

用于无线设备身份识别的射频指纹提取方法

周樾*, 唐丹红, 蔡阳, 徐之欣

(国网上海金山供电公司, 上海 200540)

摘要: 已有的射频指纹技术缺乏理论支撑, 无法摆脱对设备身份识别过程中借助信道特征的依赖。该文针对直接序列扩频的 ZigBee 信号, 提出了一种基于扩频还原的射频指纹提取方法。在 60 台 ZigBee 设备上以及 4 种不同的实验场景中进行试验, 设备分类准确率最高可达 99.8%。该方法的鲁棒性更强、计算量更小, 具有较强的实用性。

关键词: 无线物理层安全; 射频指纹; 扩频还原; 身份识别

中图分类号: TM425

文献标识码: A

RF fingerprint extraction method for wireless device identity recognition

ZHOU Xi, TANG Danhong, CAI Yang, XU Zhixin

(State Grid Shanghai Electric Power Company, Shanghai 200540, China)

Abstract: The existing radio frequency fingerprinting technique lacks theoretical support and cannot get rid of the suspicion of identifying the process of device identification with the help of channel features. In this paper, a radio frequency fingerprint extraction method based on spread spectrum reduction is proposed for ZigBee signals with direct sequence spread spectrum. The method achieves device classification accuracy up to 99.8% on 60 ZigBee devices and in four different experimental scenarios. The method is more robust and less computationally intensive, making it highly practical.

Keywords: wireless physical layer security; radio frequency fingerprint; spread spectrum restoration; identity recognition

0 引言

随着物联网和第五代通信技术的不断发展, 无线物联网设备已广泛应用于多个领域, 如智慧城市、智能家居、智慧医疗等, 极大地便利了人们的工作与生活^[1]。由于低成本的特点, 物联网设备的安全策略受到诸多限制。射频指纹反映了设备制造时射频模块的硬件差异, 可作为设备合法身份的识别依据。射频指纹具有唯一性、不可克隆性和长期稳定性等优点^[2], 在物联网安全和应用中发挥着不可替代的作用。

不同类型的信道衰落加大了从接收信号中提取射频指纹的难度^[3]。目前, 针对抗信道干扰的射频指纹技术分为两种, 一种是提取与信道无关的设备特征, 另一种是研究抑制或去除信道影响的算法。与信道无关的设备特征有载波频率偏移^[4]、I/Q 直流偏置^[5]等, 然而这些特征值在较大范围内变化, 无法达到一个稳定状态。参考文献[6]利用 Wi-Fi 两个前导符号频谱相除的方式完美削弱信道的影响, 但导致了射频指纹特征大部分被消除, 引起

识别准确度大幅下降。

1 调制技术

ZigBee 调制技术基于 IEEE 802.15.4 协议, 在中国采用工作在 2.4 GHz 频段上的偏移正交相移键控 (Offset Quadrature Phase Shift Keying, OQPSK) 调制, 共需要经过 3 个步骤: ①比特到符号映射。原始二进制数据流按照十六进制表示的方式, 每 4 比特为一组编码, 为十六进制数, 从小到大分别对应 16 个不同的 ZigBee 数据符号。②符号到码片映射。将 16 个 ZigBee 数据符号映射到 16 个不同的伪随机噪声 (Pseudorandom Noise, PN) 序列, 每个 PN 序列由 32 个码片组成。③OQPSK 调制。将 32 码片的 PN 序列进行分组, 按照原来的顺序, 奇、偶数下标分别为对一组载波进行调制, 每个码片利用半正弦波成型脉冲 $v(t)$ 表示, 表达式为:

$$v(t) = \begin{cases} \sin\left(\frac{\pi}{2T_c}t\right), & 0 \leq t \leq 2T_c \\ 0, & \text{其他} \end{cases} \quad (1)$$

周樾(工程师), 主要从事电力调度自动化和网络安全研究。

* 通信作者, E-mail: 3487787117@qq.com。

式中, T_c 表示半个码片的持续时间, $\sin$ 表示正弦函数。码片“1”表示为 $v(t)$, 码片“0”表示为 $-v(t)$ 。将奇数下标的码片映射到同相(In-phase, I)路载波, 偶数下标映射到正交相(Quadrature, Q)路载波。

2 射频指纹提取

基于扩频还原的射频指纹提取算法流程图如图 1 所示, 可分为 3 大模块: 预处理模块、扩频还原模块和射频指纹提取模块。

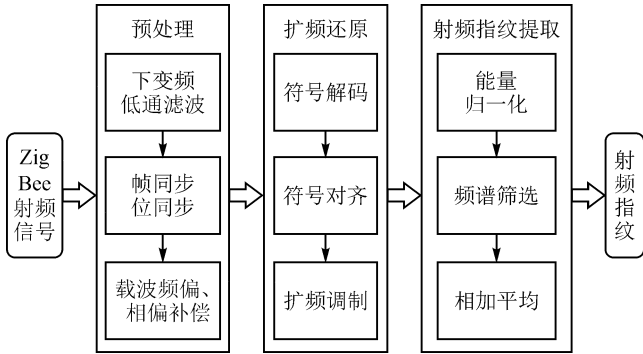


图 1 基于扩频还原的射频指纹提取算法流程图

Fig. 1 Flow chart of RF fingerprint extraction algorithm based on spread spectrum restoration

2.1 扩频还原

预处理模块分为 3 个步骤, 包括下变频低通滤波、帧同步位同步和载波频偏、相偏补偿。ZigBee 射频信号经过预处理模块后得到基本同步好的基带信号, 以便进行后续的符号解码、符号对齐和扩频调制操作。

(1) 符号解码

符号解码等价于从数据符号映射到已调信号的逆过程。在 ZigBee 调制过程中, 16 个数据符号对应于 16 个不同的 PN 序列, 也就对应于 16 个不同的时域波形。由于不同 PN 序列之间的正交性, 只有相同 PN 序列对应时域波形的互相关值最大, 因此利用这个基本思想即可实现对 ZigBee 符号的解码, 用公式表示为:

$$SSC_i = \arg\max_{t \in T} \left(\text{abs} \left(\sum_{n=1}^L s_i(n) \times r_t^*(n) \right) \right) \quad (2)$$

式中, $s_i(n)$ 表示信号帧内第 i 个符号, SSC_i 表示 $s_i(n)$ 的解码结果, $r_t(n)$ 表示编号为 t 的理想符号, L 表示一个符号长度, T 表示所有理想符号构成的编号集合, $*$ 代表复数的共轭, $\times$ 代表复数乘法, $\text{abs}()$ 表示复数取模值运算。

(2) 符号对齐

同步好的基带信号根据同步点与符号长度即可对整个信号帧进行划分, 确定每个符号准确的位置。然而, 由于物联网设备的低成本性, 发射设备晶振精密度不高导致时钟的振荡周期发生变化。每个符号的持续时间实际上不是一个固定值, 会随着设备的工作状态发生微小波动。

在采样率确定的情况下, 这种波动通常只会表现为一两个采样点的随机波动。因此考虑以同步点为参考点, 将参考点连同前后两个采样点作为可能的起始点进行比较, 计算符号互相关值最大的采样点作为当前符号的起始点, 以此来估计下一个符号可参考的起始点。按照这样的操作对整个信号帧从前向后依次遍历, 实现符号对齐。

(3) 扩频调制

经过符号解码与对齐的接收信号 $y(n)$ 可表示为:

$$y(n) = [s_1(n), s_2(n), \dots, s_N(n)] \quad (3)$$

式中, N 表示接收信号内的符号总数。根据 ZigBee 扩频调制技术原理, 每个符号 $s_i(n)$ 通过基带信号调制而成, 表达式为:

$$s_i(n) = x_i(n) \diamond P_i(n) = x_i^I(n) P_i^I(n) + j x_i^Q(n) P_i^Q(n) \quad (4)$$

式中, $x(n)$ 表示扩频调制前的基带载波信号, $\diamond$ 表示直接序列扩频调制, $x_i(n)$ 表示 $x(n)$ 中的第 i 个符号, $x_i^I(n)$ 表示 $x_i(n)$ 的 I 路, $x_i^Q(n)$ 表示 $x_i(n)$ 的 Q 路, $P_i(n)$ 表示调制到 $x_i(n)$ 的第 i 个 PN 序列, $P_i^I(n)$ 表示 $P_i(n)$ 的实部, $P_i^Q(n)$ 表示 $P_i(n)$ 的虚部。

根据符号解码的结果 SSC_i 确定每个符号对应 PN 序列的估计结果 $\tilde{P}_i(n)$, 再将该结果调制到每个符号上, 得到扩频还原后的符号结果 $\tilde{s}_i(n)$, 具体表达式为:

$$\begin{aligned} \tilde{s}_i(n) &= s_i(n) \diamond \tilde{P}_i(n) \\ &= s_i^I(n) \tilde{P}_i^I(n) + j s_i^Q(n) \tilde{P}_i^Q(n) \\ &= x_i^I(n) P_i^I(n) \tilde{P}_i^I(n) + j x_i^Q(n) P_i^Q(n) \tilde{P}_i^Q(n) \end{aligned} \quad (5)$$

式中, $\tilde{P}_i^I(n)$ 表示 $\tilde{P}_i(n)$ 的实部, $\tilde{P}_i^Q(n)$ 表示 $\tilde{P}_i(n)$ 的虚部。当 PN 序列估计准确时, 有:

$$\begin{cases} P_i^I(n) \tilde{P}_i^I(n) = E(n) \\ P_i^Q(n) \tilde{P}_i^Q(n) = E(n) \end{cases} \quad (6)$$

式中, $E(n)$ 表示元素全为 1 的序列, 那么式(5)可进一步表示为:

$$\tilde{s}_i(n) = x_i^I(n) E(n) + j x_i^Q(n) E(n) = x_i^I(n) + j x_i^Q(n) \quad (7)$$

即每个符号 $s_i(n)$ 被还原成未被扩频调制前的基带载波符号 $x_i(n)$ 的估计结果 $\tilde{s}_i(n)$ 。

2.2 射频指纹提取

对接收信号扩频还原的操作是在时域完成, 还需将其变换到频域进行最终的射频指纹提取, 包括能量归一化、频谱筛选和相加平均。

(1) 能量归一化

本文以一个符号为单位进行射频指纹的提取, 能量归一化运算具体为将一个符号除以该符号能量的均方根值, 用公式表示为:

$$\tilde{s}'_i(n) = \frac{\tilde{s}_i(n)}{\sqrt{\frac{e[\tilde{s}_i(n)]}{L_{\tilde{s}_i(n)}}}} \quad (8)$$

式中, $e[\tilde{s}_i(n)]$ 表示 $\tilde{s}_i(n)$ 的能量, $L_{\tilde{s}_i(n)}$ 表示 $\tilde{s}_i(n)$ 的符号长度。

(2) 频谱筛选

理想的基带信号是由半正弦波成型脉冲 $v(t)$ 拼接而成, 变换到频域后的结果如图 2 所示。可以观察到理想的基带载波仅包含有限个频谱值不为零的频点, 这些频点是有效频点。提取这些频点作为每个接收符号幅度谱的参考频点, 便于后续实际接收信号进行参考。

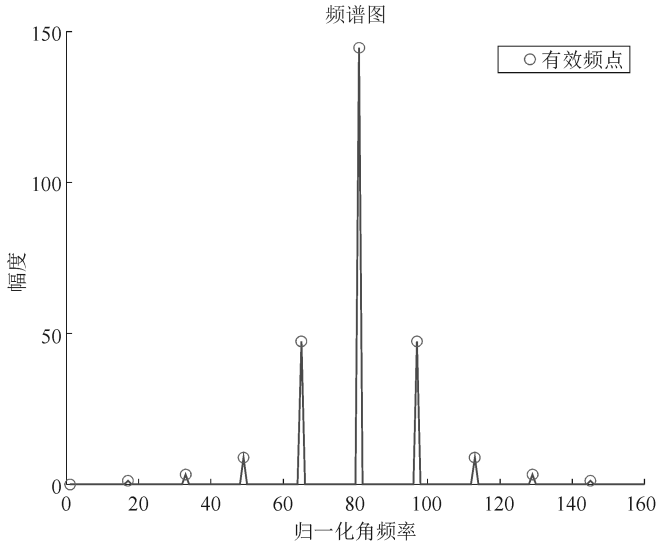


图 2 理想基带载波有效频点示意图

Fig. 2 Schematic diagram of effective frequency points of ideal baseband carrier

(3) 相加平均

从实际接收信号的每个符号中都可以提取出相应的特征序列作为一个符号的射频指纹, 然而由于射频指纹表征了设备细微的特征, 体现在每个符号上不会是一个固定的序列。类似于方差较小的高斯分布, 射频指纹体现在每个符号上时呈现一定程度的波动。此外, 一个信号帧的射频指纹应该与其中一个符号上的射频指纹相似, 因此, 将一帧信号内所有符号提取出来的特征序列叠加作为接收信号的射频指纹, 表达式为:

$$RFF = \frac{\sum_{i=1}^N RFF_i}{N} \quad (9)$$

式中, N 表示一帧信号内的符号总数, RFF_i 表示第 i 个符号的射频指纹, RFF 即为一帧信号的射频指纹表示。叠加的方式不仅可以消除轻微波动的影响, 还可以实现在一帧内提取出射频指纹, 更符合实际需求。

2.3 抗信道分析

针对式(5)阐述的扩频还原操作的抗信道性能, 将信道响应与噪声同时考虑进行推导。实际接收到的基带扩频信号 $y(n)$ 表示为:

$$y(n) = x(n) \diamond P(n) \otimes h(n) + g(n) \quad (10)$$

式中, $h(n)$ 表示信道冲激响应, $g(n)$ 表示加性噪声。在信道比较高时, $g(n)$ 可忽略, 那么进行扩频还原后的信号 $\tilde{y}(n)$ 可表示为:

$$\begin{aligned} \tilde{y}(n) &= y(n) \diamond \tilde{P}(n) \\ &= x(n) \diamond P(n) \otimes h(n) \diamond \tilde{P}(n) \\ &= x(n) \diamond P(n) \diamond \tilde{P}(n) \otimes h(n) \\ &= x(n) \otimes h(n) \end{aligned} \quad (11)$$

式中, $\otimes$ 表示卷积运算, 再将 $\tilde{y}(n)$ 变换到频域, 有:

$$\begin{aligned} \tilde{Y}(k) &= FFT(\tilde{y}(n)) \\ &= FFT(x(n) \otimes h(n)) \\ &= X(k) \times H(k) \end{aligned} \quad (12)$$

式中, $X(k)$ 表示 $x(n)$ 的频域变换结果, $H(k)$ 表示 $h(n)$ 的频域变换结果, FFT 表示傅里叶变换。由于基带载波信号 $x(n)$ 是单频正弦波, 因此在频谱中 $X(k)$ 仅有限个频率分量。复杂信道 $H(k)$ 对 $X(k)$ 的影响体现为幅度上的衰减, 通过能量归一化后可有效去除信道干扰。

3 实验与分析

3.1 ZigBee 数据集

为了验证本文提出的扩频还原方法对抗复杂无线信道影响的有效性, 使用公开的 ZIGBEE_RFF_DATASET 数据集^[7]。该数据集在 4 种不同场景下采集了 60 台 ZigBee 设备大量的无线信号, 几乎覆盖了所有的通信场景, 包括室内视距、室内非视距、室外视距和室外非视距。多种信道类型反映了不同的信道干扰影响, 可对本文方法抗信道的鲁棒性进行有效验证。

3.2 数据处理

首先, 根据图 1 所示的流程图对 ZIGBEE_RFF_DATASET 数据集进行预处理操作, 接着按照本文论述的方法将每个场景下所有设备的信号帧实行扩频还原, 最后将一帧信号内所有符号的射频指纹叠加取平均, 得到每一帧信号的射频指纹。以室内视距场景下的两个随机 ZigBee 设备为例展示提取到的射频指纹。ZigBee 每个数据符号持续时间为 16 μs , 在采样率为 10 Msps 下, 一个符号相应的采样点数为 160。通过频谱筛选后, 有效频点的数目为 10 个, 因此, 一个数据符号(也就是一帧信号)可提取 10 维的特征值作为射频指纹, 如图 3 所示。

比较图 3(a)和图 3(b), 可清晰地看出同一个设备从符号提取出的射频指纹存在一定波动, 特征呈现一定的发散。通过将一个信号帧内所有符号的射频指纹叠加取平均后, 同一个设备多个信号帧的射频指纹就比较稳定, 类内方差更小, 更加能体现设备本身的身份特征。从图 3 可以看出, 提取出的射频指纹通过肉眼就可区分, 说明本文

方法提取出的射频指纹区分度较高,因此采用二次判别分

类器进行训练与测试。

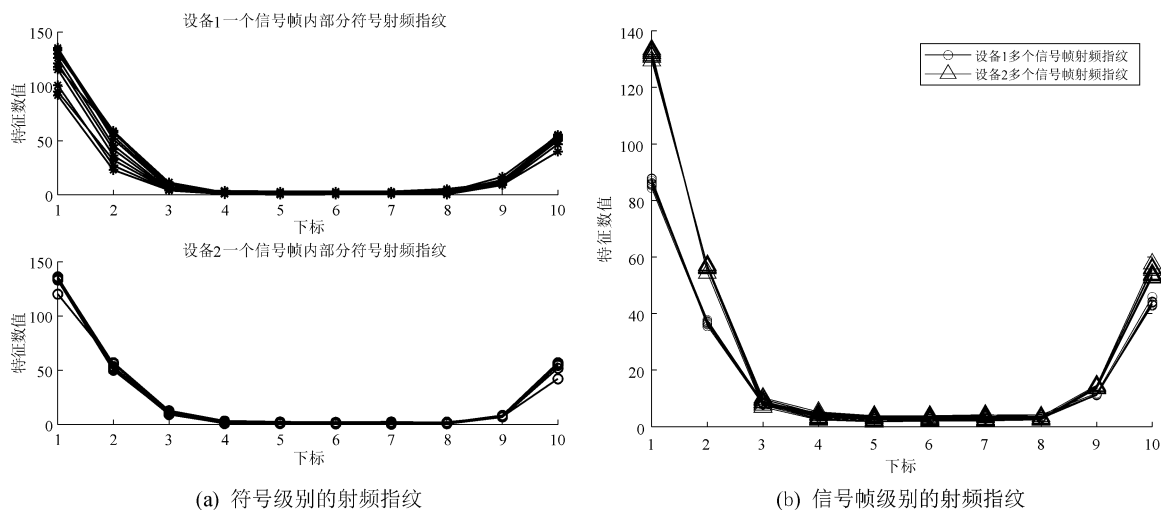


图3 两个 ZigBee 设备的射频指纹表示图

Fig. 3 RF fingerprint representation of two ZigBee devices

3.3 测试结果

将所有场景下的数据帧转换为 10 维特征值和设备标签的组合,即 11 维数据送入分类器训练。将所有数据集按照 80%作为训练集、20%作为测试集进行。为了解决训练类别数目之间不匹配的问题,所有设备都按照此比例划分,最终的测试结果如表 1 所列。

表 1 4 种场景训练与测试准确率

Table 1 Training and testing accuracy for four scenarios

场景	训练准确率/%	测试准确率/%
室内视距	100	99.6
室内非视距	99.9	99.8
室外视距	100	99.4
室外非视距	99.8	99.7

在 4 种不同场景下的测试准确率都较高,证明本文方法对抗不同类型信道影响效果显著,不仅去除了设备特征中的信道信息,还抑制了信道对射频指纹的影响,效果较好。

4 结 语

本文提出了一种基于扩频还原的射频指纹提取方法,通过简单的二次判别分类器实现了对 4 种不同场景下 60 台 ZigBee 设备的身份识别与认证。通过对实际信号数据的测试,准确率最高可达 99.8%。本文方法在对抗信道干扰方面具有较高的鲁棒性,可应对各种复杂应用环境的物联网设备的射频指纹提取,具有一定的实用性。

参考文献

[1] MEHMOOD Y, AHMAD F, YAQOOB I, et al. Internet - of -

things - based smart cities: Recent advances and challenges [J]. IEEE Communications Magazine, 2017, 55(9): 16 - 24.

[2] 袁红林, 胡爱群. 射频指纹的产生机理与唯一性[J]. 东南大学学报(自然科学版), 2009, 39(2): 230 - 233.

YUAN H L, HU A Q. The mechanism and uniqueness of radiofrequency fingerprints[J]. Journal of Southeast University (Natural Science Edition), 2009, 39(2): 230 - 233 (in Chinese).

[3] 王栋, 胡爱群, 王炎. 多径传输环境下的射频指纹估计方法[J]. 密码学报, 2020, 7(2): 249 - 260.

WANG D, HU A Q, WANG Y. RF fingerprint estimation method in multipath transmission environment[J]. Journal of Cryptography, 2020, 7(2): 249 - 260 (in Chinese).

[4] JIANG Y, PENG L, HU A, et al. Physical layer identification of LoRa devices using constellation trace figure[J]. EURASIP Journal on Wireless Communications and Networking, 2019(1).

[5] PENG L, HU A, ZHANG J, et al. Design of a hybrid RF fingerprint extraction and device classification scheme[J]. IEEE Internet of Things Journal, 2018, 6(1): 349 - 360.

[6] 陈天舒, 胡爱群, 姜禹. 基于功率谱特征的 Wi - Fi 射频指纹提取方法[J]. 信息安全学报, 2021, 6(2).

CHEN T SH, HU AI Q, JIANG Y. Wi Fi RF fingerprint extraction method based on power spectrum features[J]. Journal of Information Security, 2021, 6(2) (in Chinese).

[7] SHI J, PENG L, FU H, et al. Robust RF Fingerprint Extraction Based on Cyclic Shift Characteristic[J]. IEEE Internet of Things Journal, 2023.

(责任编辑: 薛士然 收稿日期: 2023-08-17)