基于功率谱特征的 Wi-Fi 射频指纹提取方法

陈天舒[1], 胡爱群[2,3], 姜 禹[1,3]

[1] 东南大学网络空间安全学院 南京 中国 211189
[2] 东南大学信息科学与工程学院 南京 中国 210096
[3] 网络通信与安全紫金山实验室 南京 中国 211111

摘要 近年来, 利用射频指纹(Radio Frequency Fingerprint, RFF)技术对设备进行识别认证, 构建保密通信系统成为研究的热点。相比于传统的认证体制, 射频指纹利用设备本身的硬件特性进行识别, 具有更高的安全性。与其他射频技术相比, Wi-Fi 信号频谱更宽, 应用更加广泛, 但也更容易受室内多径干扰, 造成对 Wi-Fi 射频指纹识别率下降的问题。针对这一问题, 本文提出一种基于功率谱特征的 Wi-Fi 射频指纹提取方法, 通过计算其信号帧中短导码符号和长导码符号的功率谱比值, 并以此比值作为射频指纹特征。本文采用了 27 台 Wi-Fi 路由器进行实验, 在室内场景中模拟的四个不受外界干扰的相对静止情形以及简单的移动环境中采集数据, 运用随机森林模型进行训练和测试, 识别率能达到 93.3%。理论分析和实验结果表明, 本文方法能够较好地抵抗多径效应和加性噪声对射频指纹的影响, 即使设备在相对移动的情况下, 提取的射频指纹信息也具有较好的稳健性。因此, 本文所提的功率谱特征方法在物理层设备认证和身份识别领域具有一定的应用价值。

关键词 物理层安全; 射频指纹; 设备识别; 软件无线电; 随机森林
中图法分类号 TN918 **DOI 号** 10.19363/J.cnki.cn10-1380/tn.2021.03.01

Power Spectrum Based Wi-Fi RF Fingerprint Extraction Method

CHEN Tianshu[1], HU Aiqun[2,3], JIANG Yu[1,3]

[1] School of Cyber Science and Engineering, Southeast University, Nanjing 211189, China
[2] School of Information Science and Engineering, Southeast University, Nanjing 210096, China
[3] The Purple Mountain Laboratories for Network and Communication Security, Nanjing 211111, China

Abstract In recent years, using Radio Frequency Fingerprint (RFF) technology to realize the identification and authentication of devices and build a secure communication system has become a research hotspot. Compared with the traditional authentication systems, RFF identification has higher security by utilizing the hardware characteristics of the devices themselves. Compared to other radio frequency technologies, Wi-Fi signals have broader spectrum and wider application. However, they are also more susceptible to indoor multipath interference, resulting in the decline of the recognition accuracy of Wi-Fi devices. To solve this problem, a novel Wi-Fi RFF extraction method based on power spectrum characteristics is proposed in this paper, in which the power spectrum ratio of short training symbols (STS) and long training symbols (LTS) in the signal frame is calculated as fingerprint feature. In the experiment, the frame data of 27 routers were collected in four relatively static situations without external disturbance and simple moving states simulated in indoor scene, and then they were trained and tested by using random forest model. The recognition rate can reach 93.3%. The theoretical analysis and experimental results demonstrate that this method can better resist the influences of multipath and additive noises, and the extracted RFF characteristics still have good robustness even when the devices are in motion. Therefore, the power spectrum based method presented in this paper has a certain application value in the field of physical layer device authentication and identity recognition.

Key words physical layer security; radio frequency fingerprint; device identification; software defined radio; random forest

通讯作者: 胡爱群, 博士, 教授, Email: aqhu@seu.edu.cn。
本课题得到江苏省产业前瞻与关键核心技术—竞争项目(No. BE2019109), 网络通信与安全紫金山实验室, 移动通信国家重点实验室自主研究经费(No. 2020B05)资助。
收稿日期: 2020-03-27; 修改日期: 2020-04-27; 定稿日期: 2020-12-21

1 引言

随着现代计算机技术和无线通信技术的迅猛发展, 物联网技术的逐渐成熟, 移动终端设备的大量普及, 网络通信已渗透到生产生活的方方面面。而无线网络中传输媒介的开放性、终端的移动性和通信系统的不稳定性也使得传输的可靠性和安全性受到严重的威胁[1], 对物联网中无线设备的身份认证技术等安全性方面的研究近年来正逐渐成为热点[2]。

当前, 为了解决网络中终端设备的接入认证问题, 比较常见的识别与接入控制技术是基于 MAC 地址和数字证书来实现, MAC 地址极易被伪造和篡改[3]; 数字证书认证方式则要求在终端安装数字证书客户端, 且需要较强的计算能力来实现对密钥或证书的分发[4]。而量子计算机的出现使得暴力破解密码变得更加容易, 传统的基于密码学的密钥分发协议在通信中可能会面临更大的安全挑战[5]。射频指纹是无线设备固有的物理特性, 具有唯一性和难以克隆的特性, 因此, 基于设备物理特征的认证机制将可能是物联网身份认证未来发展的趋势[6]。

通过提取无线设备的射频指纹信息进而对无线设备个体进行识别认证, 这种基于物理层特性的认证方法为上述问题提供了一套安全可靠的解决方案[7]。射频指纹是一种独特的物理信息, 可理解为无线设备本身的 DNA。由于无线发射设备中包含众多的电阻、电容、电感等电子元器件, 电子元件存在制造容差和漂移容差, 不同元器件之间或多或少会存在微小的指标差别, 自然环境的变化、器件老化、印制电路板走线等因素也会造成元件指标值的变化[8]。这些因素导致不同的设备拥有不同的硬件参数和性能指标, 并最终生成不同的射频指纹, 而且这些独有的指纹信息难以被复制[9]。射频指纹具有通用性、唯一性、短时不变性、独立性、稳健性等特征, 这些特征使得射频指纹识别系统更加安全可靠[10]。

目前已有一些常用的射频指纹提取方法, 如基于瞬态响应的提取电磁波幅度包络和相位信息、瞬时幅度和瞬时频率、瞬态强度等特征[11], 基于稳态响应的提取前导码特征[12], 基于图像特征[13]的射频指纹提取方法等等。近几年来, 针对不同的信号类型和调制方式以及环境或信道干扰带来的不稳定因素等情况, 研究者提出了一些识别正确率更高的改进方法。如针对 ZigBee 设备提出一种将差分星座轨迹图、载波频率偏移等四种特征混合的射频指纹提取和设备分类方案[14]; 为进一步提高方法的稳健性, 利用多采样卷积神经网络选择兴趣域的信噪比自适应算法[15]和将测量噪声转换为标准高斯分布并在训练阶段加入人工噪声的算法[16]被提出用于 ZigBee 设备的识别; 针对 LoRa 设备提出一种有监督的机器学习和零样本图像分类方法来提高指纹识别率[17]; 针对线性调频脉冲雷达提出了一种基于分段曲线去噪算法和一种混合射频指纹的识别方案[18]; 针对无人机在无线信号干扰条件下, 采用多级检测器分离噪声和有用信号, 从而提取有效的指纹特征进行检测和分类[19]; 针对 Wi-Fi 设备利用从 Wi-Fi 前导码中提取的基于 Gabor 变换的射频指纹, 并对该特征进行降维处理, 对设备分类的性能和对非法设备拒绝的性能均超过 90%[20], 使用瑞利衰落信道可让识别的可靠性优于 95%[21], 并进一步研究了载波频率偏移对指纹的影响[22]。

上述方法中, 针对 ZigBee、LoRa 等其他信号类型的方法大多无法直接应用到 Wi-Fi 设备上来。这是因为 ZigBee 等信号带宽较窄, LoRa 及雷达信号是线性扩频信号, 多径对这些设备的指纹影响较弱。而 Wi-Fi 属于宽带信号, 频谱较为丰富, 在复杂的室内环境中更容易受到多径、信号衰减、人物走动、墙体和障碍物对信号的反射等因素带来的干扰[23], 但上面文献中的这些方法均未对各种干扰因素的影响展开研究。一些文献对 Wi-Fi 信号的研究采用的设备样本数只有几个, 当测试的设备数量增加后, 设备之间的指纹碰撞概率会增大, 已有方法并不适用于对大量设备的识别。从目前对 Wi-Fi 射频指纹提取方法的研究现状来看, 当前提出的方法考虑还不够全面, 虽然大多考虑了加性高斯白噪声的影响, 但针对射频指纹的稳健性和鲁棒性尤其在抗多径方面还需深入研究[24]。射频指纹特征极其微小, 信号传输过程中的多径影响和噪声的干扰都极易湮没设备的指纹信息, 要提取设备的射频指纹仍然面临着极大的挑战性。如需对多径影响开展研究, 不仅要对设备在某一个静止位置上进行射频指纹的采集与识别, 还需要将设备置于不同位置并处于移动状态进行实验。Li 等人[25]在 2019 年提出一种针对 Wi-Fi 设备的基于幅度商(Amplitude of Quotient, AoQ)的射频指纹提取方法, 该方法考虑了多径因素的影响, 在不同位置测量提取的指纹特征具有不变性, 但该篇文章只研究了几个固定点处的实验数据, 没有考虑移动的情况以及其他障碍物带来的干扰, 且该方法针对不同设备的指纹特征区分度还有待进一步提高。

本文针对 IEEE 802.11 Wi-Fi 的前导码信号, 提出一种提取其功率谱特征的新的射频指纹提取方案。首先对接收到的信号完成捕获、同步、频偏估

计和校正等数据预处理步骤, 得到有效的正交频分复用(Orthogonal Frequency Division Multiplexing, OFDM)信号帧。接着从每一帧的数据中提取其功率谱特征, 得到每一帧数据的射频指纹。随后利用 K 近邻算法、朴素贝叶斯算法、随机森林算法等机器学习算法对实际测量数据中提取出的指纹进行分类识别, 验证该方法的有效性。最后对实验结果进行总结, 并提出了可进一步研究的工作。

本文要解决的问题包括:

(1)如何只利用一帧数据便可提取出有效的指纹特征;

(2)如何抵抗多径的影响, 在无线设备所处环境变化的情况下, 所提取的指纹特征依然稳定;

(3)如何抵抗噪声的干扰, 在去噪的同时不丢失指纹信息, 在不同信噪比的条件下均能提取出不变的指纹特征。

本文的主要贡献有:

(1)提出一种基于功率谱特征的射频指纹提取方法, 利用 Wi-Fi OFDM 信号同一帧中短导码符号和长导码符号的功率谱的比值形成指纹。该方法只需利用一帧数据即可提取出指纹特征, 且可以去除信道的影响, 使得设备在多径的环境下, 提取的指纹仍具有较好的不变性。

(2)对 27 台 Wi-Fi 路由器发出的 OFDM 信号按此方法提取出的指纹特征利用机器学习方法分类。运用 K 近邻、朴素贝叶斯、随机森林三种模型对特征数据进行训练和测试, 使用随机森林算法分类正确率最高, 说明随机森林模型应用于本文提出的射频指纹识别方法更具有有效性。

(3)实验中采集装置在实验室内固定的 4 个位置和移动状态下对路由器发出的信号进行采集, 并考虑了直达径和非直达径的情况。实验时, 路由器连续开机时间超过 1 小时, 在不同位置和移动情形下完成对数据的采集, 以验证该方法的鲁棒性以及提取的指纹特征是否具有时空不变性。对这 27 台路由器 4 个静止点位置的数据分类正确率能达到 97%, 对移动状态下的数据分类正确率能达到 90%, 总体平均分类正确率超过 93%。

本文剩余部分安排如下: 第 2 节介绍 Wi-Fi OFDM 信号帧的采集与捕获以及预处理流程; 第 3 节具体介绍基于功率谱的 Wi-Fi 射频指纹提取方法; 第 4 节介绍实验过程并对实验结果进行分析; 第 5 节为总结和展望。

2 Wi-Fi 信号的提取与预处理

本文以 IEEE 802.11n 协议下的 OFDM 信号为主要研究对象, 对 Wi-Fi 无线通信设备指纹提取与识别分类的流程框图如图 1 所示。

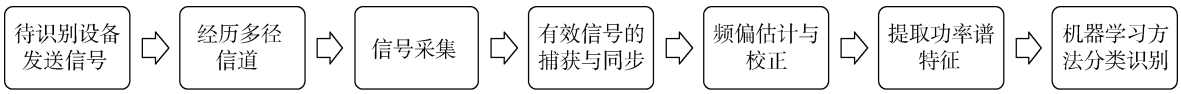


图 1 Wi-Fi 射频指纹提取与识别分类流程框图

Figure 1 Flow chart of Wi-Fi radio frequency fingerprint extraction and classification

本节首先介绍 IEEE 802.11n 信号的物理层信号帧格式以及该信号的采集与捕获方法, 并简述信号预处理的流程, 经过该流程处理过的信号消除了频偏等因素的干扰, 能提取出更加稳定的指纹。

2.1 IEEE 802.11n 物理层信号帧格式

本文将从IEEE 802.11n Wi-Fi信号物理层的帧格式中提取有效的指纹特征。IEEE 802.11n 协议在 OFDM 物理层会聚程序(Physical Layer Convergence Procedure, PLCP)子层中定义了 PLCP 协议数据单元(PLCP Protocol Data Unit, PPDU)的帧格式, 其包括 PLCP 前导码、PLCP 头、PLCP 服务数据单元(PLCP Service Data Unit, PSDU)、尾比特和填充比特[26], 具体的帧结构格式如图 2 所示。由于信号段和数据段会根据所传输的内容发生变化, 而前导码部分传输的内容是恒定不变的, 因此从前导码中提取射频指纹更具有可行性。

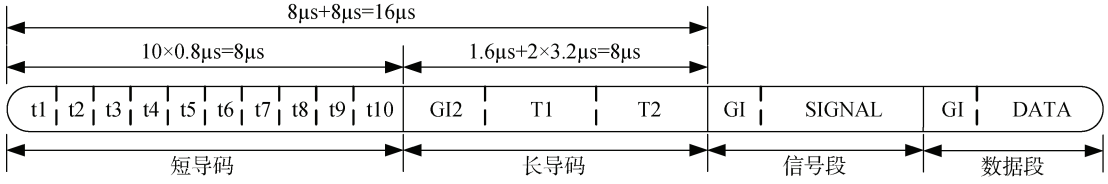


图 2 PPDU 帧格式

Figure 2 PPDU frame format

在 PLCP 前导码字段中包含 10 个重复的短导码符号和 2 个重复的长导码符号, 这些符号是由傅里叶逆变换的周期性引起的循环, 长导码符号之前有一段循环前缀 GI2。短导码符号由 12 个子载波组成, 这些子载波由序列

$$\begin{aligned} S_{-26,26} = \sqrt{\frac{13}{6}}\{&0,0,1+j,0,0,0,-1-j,0,0,0,1+j,\\ &0,0,0,-1-j,0,0,0,-1-j,0,0,0,1+j,0,0,\\ &0,0,0,0,0,-1-j,0,0,0,-1-j,0,0,0,1+j,\\ &0,0,0,1+j,0,0,0,1+j,0,0,0,1+j,0,0\} \end{aligned} \quad (1)$$

的组成元素进行调制, OFDM 短导码利用了这 52 个子载波中的 12 个, 信号由式(2)产生。

$$r_{short}(t) = w_{short}(t)\sum_{k=-N_{ST}/2}^{N_{ST}/2} S_k e^{j2\pi k\Delta_F t} \quad (2)$$

其中, $w_{short}(t)$ 为窗函数, N_{ST} 为子载波的数量, Δ_F 为子载波频率间隔。上述波形在时域是 0.8μs 的序列, 短导码共有 10 个这样的序列, 持续时间 8μs。长导码符号由 53 个子载波组成, 其中包含了一个直流零值, 这些子载波由序列

$$\begin{aligned} L_{-26,26} = \{&1,1,-1,-1,1,1,-1,1,-1,1,1,1,1,1,1,-1,-1,1,\\ &1,-1,1,-1,1,1,1,1,0,1,-1,-1,1,1,-1,1,-1,1,-1,-1,\\ &-1,-1,-1,-1,1,1,-1,-1,1,-1,1,-1,1,1,1,1\} \end{aligned} \quad (3)$$

的组成元素进行调制, OFDM 长导码信号由式(4)产生。

$$r_{long}(t) = w_{long}(t)\sum_{k=-N_{ST}/2}^{N_{ST}/2} L_k e^{j2\pi k\Delta_F (t-T_{GI2})} \quad (4)$$

其中, $w_{long}(t)$ 为窗函数, N_{ST} 为子载波的数量, Δ_F 为子载波频率间隔, T_{GI2} 为循环前缀 GI2 持续时间。上述波形在时域是 3.2μs 的序列, 为了提高信道估计的精确度, 需要发送两个周期的长导码符号, 共计 6.4μs。增加的 1.6μs 的循环前缀 GI2 是由长导码尾部的信号搬移到头部构成的, 以确保做 FFT 循环卷积运算时, 不管从哪里加窗都可以取到一个完整的信号。因此长导码部分的持续时间也是 8μs。这样的设计能够精准地对信号进行捕获和同步, 并进行频偏估计和频偏校正。每一帧 OFDM 信号都有着相同结构的前导码, 从前导码中提取的指纹特征具有更好的一致性。

2.2 Wi-Fi OFDM 信号帧的采集与捕获

采集设备采集到无线 Wi-Fi 信号后需要对每一 OFDM 帧进行分离, 而其前导码有着固定的重复结构, 适合用于对每一帧信号进行快速捕获。如果能从接收到的信号中找到前导码的位置, 也就定位到了每一帧的数据。OFDM 前导码拥有 10 个重复的短导码符号, 可将采集到的信号逐段顺次输入处理器, 计算其与本地短导码符号的相关值, 当连续检测到有 8 段信号的相关值超过设定的门限, 就说明找到了 8 个短导码, 可认为捕获成功。

在 20MHz 的采样率下, 一个持续 0.8μs 的短导码有 16 个采样点。设本地短导码符号为 $x(n)$, 每次送入处理器的长度为 16 个样本点的接收信号为 $y(n)$, 计算两者的复数共轭相关值

$$r = \sum_{n=0}^{15} y(n)x^*(n) \quad (5)$$

其中, $x^*(n)$ 是 $x(n)$ 的共轭。根据实验结果设定一个门限值 TH, 当 $r > TH$, 则认为捕获到一个短导码。考虑到实际信号传输当中首尾两端的符号可能会受到干扰造成波形不稳定的情况, 当连续捕获到 8 个短导码符号, 便可认为捕获到一组 OFDM 信号帧的前导码, 即找到一帧 OFDM 信号。由于后续对信号都是以帧为单位处理的, 因此按上述方法将采集到的信号按帧分离开后, 更加便于后续操作的进行。

2.3 信号预处理

在对每一帧 OFDM 信号预处理的流程中, 主要包括时间同步、频偏估计与校正两大步骤。在之前对信号的捕获步骤中, 只找到了帧起始的大致位置, 而在提取射频指纹时需要定位到每一帧信号如图 2 所示前导码各字段的准确时刻, 即使有一个采样点的偏差, 都会对提取的指纹特征产生较大的影响。同时, 精准的时间同步也为后续的频偏估计和频偏校正服务。

时间同步是将接收信号以滑动接收的方法与本地信号的整个前导部分进行复数共轭相关来实现。整个前导部分由 8μs 的短导码和 8μs 的长导码共计 16μs 组成, 在 20MHz 的采样率下共计 320 个采样点。由于捕获过程找到的帧起始位置可能有偏差, 因此对捕获到的每帧信号前后各增加 16 个采样点, 共截取出 352 个采样点用于同步计算。设本地 320 点前导码信号为 $x_r(n)$, 352 点接收信号为 $y_r(n)$, 计算两者的复数相关值

$$r_{xy}(m) = \sum_{n=0}^{319} y_r(n+m)x_r^*(n) \quad (6)$$

其中 $m = 0,1,\cdots,31$, $x_r^*(n)$ 是 $x_r(n)$ 的共轭。找出 $r_{xy}(m)$ 取最大值时 m 的值, 以此推出每一帧信号精准的起始时刻, 从而实现精准的时间同步。

由于信号在发送端和接收端之间通信时, 载波存在一定角频率偏差, 从而对提取出的指纹特征产

生一定误差。可通过计算相邻两个重复导码间的相位差, 估计出频偏值, 以此消除频偏的影响。短导码部分有 10 个重复相同的长度为 16 个采样点的符号序列, 计算前后相邻的两个短导码符号序列对应点之间相位差的平均值

$$\Delta\theta_i=\frac{1}{16}\sum_{n=0}^{15}\left[\theta_{y_{i+1}}(n)-\theta_{y_i}(n)\right] \tag{7}$$

其中 $i=1,2,\cdots,9$, $\theta_{y_i}(n)$ 是第 i 个短导码符号序列的相位值。接下来, 计算 9 个 $\Delta\theta_i$ 的平均值

$$\Delta\theta=\frac{1}{9}\sum_{i=1}^{9}\Delta\theta_i \tag{8}$$

得到相位偏差平均值, 再由公式

$$\Delta\theta=2\pi\Delta f nT_s \tag{9}$$

$$T_s=\frac{1}{f_s} \tag{10}$$

可估计出频率偏差平均值 Δf, 式中间隔的采样点个数 n 取 16, T_s 指采样间隔, 采样频率 f_s 取 20MHz。

频偏校正过程是根据频偏估计的结果对接收信号实施频偏的校正, 消除频偏对指纹的影响。对于精确同步后的信号, 将 320 点的前导码整体进行频偏校正, 得到校正后的信号

$$y_{freq}(n)=y_{syn}(n)e^{-j2\pi\Delta f nT_s} \tag{11}$$

其中 $y_{syn}(n)$ 是已同步好的信号, $n=0,1,\cdots,319$。对采集到的每一帧 OFDM 前导码均按照上述预处理流程处理完后, 便可从中提取功率谱特征形成指纹, 具体方法将在下一节中介绍。

3 功率谱指纹特征的提取

针对 IEEE 802.11n 协议下的 OFDM 信号帧, 本文重点研究多径条件下射频指纹的提取识别方法。

对于 20MHz 采样率下长度为 320 点的 OFDM 前导码, 取第 2 至第 5 个短导码符号作为序列 STF1, 第 7 至第 10 个短导码符号作为序列 STF2, 取第 1 个长导码符号作为序列 LTF1, 第 2 个长导码符号作为序列 LTF2。根据第 2.1 节的介绍, 取出的这四段序列持续时间均为 3.2μs, 即 64 个采样点。对于理想的本地信号, STF1 与 STF2 完全相同, LTF1 与 LTF2 完全相同。

当终端设备发出无线信号后, 经信道传输, 由接收机接收到的信号会受到多径效应以及加性噪声的影响, STF1、STF2、LTF1、LTF2 四段序列的接收信号可表示为

$$y_{STF1}(n)=s_{STF1}(n)*h(n)+v_{STF1}(n) \tag{12}$$

$$y_{STF2}(n)=s_{STF2}(n)*h(n)+v_{STF2}(n) \tag{13}$$

$$y_{LTF1}(n)=s_{LTF1}(n)*h(n)+v_{LTF1}(n) \tag{14}$$

$$y_{LTF2}(n)=s_{LTF2}(n)*h(n)+v_{LTF2}(n) \tag{15}$$

其中 $n=0,1,\cdots,63$, $y(n)$ 为接收信号, $s(n)$ 为发送信号, 假设这四段前导序列经历的多径信道相同, 均为 $h(n)$, $v(n)$ 为加性噪声。受位置变动和外界干扰的影响, 信道特性处于不断变化中, 从而对射频指纹的提取造成不利的影响。OFDM 的 PPDU 帧格式定义中对前导码结构巧妙的设计再次为信道特性和噪声的消除创造了条件。

加性噪声 $v(n)$ 可通过互相关运算消除。$y_{STF1}(n)$ 与 $y_{STF2}(n)$ 的互相关函数为

$$\begin{aligned}r_{STF}(m)&=\sum_{n=0}^{63}y_{STF1}(n)y_{STF2}^{*}(n+m)\\&=\sum_{n=0}^{63}\left[s_{STF1}(n)*h(n)+v_{STF1}(n)\right]\\&\quad\left[s_{STF2}(n+m)*h(n+m)+v_{STF2}(n+m)\right]^{*}\end{aligned} \tag{16}$$

其中 $m=-63,-62,\cdots,-1,0,1,\cdots,62,63$, 加性噪声 $v(n)$ 可认为是均值为零的平稳随机过程, 有用信号与其互不相关, 两者互相关系数为 0, 因此

$$\begin{aligned}r_{STF}(m)&=\sum_{n=0}^{63}\left[s_{STF1}(n)*h(n)\right]\\&\quad\left[s_{STF2}(n+m)*h(n+m)\right]^{*}\end{aligned} \tag{17}$$

同理可得

$$\begin{aligned}r_{LTF}(m)&=\sum_{n=0}^{63}\left[s_{LTF1}(n)*h(n)\right]\\&\quad\left[s_{LTF2}(n+m)*h(n+m)\right]^{*}\end{aligned} \tag{18}$$

由此可见, 两段短导码序列之间、两段长导码序列之间的互相关系数均与加性噪声无关。

信道特性 $h(n)$ 需要变换到频域来去除。在信号是平稳随机过程的条件下, 互相关函数的傅里叶变换是互功率谱密度, 上式(17, 18)的频域表达式为

$$Z_{STF}(k)=P_{STF}(k)\left|H(k)\right|^2 \tag{19}$$

$$Z_{LTF}(k)=P_{LTF}(k)\left|H(k)\right|^2 \tag{20}$$

其中, $k=0,1,\cdots,126$, $H(k)$ 是 $h(n)$ 的频域表达式, $P_{STF}(k)$ 是 $s_{STF1}(n)$ 和 $s_{STF2}(n)$ 的互功率谱, $P_{LTF}(k)$ 是 $s_{LTF1}(n)$ 和 $s_{LTF2}(n)$ 的互功率谱。将上式(19, 20)相除即可得到具有功率谱特征的射频指纹表

达式

$$RFF(k)=\frac{Z_{STF}(k)}{Z_{LTF}(k)}=\frac{P_{STF}(k)\left|H(k)\right|^2}{P_{LTF}(k)\left|H(k)\right|^2}=\frac{P_{STF}(k)}{P_{LTF}(k)} \tag{21}$$

由此可见, 由式(21)计算得到的 $RFF(k)$ 的表达式中已去除了信道的影响, 只有发送信号 $s(n)$ 的功率谱特征, 可作为待识别发送设备的指纹表达式。但在实际信号当中, 虽然前导码部分包含循环前缀, 但 $h(n)$ 起始点之前的部分为 0, 在计算 RFF 值时信道特性不会完全去除, 存在卷积残留, 同时在上述运算过程中也会消除部分指纹特征。另外, 由于实际信号只能近似看成平稳随机过程, 达不到理想的情况, 最终也难以得到理想的纯粹的指纹, 仍会夹带一定干扰, 但本方法已基本实现了对信道和噪声干扰最大化的去除以及对设备指纹信息最大化的保留。

4 实验流程与结果分析

本节将结合实验探讨基于功率谱特征的射频指纹特征提取方法的有效性和实用性。

4.1 实验器材和环境

实验采用的通用软件无线电外设 USRP 是 Ettus 公司的 USRP N210, 信号采集所用的计算机安装 Linux Ubuntu 16.04 操作系统, 并装有开源的软件定义无线电平台 GNU Radio 等软件。为方便移动, 上述设备由不间断电源供电。本实验所用的采集设备实物图如图 3 所示。信号处理与指纹提取识别所用的计算机安装 Windows 10 操作系统并装有 MATLAB 软件, 主机采用 Intel 酷睿 i5-7500 处理器。

图 3 实验所用的 USRP 和计算机实物图

Figure 3 Picture of USRP and computer used in the experiment

实验研究的待识别设备为 27 台 Wi-Fi 路由器, 其中包含 20 台 MERCURY 品牌 MW305R 型号路由器和 7 台 DLINK 品牌 DWL-2000AP+A 型号路由器, 具体信息见表 1。实验过程中, 路由器参数设置如表 2 所示, IEEE 802.11n 是当前 Wi-Fi 设备的常用协议, 路由器工作在 802.11n 模式下能产生以 OFDM 方式调制的信号, 本文提出的方法将从 OFDM 的前导码中提取出射频指纹特征。

表 1 实验所用的 Wi-Fi 路由器信息

Table 1 Information of Wi-Fi devices used in the experiment

生产厂商	型号	数量	编号
MERCURY	MW305R	20	MERCURY001-020
DLINK	DWL-2000AP+A	7	DLINK001-007

表 2 路由器参数设置

Table 2 Parameter configuration of Wi-Fi devices

选项	参数值
工作模式	802.11n
工作频段	2.4GHz
信道	13(2472MHz)
频段带宽	20MHz

实验在 10m×9m 的实验室中进行, 实验室场地平面示意图如图 4 所示。实验场地宽阔, 能模拟不同通信距离下的情况, 场地内部多径环境较为丰富, 拥有桌椅等障碍物和地面墙面的反射, 可较为完整地模拟真实的多径通信场景。

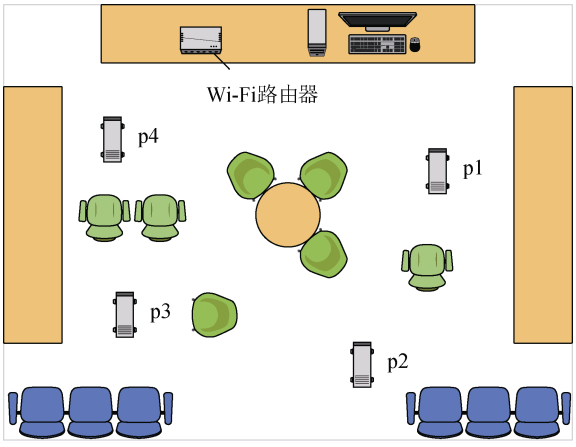


图 4 实验室场地平面示意图

Figure 4 Diagrammatic sketch of laboratory

实验时, 路由器放在固定位置上, USRP 等采集设备放在移动推车上, 在静止和移动两种状态下分别采集信号并进行分析, 两种状态的具体实验环境

和信号采集方式描述如下。

(1)静止状态: 采集设备在图 4 所标注的四个固定的静止点 p1-p4 处采集信号。这四个位置与路由器之间的距离各不相同, 每个位置测量信号的信噪比也不同, 且路径上考虑了有障碍物和无障碍物的情况。在这四个固定点采集信号时, 移动推车置于指定地点后不再移动, USRP 设备及其天线朝向同样保持不变, 实验者在信号直达径之外的地方保持静止不动, 且实验室中没有其他移动设备, 以模拟出相对静止的不受外界干扰的情形。

(2)移动状态: 在移动状态采集信号时, 由实验者推动移动推车在图 4 所示实验室内以平均 1m/s 的步行速度匀速移动, 移动路线不固定, 遍布实验室每一处空地, 包括沿墙边直线行驶、围绕桌椅等障碍物做 8 字形或 S 形运动、原地旋转等, 移动方向随机选择。在推动推车的过程中有一些时刻存在人遮挡天线的情况, 产生阴影区域, 从而模拟没有直达径的情形。同时另一名实验人员在实验室内任意走动, 并多次在 USRP 与路由器之间穿梭, 以模拟受外界干扰的情形。

在整个实验过程中, 同一时间内仅允许一台路由器上电, 其余路由器保持断电状态, 且实验室内其他非实验所需设备也全部关闭, 以免设备之间互相干扰。实验时, 以向路由器 IP 地址发送 ping 指令的方式发起请求, 从而获得路由器发出的信号帧数据。每台路由器连续开机时间需要超过 1 小时, 前半小时在 p1 处以静止状态进行测量, 研究在连续工作情形下指纹是否具有短时不变性, 后半小时内在四个静止点测量, 并让采集设备处于不断移动的情况下连续采集信号, 每个静止点测量 2 分钟, 移动状态测量 6 分钟, 研究空间变化和多径信道对指纹产生的影响。

4.2 实验结果和讨论

实验中, 在模拟的四个不受外界干扰的静止情形和简单的缓慢匀速移动情形下共捕获到 145230 个有效的 OFDM 帧结构, 平均每台设备约 5400 个, 其中每台设备在每个静止点大约捕获到 750 帧数据, 移动状态下约捕获到 2400 帧数据。对以上数据按照第 2 节描述的信号预处理流程完成时间同步、频偏估计与校正两大步骤, 随后按照第 3 节介绍的基于功率谱的方法消除多径影响, 提取待识别设备的射频指纹特征。由于 OFDM 短导码只包含 12 个子载波, 即体现在短导码的频谱图上只有 12 个位置处有值, 其余均为零值点, 计算出的 STF1 和 STF2 两段的互功率谱也是只有 12 个点处值较大, 其余点处的值都很微小。为避免幅值较小点处占比较大的噪声分量带来负面影响, 在计算得到的 RFF 序列中仅选取短导码有数值的 12 个子载波位置处的特征值形成 12 维复数特征向量作为该帧的射频指纹。

图 5~6 所示为实验中某几台待识别设备提取的这 12 维复数特征的幅度值, 同一设备在同一位置处采集到的全部 OFDM 信号帧提取出的特征值均画到了同一张图上。图 5 为编号是 DLINK002 的路由器在四个静止点采集的信号特征, 可见这四张图之间

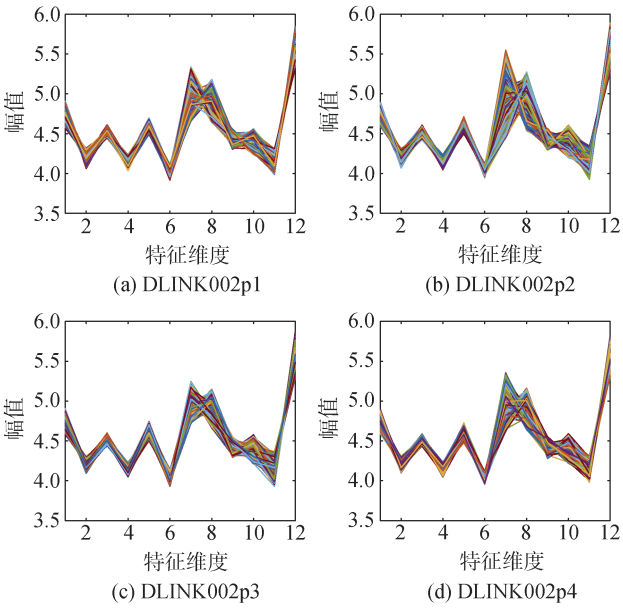


图 5 DLINK002 设备在四个固定点处的功率谱特征

Figure 5 Power spectrum characteristics of device DLINK002 in four fixed positions

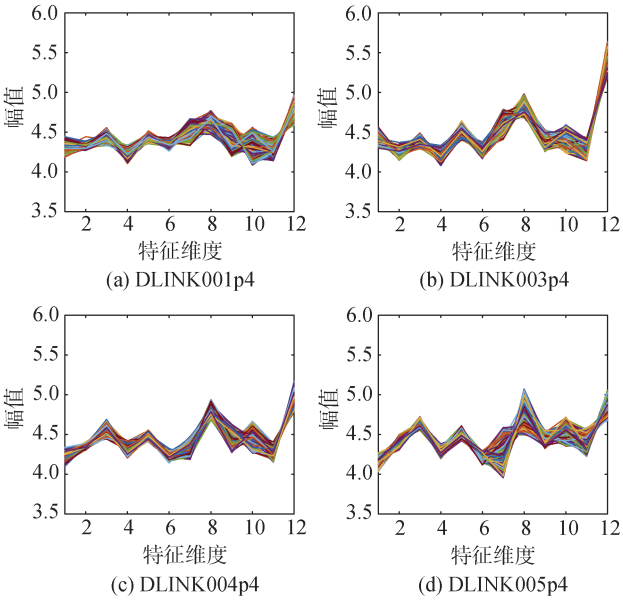


图 6 四台 DLINK 设备在固定点 p4 处的功率谱特征

Figure 6 Power spectrum characteristics of four DLINK devices in fixed position 4

相似度很高, 说明同一设备在不同位置处的特征具有一致性。图 6 为另外四台 DLINK 路由器在同一位置 p4 处采集的信号特征, 可明显地看到这四张图之间差异性较大, 说明即使是同一型号同一批次的不同设备在同一位置处的特征也有一定的区分度。

随后, 在MATLAB 软件上对这27台路由器提取的 145230 个样本特征值利用机器学习方法分类, 本次实验选用 K 近邻、朴素贝叶斯、随机森林三种较为常见的模型对特征数据进行训练和测试, 对于每一台设备在每个固定点位置和移动情况下采集的信号帧各选取 100 帧数据作为训练集样本, 每组数据中剩余数据作为测试集样本, 训练集样本和测试集样本没有重叠。模型参数综合分类结果的正确率和运算时间设定, K近邻模型中邻近点k值设为1, 距离度量类型选择欧氏距离, 随机森林模型中树的数量设为 100。在上述三种模型中, 使用随机森林算法模型进行训练并测试, 得到的分类效果最优, 分类正确的样本数为 135498 个, 正确率为 93.3%, 而使用 K 近邻算法的分类正确率为 93.1%, 朴素贝叶斯算法的分类正确率为 86.6%。三种模型下, 四个静止点位置和移动状态下具体的分类结果如表 3 所示。总体而言, 本文提出的基于功率谱的射频指纹提取方法针对不同设备能提取出具有较高区分度的指纹特征, 分类结果也有着较高的正确率, 而本次实验选取的同一品牌的路由器中都是同一型号同一生产线上的产品, MAC 地址也是相近的, 设备本身相似度就很高, 能达到超过 90%的正确率说明方法本身具备较好的可行性。

表 3 三种模型下 27 台设备基于功率谱特征的分类正确率

Table 3 Classification accuracy of 27 devices based on power spectrum characteristics under three models %

位置 \ 模型	K 近邻	朴素贝叶斯	随机森林
p1	94.6	87.2	93.9
p2	91.1	90.2	94.0
p3	97.1	88.6	96.6
p4	97.5	89.7	97.5
移动	90.4	83.8	90.4
总体	93.1	86.6	93.3

另外, 从表中可知, 固定点位置采集的数据分类正确率要高于移动状态下采集的数据, 说明提取的指纹特征中仍残留部分多径信息, 实际信号在运算过程中无法完全消除信道特性, 外界的扰动会对指纹的识别率产生一定影响。在使用 K 近邻和随机森林两种模型下, p4 位置处数据的正确率要高于 p2 位置处, 在这四个点处, p4 处的接收信号强度最高, 为–44dBm, p2 处的接收信号强度最低, 为–54dBm, p4 处接收信号的信噪比略高于 p2 处, 说明噪声对设备指纹的提取也有一定影响, 在信噪比高的情况下, 提取出的指纹特征受干扰程度小, 判决正确率也越高。

实验中在每台设备开机前半小时内持续采集 OFDM 帧信号, 以检验设备连续工作一段时间指纹特征是否稳定, 指纹的识别率是否下降。本次实验共对 7 台 DLINK 路由器在上电 30min 内采集了 69234 帧有效数据, 平均每台设备近 10000 帧, 对每一帧数据按上文所述同样的步骤分别提取出 12 维的功率谱特征向量。每台设备前 1min 内采集到的信号帧特征向量作为训练集样本, 1~30min 的数据作为测试集样本, 使用上文正确率最高的随机森林算法模型进行训练并测试, 参数配置同上, 7 台设备测试集样本的平均分类正确率为 97.8%。选取其中五台设备计算 1~30min 的每 1min 内的识别正确率, 观察变化趋势。图 7 为这 5 台 DLINK 设备在上电 30min 内基于功率谱特征的逐分钟识别率折线图。从图中可看出, 这 5 台设备的识别正确率都很高, 每个时间点都超过了 97%。虽然只用了第 1min 内的数据进行训练, 但随着开机时间的推移, 识别率并没有出现明显下降的趋势, 均在±1.5%内上下波动。说明提取出的功率谱特征值在一定时间内能保持一定的稳定性, 同一设备的特征向量集中在一定的范围内, 没有出现较大的偏差。

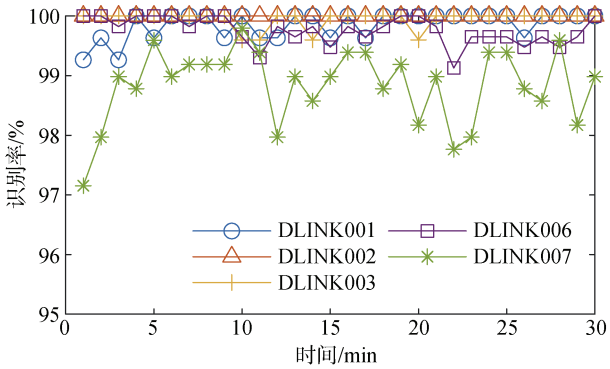


图 7 5 台 DLINK 设备上电 30min 内的逐分钟功率谱特征识别率折线图

Figure 7 Identification accuracy line chart of 5 DLINK devices based on power spectrum characteristics in each minute within 30 minutes after power on

下文将对本文方法与文献[25]所述的提取 AoQ 特征的方法进行对比。针对实验中的待识别设备利

用这两种方法分别提取出 12 维复数特征，并标记在一圆中。将每一维特征的相位归一化到 30°角的区间内，幅度保持不变。图 8 所示为 4 台 DLINK 设备在移动状态下提取的射频指纹特征，同一设备采集到的多个信号帧的特征值叠加在了同一张图上。左边一列是利用本文所提的功率谱方法得到的特征，右边一列是利用 AoQ 方法得到的特征。从图中可观察到，本文方法提取出的每一维特征之间幅度变化较大，不同设备的图形之间区分度也较大，不同设备的同一维度的幅值有着明显的差异。而 AoQ 方法提取出的特征，每一维度的幅值都很接近，基本落在同一半径的圆上，不同设备间的图形相似度很高，区分度不明显。此外，两种方法提取的射频指纹特征值都比较稳定，同一设备的特征向量均较为集中，方差较小，离散程度低，虽然设备处于移动状态当中，但受多径等外界因素影响小，提取的指纹一致性高。而提取 AoQ 特征的方法虽然能很好地消除信道的影响，但同时也消除了设备的指纹信息。相比于该方法，本文所提的方法不仅实现了抗多径、抗干扰的目标，同时也保留住了更多的射频指纹特征。

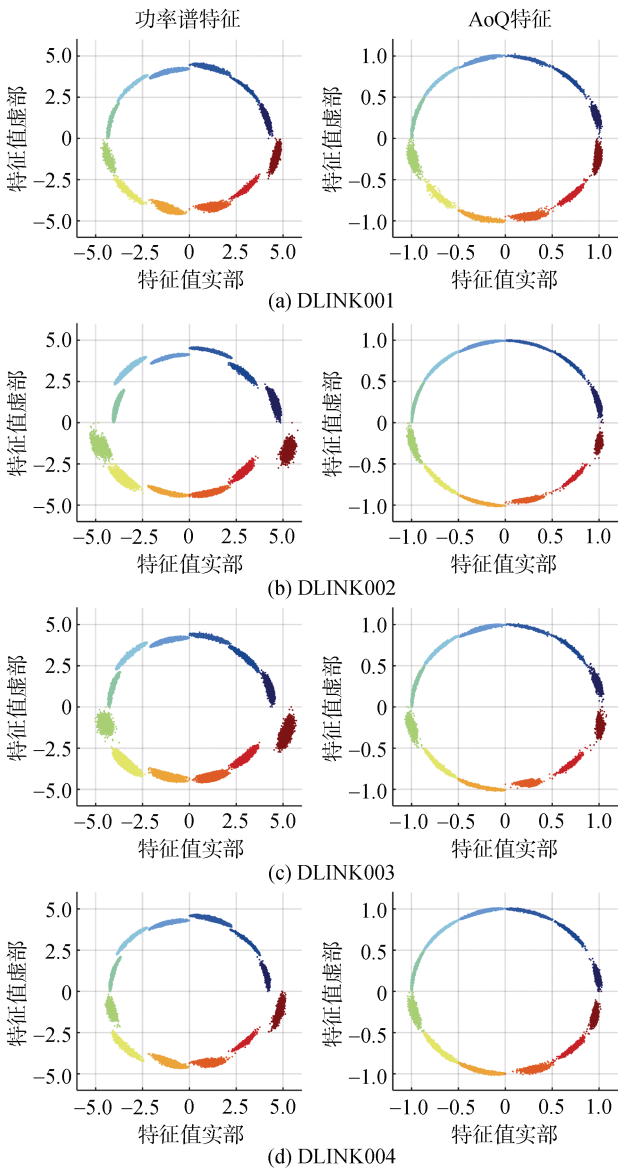


图 8　四台 DLINK 设备在移动状态下通过两种方法提取的特征归一化示意图

Figure 8　Normalized characteristics diagrams of 4 DLINK devices in moving states under 2 methods

接下来，使用识别正确率这一指标定量地进行比较分析。文献[25]中利用长导码部分计算 AoQ 值，本文再将短导码部分也考虑进去，利用长导码和短导码共同计算 AoQ 值，作为另一种提取指纹特征的方法进行性能的比较。针对以上三种方法，使用同样的原始数据和机器学习模型进行训练和测试。本次对比实验中，选用了 7 台 DLINK 路由器采集的 35232 帧数据，对每台设备在每个位置状态下的数据选取 200 帧数据作为训练集样本，每组数据中的剩余数据作为测试集样本，使用效果较好的随机森林模型，分类结果如表 4 所示。从表中可看出，利用本文所提的基于功率谱的射频指纹特征对实验设备进行分类识别能达到97.3%的正确率，要明显优于利用 AoQ 特征分类的方法。基于长导码的 AoQ 识别正确率仅有 41.5%，在此基础上加入短导码计算 AoQ 的改进方法虽然能提高到93.0%的正确率，但依然比本文方法低 4%以上。

表 4　三种不同方法下 7 台 DLINK 设备的分类正确率

Table 4　Classification accuracy of 7 DLINK devices under 3 different methods　%

方法 / 位置	提取长导码 AoQ 特征	提取短导码和长导码 AoQ 特征	提取功率谱特征
p1	52.5	95.6	99.0
p2	43.8	93.5	97.3
p3	47.9	94.7	98.3
p4	37.5	94.7	98.1
移动	35.3	89.7	96.2
总体	41.5	93.0	97.3

总体来看，虽然部分信道特性和噪声干扰没有完全消除，但本文所提的基于功率谱特征的新方法无论是在静止还是移动状态下，较已有的方法在性能上均有很大的提高，且以往的研究大多没有将多径因素重点考虑，缺少在设备移动情况下的实验。本文所做的研究特别考虑了相对移动的情况，采集数据时将采集设备一直处于随机路径的运动当中，运动轨迹遍布整个实验室空间，实验所用的 27 台设备

在移动情形下的分类正确率超过了 90%, 7 台同一流水线生产的 DLINK 设备在移动情形下取得 96.2%的分类正确率, 由此可见本方法体现出较强的抗移动、抗多径性能。

5 结论

本文针对 IEEE 802.11n 协议的设备提出一种新的射频指纹提取方案, 利用设备发出的 OFDM 信号帧数据, 提取其前导码功率谱特征作为设备指纹, 实现对不同设备的区分。本方法重点考虑了信道参数对射频指纹的影响, 对于不同时间段内采集的数据以及训练和测试地点变换的情况下, 通过该方法提取的 Wi-Fi 射频指纹特征具有很好的稳定性和时空不变性, 受多径信道影响小, 在身份认证和识别领域具有一定的实用价值。本文提出的基于功率谱特征的方法对 7 台同一型号同一批次的 Wi-Fi 路由器识别准确率能达到 97.3%, 对 27 台分属两种型号的路由器识别准确率超过 93%, 在设备数不大的情况下, 该方法的识别准确率要高于现有的其他方法。但本文对移动场景的研究还存在不足, 模拟的实验环境较为理想。在今后的工作中, 可进一步增加待识别设备的数量, 在更复杂的信道多变的移动环境中进行测试, 如考虑隔墙的情况, 对移动状态下的加速度、瞬时速度、平均速度等参数定量分析, 将外界因素干扰纳入考虑之中, 并研究如何消除残留多径的影响, 寻求特征区分度更大的射频指纹提取方法。

参考文献

[1] Hu A Q, Li G Y. Physical Layer Security in Wireless Communication: Survey[J]. *Journal of Data Acquisition & Processing*, 2014, 29(3): 341-350.
(胡爱群, 李古月. 无线通信物理层安全方法综述[J]. 数据采集与处理, 2014, 29(3): 341-350.)

[2] Zou Y L, Zhu J, Wang X B, et al. A Survey on Wireless Security: Technical Challenges, Recent Advances and Future Trends[EB/OL]. 2015: arXiv:1505.07919[cs.IT]. https://arxiv.org/abs/1505.07919

[3] Zhang K, Liang X H, Lu R X, et al. Sybil Attacks and Their Defenses in the Internet of Things[J]. *IEEE Internet of Things Journal*, 2014, 1(5): 372-383.

[4] Yu J B, Hu A Q, Zhou F, et al. Radio Frequency Fingerprint Identification Based on Denoising Autoencoders[EB/OL]. 2019: arXiv:1907.08809[eess.SP]. https://arxiv.org/abs/1907.08809

[5] Li G Y, Sun C, Zhang J Q, et al. Physical Layer Key Generation in 5G and beyond Wireless Communications: Challenges and Opportunities[J]. *Entropy*, 2019, 21(5): 497.

[6] Wang N, Wang P, Alipour-Fanid A, et al. Physical-Layer Security of 5G Wireless Networks for IoT: Challenges and Opportunities[J]. *IEEE Internet of Things Journal*, 2019, 6(5): 8169-8181.

[7] Zhang J Q, Rajendran S, Sun Z, et al. Physical Layer Security for the Internet of Things: Authentication and Key Generation[J]. *IEEE Wireless Communications*, 2019, 26(5): 92-98.

[8] Yuan H. Research on the Key Technologies of Wireless Network Physical Layer Authentication Based on RF Fingerprint[D]. Nanjing: Southeast University, 2011.
(袁红林. 基于射频指纹的无线网络物理层认证关键技术研究[D]. 东南大学, 2011.)

[9] Tian Q, Lin Y, Guo X H, et al. New Security Mechanisms of High-Reliability IoT Communication Based on Radio Frequency Fingerprint[J]. *IEEE Internet of Things Journal*, 2019, 6(5): 7980-7987.

[10] Yu J B, Hu A Q, Zhu C M, et al. RF Fingerprinting Extraction and Identification of Wireless Communication Devices[J]. *Journal of Cryptologic Research*, 2016, 3(5): 433-446.
(俞佳宝, 胡爱群, 朱长明, 等. 无线通信设备的射频指纹提取与识别方法[J]. *密码学报*, 2016, 3(5): 433-446.)

[11] Ali A M, Uzundurukan E, Kara A. Improvements on Transient Signal Detection for RF Fingerprinting[C]. *2017 25th Signal Processing and Communications Applications Conference*, 2017: 1-4.

[12] Yuan H L, Hu A Q. Preamble-based Detection of Wi-Fi Transmitter RF Fingerprints[J]. *Electronics Letters*, 2010, 46(16): 1165.

[13] Cui Z Y, Hu A Q, Peng L N. A Method of RF Fingerprint Recognition Based on Contour Feature[J]. *Netinfo Security*, 2017(10): 75-80.
(崔正阳, 胡爱群, 彭林宁. 一种基于轮廓特征的射频指纹识别方法[J]. *信息网络安全*, 2017(10): 75-80.)

[14] Peng L N, Hu A Q, Zhang J Q, et al. Design of a Hybrid RF Fingerprint Extraction and Device Classification Scheme[J]. *IEEE Internet of Things Journal*, 2019, 6(1): 349-360.

[15] Yu J B, Hu A Q, Li G Y, et al. A Robust RF Fingerprinting Approach Using Multisampling Convolutional Neural Network[J]. *IEEE Internet of Things Journal*, 2019, 6(4): 6786-6799.

[16] Zhou X, Hu A, Li G, et al. Design of a Robust RF Fingerprint Generation and Classification Scheme for Practical Device Identification[C]. *2019 IEEE Conference on Communications and Network Security*, 2019: 196-204.

[17] Robyns P, Marin E, Lamotte W, et al. Physical-layer Fingerprinting of LoRa Devices Using Supervised and Zero-shot Learning[C]. Proceedings of the 10th ACM Conference on Security and Privacy in Wireless and Mobile Networks, 2017: 18-20.

[18] Xing Y X, Hu A Q, Zhang J Q, et al. Design of a Robust Radio-Frequency Fingerprint Identification Scheme for Multimode

LFM Radar[J]. *IEEE Internet of Things Journal*, 2020, 7(10): 10581-10593.

[19] Ezuma M, Erden F, Kumar Anjinappa C, et al. Detection and Classification of UAVs Using RF Fingerprints in the Presence of Wi-Fi and Bluetooth Interference[J]. *IEEE Open Journal of the Communications Society*, 2020, 1: 60-76.

[20] Reising D R, Temple M A, Jackson J A. Authorized and Rogue Device Discrimination Using Dimensionally Reduced RF-DNA Fingerprints[J]. *IEEE Transactions on Information Forensics and Security*, 2015, 10(6): 1180-1192.

[21] Fadul M, Reising D, Loveless T, et al. RF-DNA Fingerprint Classification of OFDM Signals Using a Rayleigh Fading Channel Model[C]. *2019 IEEE Wireless Communications and Networking Conference*, 2019: 1-7.

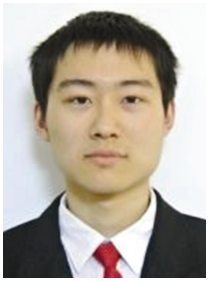
[22] Wheeler C, Reising D. Assessment of the Impact of CFO on RF-DNA Fingerprint Classification Performance[C]. *2017 International Conference on Computing, Networking and Communications*, 2017: 110-114.

[23] Zhou M, Arigye W, Tian Z S, et al. ScOFi: Schematic Assisted Optimum Fingerprinting for Wi-Fi Indoor Localization Using Peer Hand-shake[J]. *Physical Communication*, 2017, 25: 399-411.

[24] Zeng Y H, Chen X, Lin Y, et al. Review of Radio Frequency Fingerprinting Identification[J]. *Chinese Journal of Radio Science*, 2020, 35(3): 305-315.
(曾勇虎, 陈翔, 林云, 等. 射频指纹识别的研究现状及趋势[J]. *电波科学学报*, 2020, 35(3): 305-315.)

[25] Li G Y, Yu J B, Xing Y X, et al. Location-Invariant Physical Layer Identification Approach for WiFi Devices[J]. *IEEE Access*, 2019, 7: 106974-106986.

[26] ISO/IEC/IEEE International Standard - Information Technology —— Telecommunications and Information Exchange between Systems —— Local and Metropolitan Area Networks —— Specific Requirements —— Part 15-6: Wireless Body Area Network[S]. IEEE, . DOI:10.1109/ieeestd.2018.8323448



陈天舒 于 2018 年在南京理工大学电子信息工程专业获得学士学位。现在东南大学网络空间安全专业攻读硕士学位。研究领域为物理层安全。研究兴趣包括: 物联网安全、机器学习。Email: iamtianshu@seu.edu.cn

姜禹 于2009年在东南大学信号与信息处理专业获得博士学位。现为东南大学网络空间安全学院副教授。研究领域为物理层安全、无线网络安全、RFID 技术、物联网技术。Email: jiangyu@seu.edu.cn

胡爱群 于 1992 年在东南大学信号与信息处理专业获得博士学位。现为东南大学信息科学与工程学院、移动通信国家重点实验室教授/博导。研究领域为物理层安全、无线通信安全。Email: aqhu@seu.edu.cn