma = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2} \quad (4)$$

步骤 3 对于所有的 $i=1,2,\dots,n$ ，如果 $|V_i| > 3\sigma$ ，那么认为 X_i 为异常值，并使用算术平均值 \bar{X} 替换 X_i 。

步骤 4 重复上述 3 个步骤，直到检测完所有的采集样本。

2) 数据归一化

为了提高检测精度，在进行 RF 分类操作时应该将输入值限制在(0,1)。因此，选择对 CSI 数据的振幅进行归一化，归一化后的值记为 Y_i ，计算方法如式(5)所示。

$$Y_i = \frac{X_i - X_{\text{mean}}}{X_{\text{max}} - X_{\text{min}}} \quad (5)$$

其中， X_i 表示原始数据， X_{mean} 是振幅的平均值， X_{max} 和 X_{min} 分别是在去除 CSI 数据异常值后的一段时间内的振幅的最大值和最小值。

3) 消除噪声

在应用特征提取技术之前，选择切比雪夫 II 型滤波器进一步去除环境噪声。定义切比雪夫 II 型滤波器的响应函数为

$$|H(j\omega)|^2 = \frac{\varepsilon^2 C_N^2\left(\frac{\omega_s}{\omega}\right)}{1 + \varepsilon^2 C_N^2\left(\frac{\omega_s}{\omega}\right)} \quad (6)$$

其中， $\varepsilon (0 < \varepsilon < 1)$ 是阻带中的振幅频率的波动， ω_s 是描述频率的标定常数， N 是多项式 $C_N^2\left(\frac{\omega_s}{\omega}\right)$ 的阶数，计算方法如式(7)所示。

$$C_N(x) = \begin{cases} \cos(N \cos^{-1}(x)), & |x| \leq 1 \\ \cosh(N \cos^{-1}(x)), & |x| > 1 \end{cases} \quad (7)$$

3.2.2 特征提取

经过预处理的 CSI 数据不能直接用于分类, 还需要提取有代表性的特征。在储粮害虫异常检测方法中, 首先使用 PCA 方法进行特征提取, 该技术不仅可以集中在主要的数据特征上, 还可以减少输入数据的维度。在特征提取阶段, 应用离散小波变换 (DWT, discrete wavelet transformation) 分离高频信息和低频信息, 并应用功率谱密度 (PSD, power spectral density) 将在一段时间内收集的 CSI 数据转换为其功率密度, 可以得出害虫的移动主要影响低频率信号的结论。对于 CSI 振幅数据, 可以使用 PCA 方法计算每一个 CSI 序列的主分量 p , 因此可以得到一个大小为 $p \times n$ 的最佳子载波特征矩阵。其中, n 是数据分组的总个数, p 是子载波信号波动幅度变化比较小的最佳子载波簇数量。实验过程中存在由不确定因素引起的子载波幅值变化, 排除不是由活体害虫密度引起的子载波幅值变化, 所以选择一个合理的 p 值是最重要的。通过求取子载波的平均

绝对偏差来选择一组在测量过程中幅度值波动较小的子载波簇, 子载波的平均绝对偏差曲线如图 2 所示, 然后再结合子载波累计贡献率 (如图 3 所示), 采取折中的方案, 可以看出前 38 个子载波的波动变化较小, 且子载波累计贡献率高达约 98%。通过实验数据分析, $p = 38$ 为测量过程中引起幅度变化较小的最佳子载波簇的数量。因此, 在本文所有实验过程中, 设置 $p = 38$, 该过程描述如下。

步骤 1 预处理, 使用经过数据预处理的 CSI 振幅数据创建 CSI 矩阵, 表示为

$$\mathbf{Z} = \begin{bmatrix} Z_{11} & Z_{12} & Z_{13} & \cdots & Z_{1n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ Z_{m1} & Z_{m2} & Z_{m3} & \cdots & Z_{mn} \end{bmatrix} \quad (8)$$

其中, m 是子载波的数量, Z_j 表示为分组 j 记录的子载波处理后的 CSI 振幅。

步骤 2 计算相关矩阵, 计算 $\frac{1}{n} \mathbf{Z}^T \mathbf{Z}$ 得到关联

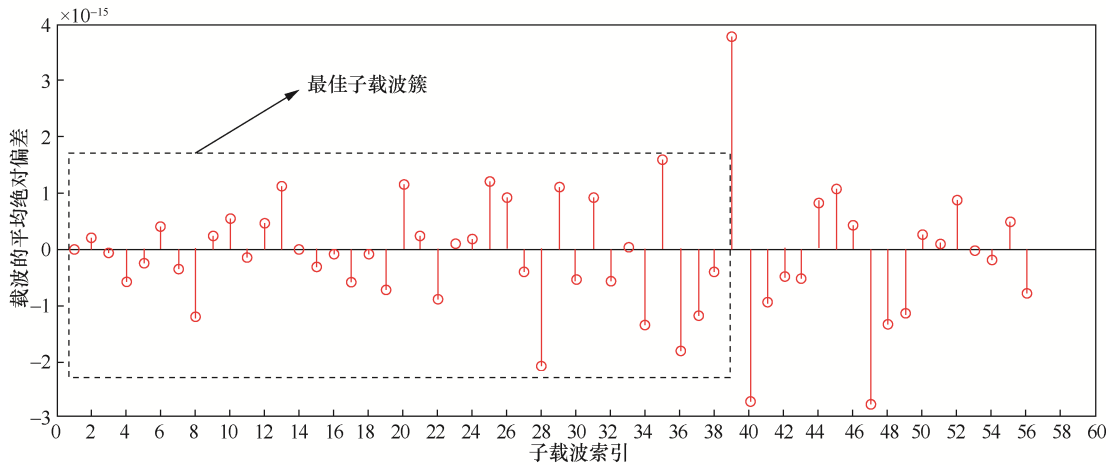


图2 子载波的平均绝对偏差曲线

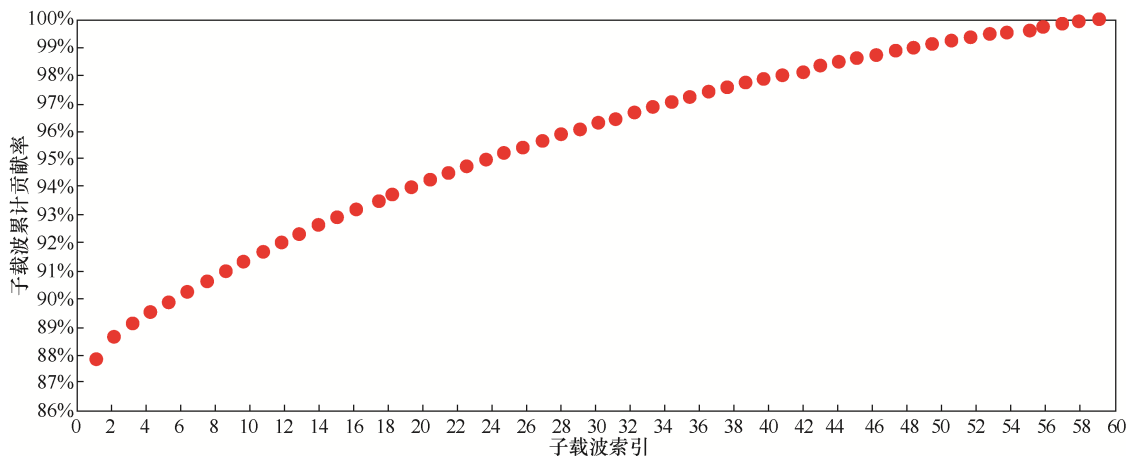


图3 子载波累计贡献率

矩阵, 其大小为 $n \times n$ 。

步骤 3 计算特征向量, 使用相关矩阵 $\frac{1}{n} \mathbf{Z}^T \mathbf{Z}$, 应用特征分解计算特征向量 $\mathbf{R}_i (i=1, 2, \dots, p)$ 。

步骤 4 重建矩阵, 使用关联矩阵和特征向量创建一个新的 CSI 矩阵, 即 $\mathbf{Z}_i = \mathbf{R}_i \mathbf{Z}$, 其中, \mathbf{Z}_i 是第 i 个主成分, \mathbf{R}_i 是第 i 个特征向量。

3.2.3 RF 分类

RF^[34] 分类算法是一种集成算法, 它属于 Bagging 类型, 组合多个弱分类器, 最终结果通过投票或取均值获得, 使得整体模型的结果具有较高的精确度和泛化性能。“随机”使它具有抗过拟合能力, “森林”使它更精确。提取主特征后的数据要划分为训练数据和测试数据两部分, 随机从训练数据中抽取数据, 然后用抽取的数据构建一些决策树, 决策树最终投票支持测试数据的分类。该算法的训练过程如下。

步骤 1 计算采样数据的基尼系数, 计算方法为

$$\text{Gini}(S) = 1 - \sum_{i=1}^m p_i^2 \quad (9)$$

其中, p_i 是要素类 i 中样本的频率, m 是数字要素类, S 是采样数据集。

步骤 2 拆分数据集 S 中的节点, 并计算拆分数据集的基尼系数, 如下

$$\text{Gini}_{\text{split}}(S) = \frac{|S_1|}{S} \text{Gini}(S_1) + \frac{|S_2|}{S} \text{Gini}(S_2) \quad (10)$$

步骤 3 继续通过最小 Gini 索引分割集合 S 中的节点。在拆分数据集上, 继续应用步骤 1 和步骤 2, 直到所有节点都成为树的叶子。

该算法的分类过程如下。

步骤 1 给定测试数据 X 和决策树 k 的数量, 得到

$$R(X) = \max \sum_{i=1}^k I(r_i(X) = y) \quad (11)$$

其中, $R(X)$ 是 RF 分类算法的分类结果, $r_i(X)$ 是决策树 i 的分类结果, y 代表分类的目标, $I(\cdot)$ 代表特征函数。

步骤 2 根据投票计算 $R(X)$ 的比例。

4 基于 CSI 检测储粮害虫的可行性验证

Wi-Fi 设备已经成功应用在许多领域, 但是还未用于检测储粮害虫。本文首次提出利用 Wi-Fi 信号 CSI 振幅数据检验储粮害虫的方法, 是一次

新的尝试。为了实现这一检测方法, 需考虑 Wi-Fi 信号的变化是否是由活体害虫和害虫密度引起的, 为验证使用 CSI 数据检测储粮害虫的可行性, 需要实验验证玉米象活体害虫和害虫密度对 Wi-Fi 信号的影响。

4.1 实验环境及材料

由于 CSI 信号对环境感知较敏感, 在采集实验数据时, 要确保采集的实验数据是由活体害虫和害虫密度不同所引起的。采集原始的实验数据, 为了避免室内及室外其他因素的影响, 室温控制在 25°C , 相对湿度控制在 $53\% \text{RH}$, 实验安排在上午 8—11 点, 在 LOS 场景下进行。实验材料如图 4 所示, 包括河南省本地相同品种含水量的实验小麦, 一个盛放 3.5 kg 小麦、体积大小为 $35 \text{ cm} \times 15 \text{ cm} \times 10 \text{ cm}$ 的有机玻璃盒子, 以及在养虫器皿中饲养的实验所用玉米象。玉米象是中国储粮的头号害虫, 也是世界性的重要储粮害虫之一, 因此选择玉米象活虫作为实验的研究对象。



(a) 实验小麦 (b) 有机玻璃盒子 (c) 养虫器皿

图 4 实验材料

4.2 实验方案

为了验证使用 CSI 数据检测储粮害虫的可行性, 制定如下两种实验室环境中的实验方案, 来验证活体害虫和害虫密度对 Wi-Fi 信号的影响。

4.2.1 可行性实验方案 1

在模拟小麦粮堆害虫的研究中, 前期已经进行了 4 组接收天线与发射天线之间距离不同的实验。当 Wi-Fi 信号穿过活体害虫小麦混合物时, 分别收集了发射天线与接收天线距离为 15 cm 、 30 cm 、 60 cm 、 100 cm 的 CSI 数据, 用来分析 Wi-Fi 信号近似的检测范围。通过分析可知, 在 LOS 为 30 cm 以内的场景下, 活体害虫对 Wi-Fi 信号才会产生影响。所以选择把发射天线与接收天线固定为 30 cm 来制定实验室环境中的第 1 种实验方案, 验证玉米象活体害虫对 Wi-Fi 信号的影响, 实验方案 1 如表 1 所示。

表 1

实验方案 1

对比类型	小麦虫口密度/(头·kg ⁻¹)	实验内容
无虫与大量活虫	0 与 40	向装满小麦的塑料瓶子放入约 140 头玉米象活体害虫, 并将其埋入实验有机玻璃盒子, 然后把未放入害虫的塑料瓶子埋入实验有机玻璃盒子, 分别收集 CSI 数据
无虫与大量死虫	0 与 40	向装满小麦的塑料瓶子放入约 140 头玉米象死虫, 并将其埋入实验有机玻璃盒子, 然后把未放入害虫的塑料瓶子埋入实验有机玻璃盒子, 分别收集 CSI 数据

4.2.2 可行性实验方案 2

实际储粮环境中的害虫不可能固定在塑料瓶子中。粮食中的大部分害虫都是分布在粮粒表面、粮粒间隙或者粮粒内部的。设计的储粮害虫检测方法只对在小麦间隙或者表面运动的玉米象活体害虫有效。由于实验方案 1 的局限性, 又模拟了一个害虫正常运动的储粮环境, 把玉米象活体害虫放到装入小麦的有机玻璃盒子里面, 让其自由正常运动。为了防止有机玻璃盒子里的玉米象跑出, 影响实验数据的准确性, 提前把盒子的上层内壁涂上聚四氟乙烯溶液。做实验之前严格按照虫粮调配比例等级标准(基本无虫粮、一般虫粮、严重虫粮 3 种标准)进行小麦虫粮的调配。实验方案 1 验证了玉米象活体害虫对 Wi-Fi 信号的影响, 接下来验证不同的害虫密度对 Wi-Fi 信号的影响。同样选择把发射天线与接收天线的距离固定为 30 cm 来制定第 2 种实验方案, 实验方案 2 如表 2 所示。

4.2.3 实验方案 1 可行性结果分析

本实验通过收集的 CSI 数据, 分析了相位差的对比图形, 但是得到的是一些无规律的相位差图形。所以本实验没有采用分析 CSI 相位差的方法来检测储粮害虫, 而是选择分析 CSI 振幅的害虫检测方法。为了验证使用 CSI 振幅数据进行玉米象活体害虫检测的可行性, 进行以下两种场景下的实验。无虫与大量活虫的 CSI 振幅对比如图 5 所示, 玉米象活体害虫放入塑料瓶子时, CSI 振幅值发生了很大变化。无虫与大量死虫的 CSI 振幅对比如图 6 所示, 玉米象死虫放入塑料瓶子时, CSI 振幅值基本无变化。通过对比, 一方面可以说明玉米象活体害虫引起 CSI 振幅值的变化, 另一方面也排除了虫子

自身的成分引起 CSI 振幅的变化。因此, 通过可行性实验 1 得出, 可以通过虫子对 CSI 振幅数据的变化影响, 实现对玉米象活体害虫的检测。

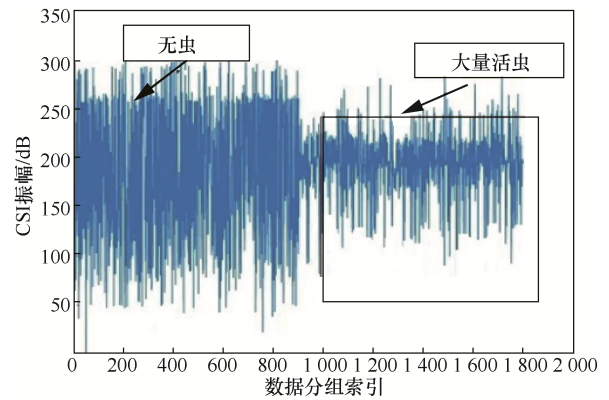


图 5 无虫与大量活虫的 CSI 振幅对比

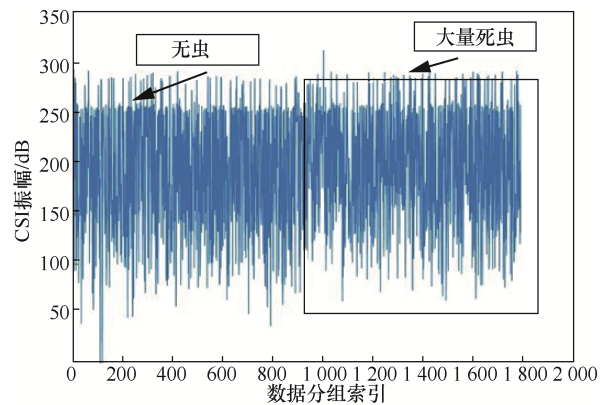


图 6 无虫与大量死虫的 CSI 振幅对比

4.2.4 实验方案 2 可行性结果分析

当 Wi-Fi 信号穿过活体害虫小麦堆时, 首先收集了 4 种不同虫口密度小麦样本的 CSI 振幅数据。根据理论可知, 如果小麦中大量玉米象的移动对

表 2

实验方案 2

对比类型	小麦虫口密度/(头·kg ⁻¹)	实验内容
无虫粮与基本无虫粮	0 与 4	实验用有机玻璃盒子可以盛放约 3.5 kg 小麦, 按照 0、4 头/kg、20 头/kg、40 头/kg 配置小麦
无虫粮与一般虫粮	0 与 20	样本, 共配置了 0、14 头/盒、70 头/盒、140 头/盒共计 4 种不同虫口密度比例的小麦样本, 然
无虫粮与严重虫粮	0 与 40	后分别收集这 4 种小麦样本 LOS 场景下的 CSI 数据

Wi-Fi 信号没有产生影响,少量玉米象的移动基本也不会对 Wi-Fi 产生影响。小麦虫口密度变化时 CSI 振幅变化如图 7 所示,小麦虫口密度从 4 头/kg、20 头/kg 逐渐增加到 40 头/kg 时,CSI 振幅数据发生了变化。其中,小麦虫口密度为 0 与 40 头/kg 振幅数据对比最明显。因此,通过可行性实验 2 得出,可以利用 CSI 振幅数据进行储粮害虫密度的检测。

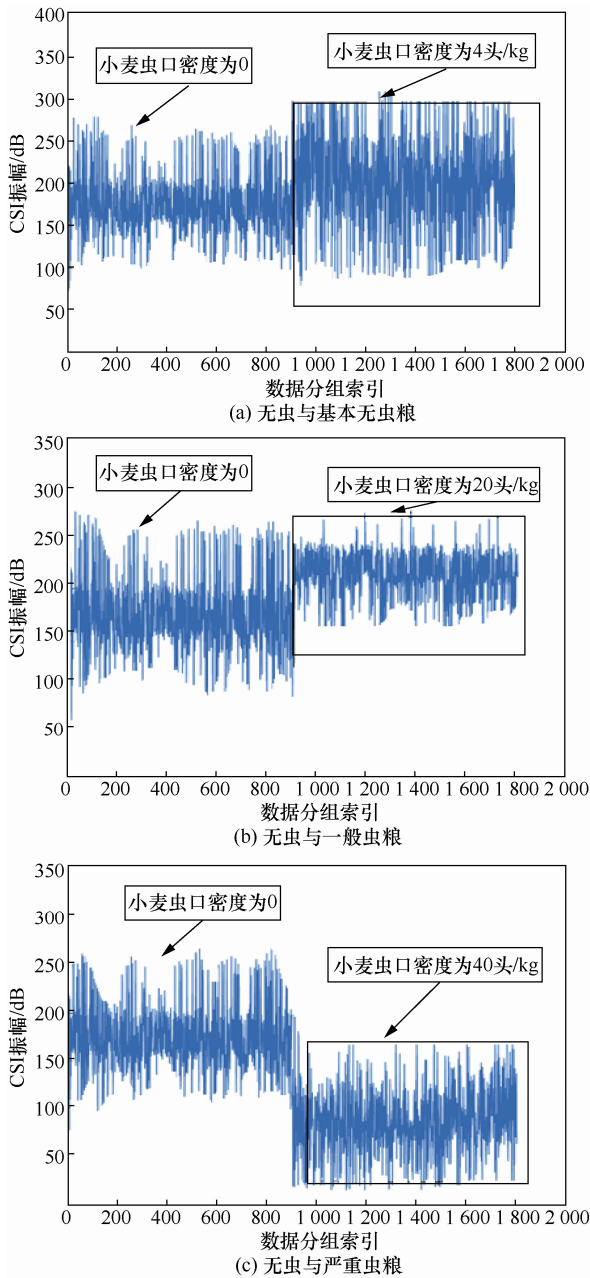


图7 小麦虫口密度变化时 CSI 振幅变化

本实验是基于 CSI 数据的储粮害虫检测方法,通过利用从 Wi-Fi 设备提取 CSI 振幅数据来检测储粮害虫。通过以上两个可行性验证实验可以看出,

能利用 CSI 振幅数值对活体害虫和害虫密度进行检测。这为储粮害虫检测的新方法提供了有希望的思路和亮点,也为后续研究基于 CSI 的储粮害虫区域定位奠定了基础。

5 Wi-Pest 系统性能分析

通过以上两个可行性实验可知,可以利用 CSI 振幅进行玉米象活体害虫的检测和害虫密度的检测。但具体分类精度如何,尚未验证。接下来,进行 Wi-Pest 系统性能精度验证分析。在分类算法上,对比 RF^[34]分类算法和支持向量机^[35](SVM, support vector machine) 分类算法的分类精度。在实验数据收集阶段,每收集一次数据大约得到 1 000 个数据分组,一共收集得到 4 000 个数据分组。通过有监督的机器学习过程,分别给收集的 4 次数据添加标签,0、4、20、40 分别代表无虫粮、基本无虫粮、一般虫粮、严重虫粮 4 种不同的虫口密度数据分组。带有标签 0 的正常数据分组分别与带有标签 4、20、40 的 3 种异常的数据分组经过训练,测试得出 3 组分类精度值。用相同的数据集进行训练和测试过程,选择用 CSI 数据中 70% 的样本训练分类器,用剩余 30% 的样本进行测试。主要害虫不多于 2 头/kg 的小麦都是正常的虫口密度小麦,在本实验中,定义 0 为正常的临界虫口密度。临界虫口密度实验与异常情况下的虫口密度实验作对比,来判断粮堆害虫的异常情况。当处于临界虫口密度时,小麦可以安全地存放很长时间,达到粮食的保质、保量、保鲜要求。选择 0 作为 RF 分类和 SVM 分类的临界值,如果小麦虫口密度大于临界值,则被认为是异常的。

虫子运动范围固定 LOS 场景如图 8 所示,虫子自由移动 LOS 场景如图 9 所示,RF 与 SVM 振幅分类精度值对比如表 3 所示,RF 与 SVM 振幅分类精度对比如图 10 所示。从图 10 可以看出,RF 和 SVM 分类器都能实现分类的目的,且 CSI 振幅分类精度都在 96% 以上。选取 0 作为分类的临界值,分别与 4 头/kg、20 头/kg、40 头/kg 虫口密度作对比,即为 0 与 4 头/kg、0 与 20 头/kg、0 与 40 头/kg 的 3 组不同虫口密度的对比。通过 SVM 分类算法分类精度可以看出,这 3 组分类精度分别为 96.7%、97.2%、96.8%,RF 算法的 3 组分类精度分别为 98.2%、98.5%、98.9%。不难看出,当采用 SVM 分类算法时,0 与 20 头/kg 这组分类精度最高。采用 RF 和 SVM 分类器时,只要分类精度达 90%

以上,算法都是可行的。实验结果表明,当采用 SVM 分类算法时,分类精度值与虫口密度的关系有可能不是假设的单调递增或递减关系,有可能是一种多项式关系。下一步工作需要探究完善这种多项式关系,从而发现分类精度值与虫口密度之间的规律。在相同虫口密度下,RF 分类算法的分类精度比 SVM 分类算法的分类精度高,且分类精度值与虫口密度的关系跟假设的一样,是一种单调递增的关系。可以看出使用 CSI 振幅数据时,提出的储粮害虫检测系统可以实现在 LOS 场景具有较高的分类精度。



图8 虫子运动范围固定 LOS 场景



图9 虫子自由移动 LOS 场景

表3 RF 与 SVM 振幅分类精度值对比

不同虫粮对比	SVM 分类精度	RF 分类精度
无虫粮与基本无虫粮	96.7%	98.2%
无虫粮与一般虫粮	97.2%	98.5%
无虫粮与严重虫粮	96.8%	98.9%

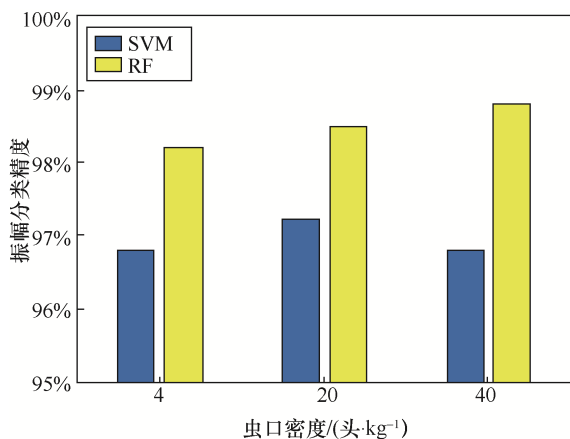


图10 RF 与 SVM 振幅分类精度对比

RF 和 SVM 分类算法在储粮害虫分类性能好坏方面,从样本数据集大小、模型参数调试、模型分类精度 3 个方面做对比分析。

1) 样本数据集大小方面

当有很多训练数据样本时,SVM 分类算法处理效率低,只适合小样本处理。而收集的样本数据是 4 000 个数据分组,此时 RF 分类算法比 SVM 分类算法更高效。

2) 模型参数调试方面

SVM 分类模型要调整很多参数,其中有两个重要的参数 C 与 g , 参数 C 和 g 的选择对 SVM 分类性能的影响很大。 C 是惩罚系数, C 越高,越容易过学习,即容易过拟合; C 越小,容易欠拟合。 g 是核函数到达 0 的速率, g 越小,函数下降越快,也容易造成过拟合。而对于 RF 分类模型,大部分参数使用默认的参数即可,涉及调整的参数主要是 RF 决策树的棵数 (ntree) 及每棵树选择分裂特征数 (ntry),只需要调节好这两个参数即可。

3) 模型分类精度方面

通过以上实验结果的验证得知,在害虫分类方面,RF 分类算法比 SVM 分类算法的分类精度更高。通过以上 3 方面的对比分析可知,在粮食害虫分类方面,RF 分类算法比 SVM 分类算法的综合性能更好,RF 算法能实现更高的分类精度。

6 结束语

本文提出了一种基于 CSI 的非接触、快速、低成本的储粮害虫检测方法。首先,设计验证使用 CSI 振幅数据进行储粮害虫检测的可行性,这是利用 CSI 数据进行害虫检测的第一项工作。其次,设计 Wi-Pest 系统框架模型,在两款带有无线网卡的笔记本电脑上成功部署了该系统,可以从接收器提取 CSI 振幅数据。模型包含 CSI 数据提取和数据处理两部分,在数据处理的过程中,采用了一个机器学习的监督学习模型,即 RF 分类以检测害虫是否存在。最后,进行实验验证 Wi-Pest 系统模型的性能,实验结果表明,该系统可以在 LOS 为 30 cm 以内的场景下保证 97% 以上的检测精度。

本文首次提出的基于 CSI 的储粮害虫解决方法不需要任何专用设备,具有快速、非接触、成本低、易于部署的优点。另一方面,与其他储粮害虫检测方法不同的是,Wi-Pest 检测方法不需要照明环境,

可以在黑暗的环境下工作。从整个粮库发展来讲, Wi-Pest 检测方法一方面可以帮助粮库管理人员制定综合害虫管理计划, 提高管理活动的有效性, 更方便地检测储粮害虫; 另一方面也有利于大力发展粮食信息化, 推进智慧粮库建设。但是, 本文设计的检测方法也有一定的不足之处, 目前只能检测活体害虫, 对死虫没有效果, 且只能在 LOS 为 30 cm 的距离内检测到实验室环境中模拟小麦粮堆害虫的密度异常情况。目前, 只是在实验室环境中检测小麦粮堆中的害虫, 难以满足实际粮库检测害虫的要求, 在粮库内检测害虫难度较大, 在实际应用中的效能有一定的局限性。

参考文献:

- [1] Food and Agriculture Organization of United Nations. Save food: global initiative on food loss and waste reduction[EB/OL]. 2014.
- [2] PARFITT J, BARTHEL M, MACNAUGHTON S. Food waste within food supply chains: quantification and potential for change to 2050[J]. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 2010, 365(1554): 3065-3081.
- [3] VASISHT D, KAPETANOVIC Z, WON J, et al. FarmBeats: an IoT platform for data-driven agriculture[C]//The 14th USENIX Symposium on Networked Systems Design and Implementation. 2017: 515-529.
- [4] JAYAS D S. Storing grains for food security and sustainability[J]. *Agricultural Research*, 2012, 1(1): 21-24.
- [5] ASRAR M, ASHRAF N, HUSSAIN S M, et al. Toxicity and repellence of plant oils against *Tribolium castaneum* (Herbst), *Rhyzopertha dominica* (F.) and *Trogoderma granarium* (E.)[J]. *Pakistan Entomologist*, 2016, 38(1): 55-63.
- [6] 张红涛, 胡玉霞, 邱道尹. 储粮害虫检测现状[J]. *河南农业科学*, 2006(3): 66-68.
ZHANG H T, HU Y X, QIU D Y. Current situation of stored grain pest detection[J]. *Journal of Henan Agricultural*, 2006(3): 66-68.
- [7] 杨自力, 蒋天科, 李伟, 等. 探管诱集与取样筛检两种方法检测储粮害虫数量比较[J]. *粮食储藏*, 2007, 36(3): 16-19.
YANG Z L, JIANG T K, LI W, et al. Comparison of probe-tube traps and sampling inspection to detect stored grain pests[J]. *Grain Storage*, 2007, 36(3): 16-19.
- [8] 秦昕, 郭敏. 2 种储粮害虫活动声信号的检测及其功率谱分析[J]. *华中农业大学学报*, 2012, 31(5): 656-660.
QIN X, GUO M. Detection of activity sound from two species stored grain insects and analysis of its power spectrum[J]. *Journal of Huazhong Agricultural University*, 2012, 31(5): 656-660.
- [9] LIU T, CHEN W, WU W, et al. Detection of aphids in wheat fields using a computer vision technique[J]. *Biosystems Engineering*, 2016, 141: 82-93.
- [10] 赵彬宇, 周慧玲, 李江涛, 等. 储粮害虫智能图鉴及图像识别 App 软件设计[J]. *粮食储藏*, 2019, 48(3): 42-46.
ZHAO B Y, ZHOU H L, LI J T, et al. Mobile App design for stored grain insect illustration guide and image recognition[J]. *Grain Storage*, 2019, 48(3): 42-46.
- [11] ELIZABETH B M, DOWELL F E, BAKER J E, et al. Detecting single wheat kernels containing live or dead insects using near infrared reflectance spectroscopy[C]//American Society of Agricultural and Biological Engineers Annual Meeting. 2002.
- [12] 邵小龙, 徐水红, 徐文. 基于软 X 射线与低场核磁检测小麦隐蔽性害虫玉米象[J]. *中国粮油学报*, 2019, 34(7): 101-106.
SHAO X L, XU S H, XU W. Detection of wheat hidden pest *Sitophilus zeamais* based on soft X-ray and low field nuclear magnetic resonance[J]. *Journal of the Chinese Cereals and Oils Association*, 2019, 34(7): 101-106.
- [13] MANKIN R W. Microwave radar detection of stored-product insects[J]. *Journal of Economic Entomology*, 2004, 97(3): 1168-1173.
- [14] PEARSON T C, BRABEC D L. Automated detection of hidden internal insect infestations in wheat kernels using electrical conductance[C]//American Society of Agricultural and Biological Engineers Annual Meeting. 2002.
- [15] XIAO J, WU K, YI Y, et al. FIFS: fine-grained indoor fingerprinting system[C]//International Conference on Computer Communication Networks. IEEE, 2012: 1-7.
- [16] WANG X Y, GAO L J, MAO S W, et al. CSI-based fingerprinting for indoor localization: a deep learning approach[J]. *IEEE Transactions on Vehicular Technology*, 2017, 66(1): 763-776.
- [17] WANG X Y, GAO L J, MAO S W. CSI phase fingerprinting for indoor localization with a deep learning approach[J]. *IEEE Internet of Things Journal*, 2016, 3(6): 1113-1123.
- [18] WANG X Y, GAO L J, MAO S W. BiLoc: Bi-modal deep learning for indoor localization with commodity 5 GHz Wi-Fi[J]. *IEEE Access*, 2017, 5: 4209-4220.
- [19] WANG X Y, WANG X Y, MAO S W. CiFi: Deep convolutional neural networks for indoor localization with 5 GHz Wi-Fi[C]//2017 IEEE International Conference on Communications (ICC). IEEE, 2017: 1-6.
- [20] 党小超, 司雄, 郝占军, 等. 复杂室内环境下基于 CSI 的定位方法[J]. *物联网学报*, 2018, 2(4): 68-77.
DANG X C, SI X, HAO Z J, et al. Indoor localization method based on CSI in complex environment[J]. *Chinese Journal on Internet of Things*, 2018, 2(4): 68-77.
- [21] WANG Y, LIU J, CHEN Y Y, et al. E-eyes: Device-free location-oriented activity identification using fine-grained Wi-Fi signatures[C]//The 20th Annual International Conference on Mobile Computing and Networking. ACM, 2014: 617-628.
- [22] WANG G H, ZOU Y P, ZHOU Z M, et al. We can hear you with Wi-Fi! [J]. *IEEE Transactions on Mobile Computing*, 2016, 15(11): 2907-2920.
- [23] WANG W, LIU A X, SHAHZAD M, et al. Understanding and modeling of WiFi signal based human activity recognition[C]//The 21st Annual International Conference on Mobile Computing and Networking. 2015: 65-76.
- [24] WANG Y X, WU K S, NI L M. WiFall: device-free fall detection by wireless networks[J]. *IEEE Transactions on Mobile Computing*, 2017, 16(2): 581-594.
- [25] WANG H, ZHANG D Q, WANG Y S, et al. RT-fall: a real-time and contactless fall detection system with commodity Wi-Fi devices[J]. *IEEE Transactions on Mobile Computing*, 2017, 16(2): 511-526.

- [26] WANG X Y, YANG C, MAO S W. PhaseBeat: exploiting CSI phase data for vital sign monitoring with commodity Wi-Fi devices[C]//2017 IEEE 37th International Conference on Distributed Computing Systems (ICDCS). IEEE, 2017: 1230-1239.
- [27] WANG X Y, YANG C, MAO S W. TensorBeat: tensor decomposition for monitoring multi-person breathing beats with commodity Wi-Fi[J]. ACM Transactions on Intelligent Systems and Technology, 2017, 9(1): 1-27.
- [28] ZHONG S X, HUANG Y Z, RUBY R, et al. Wi-fire: Device-free fire detection using Wi-Fi networks[C]//2017 IEEE International Conference on Communications (ICC). IEEE, 2017: 1-6.
- [29] WU K S. Wi-metal: Detecting metal by using wireless networks[C]//2016 IEEE International Conference on Communications (ICC). IEEE, 2016: 1-6.
- [30] WANG Z, HAN J S, XI W, et al. Efficient and secure key extraction using channel state information[J]. The Journal of Supercomputing, 2014, 70(3): 1537-1554.
- [31] 郭峰. 浅谈无线通信系统中的传输技术[J]. 通讯世界, 2019, 26(10): 91-92.
GUO F. Talk about the transmission technology in wireless communication system[J]. Telecom World, 2019, 26(10): 91-92.
- [32] GONG M X, HART B, MAO S W. Advanced wireless LAN technologies[J]. GetMobile: Mobile Computing and Communications, 2015, 18(4): 48-52.
- [33] WANG X Y, MAO S W, GONG M X. A survey of LTE Wi-Fi coexistence in unlicensed bands[J]. GetMobile: Mobile Computing and Communications, 2017, 20(3): 17-23.
- [34] 王奕森, 夏树涛. 集成学习之 RF 算法综述[J]. 信息通信技术, 2018, 12(1): 49-55.
WANG Y S, XIA S T. A survey of random forests algorithms[J]. Information and Communications Technologies, 2018, 12(1): 49-55.
- [35] SUYKENS J, LUKAS L, VAN P, et al. Least squares support vector machine classifiers: a large scale algorithm[J]. Neural Processing Letters, 1999, 9(3): 293-300.

[作者简介]



单少伟(1992-),男,河南郑州人,河南工业大学硕士生,主要研究方向为无线网络、物联网等。



杨卫东(1977-),男,河南郑州人,博士,河南工业大学教授,主要研究方向为无线网络安全、隐私保护、车载自组织网络、物联网、车联网等。



肖乐(1972-),男,河南开封人,河南工业大学副教授,主要研究方向为人工智能、智能计算等。



王珂(1980-),男,河南郑州人,河南工业大学讲师,主要研究方向为大数据、物联网。