 DOI: https://doi.org/10.26549/bdai.v3i1.11439

Fingerprint Recognition of Wireless Communication Devices Based on Deep Learning

Zhiyuan Yu

Beijing Broadthinking Technology Limited, Beijing, 100047, China

Abstract

Intercepting and collecting the transmitted signals of wireless devices and extracting the radio frequency fingerprints of the device from these transmitted signals, which are widely used in the identification of the device, effectively protecting the technical and physical layer means of communication security. RF fingerprint is a physical layer feature in various wireless communication devices and devices, which has the advantages of uniqueness, short-term stability, and independence. Just like different people will have different gestures and fingerprints, different wireless communication devices will also have different radio frequency fingerprints, such as widely used in the identification and certification of equipment, which has important scientific theoretical significance and practical use value.

Keywords

RF fingerprint; device identification; physical layer security

基于深度学习的无线通信设备指纹识别

於志渊

北京博思汇众科技有限公司，中国·北京 100047

摘　要

截取和采集无线设备的发送信号并从这些发送信号中提取出该设备的射频指纹，将其广泛应用于对设备的识别，有效地保护了通信安全的技术和物理层手段。射频指纹是各种无线通信装置和设备中的一种物理层特征，其具有唯一性、短时稳定性、独立性等优势。就好像不同的人都会拥有不一样的手势和指纹，不一样的无线通信设备也会拥有不一样的射频指纹，如广泛地应用于对设备的识别与认证，具有重要的科学理论意义和实际的使用价值。

关键词

射频指纹；设备识别；物理层安全

1 引言

随着中国互联网技术的高速进步与发展以及通信装置的不断增多，无线通信不仅体现在了军事领域，还在民用领域发挥了一个不可取代的重要作用。然而，无线网络因为其本身具有的开放性，相对传统的有线通信网络来说，它更容易被用户入侵和遭受攻击。

目前使用最多的保护无线通信安全的方案，是通过使用基于密码体系的安全协议来实现对数据机密性和完整性的保护。然而，目前的篡改、伪造软硬件信息等技术方案都已经比较成熟。另外，一旦密钥被泄露，现有的安全防护机制将导致不能做到对个人身份认证。因此，现在迫切地寻找一种新型的安全机制有效地识别用户的身份，以减少来自不同恶意用户的潜在风险。

【作者简介】於志渊（1977–），男，中国上海人，硕士，从事智能卡、安全芯片、安全通信、物联网解决方案、边缘计算、云计算及虚拟化等研究。

目前提出的一种可行性较高的方法是将物理层特征作为射频指纹应用于设备的识别。在无线网络通信过程中，采集系统采集到的设备信号除了发射机信号的特征之外还存在着传输信道所带来的影响。传输信道带来的影响主要有设备处于不同位置所带来的多径效应、设备移动带来的信号干扰等。而无线通信设备产生的可作为射频指纹特征的主要来源于电子元器件存在的微小差异，这部分差异是射频指纹存在的关键，例如元器件生产过程中的工艺技术、发送和接收天线实际方向的差异、供电电源的特性等。正因为有了这些差异，使得射频指纹识别技术的研究有了根本的保障。待到未来的射频指纹研究方案成熟之后，它将能有效抵抗目前已有和正在发展的伪造和欺骗技术 [1]。

2 射频指纹技术原理

2.1 射频指纹的产生机理

射频指纹（radiofrequencyfingerprint，RFF）的主要来源是无线发射机的电子元件容差，电子元件的容差分为制造容差和漂移容差。其中，设备的材料和加工工艺的生产

过程中产生的元器件电参数与标称值的误差被称为制造容差。而漂移误差是指元器件因长时间工作导致的器件参数退化。无论是什么类型的电路，其组成必然存在电子元器件，而电子元器件的容差正是射频指纹存在的根本之处，也正是因为有这些容差才使得每个设备的射频指纹是独一无二的。

2.2 射频指纹的基本特征

射频指纹是采集无线通信设备在通信过程中的发射机信号并提取出一种或多种特征向量集合，将其应用于识别发射源。在实际应用中，应具有如下特点。

2.2.1 通用性

通用性是指绝大多数的无线通信设备都可以通过采集其发送信号来提取射频指纹，不会受到发送环境、设备的系统结构等影响。

2.2.2 唯一性

正如不同的人有不同的指纹，不一样的通信设备也有不一样的射频指纹。射频指纹能够区分开不同的发射机，即使是同一类型同一批次的发射机也能找到或大或小的差异。但是随着科技的高速发展，电子元件之间的集成度越来越高，再加上噪声的影响，其区别也会变得更小。所有接下来如何保证射频指纹的唯一性是未来的一个重要议题。

2.2.3 短时不变性

设备长时间使用后会存在器件老化、各项参数退化等问题，从而会导致与射频指纹库中登记的有所不同。但是元器件的老化是一个漫长的过程，因此射频指纹需要保持着短时不变性。

2.2.4 独立性

射频指纹仅仅与一个发射器的硬件功能特性相关，并且不会直接受到信号调制的样式及其传输内容的影响。

2.2.5 稳健性

多径信道、接收设备距离、噪声干扰、电压及温度的改变都会给射频指纹提供一定的影响。射频指纹应该是能够有效地对抗这些要素的作用，稳健地进行射频指纹技术应用，从而能够促进射频指纹识别系统的发展。

2.3 射频指纹识别系统框架

如图 1 所示，无线通信设备的射频指纹提取与识别流程一般包括五部分，分别是：收集信号、信号预处理、特征提取、分类和识别。在收集信号方面首先要确定研究的设备，原则上来说只要是能发射出信号的设备都能成为研究对象，在收集信号的过程中会受到周围环境的干扰，这些影响或多或少会对研究成果造成困扰，所以要尽可能地减少多径信道的影响。收集到信号后要对信号进行预处理才能到提取特征，如要对信号进行降噪、降维归一化等。接下来对信号进行特征提取，在可识别信号方面分为两部分，一部分是基于瞬态信号，另一部分是基于稳态信号。不同的信号可提取的特征各不相同。然后就是分类器的选取，在这方面可用基于传统机器学习的分类方法也可以用基于深度学习的分类方法。分类过程中会将发射机设备和自己的射频指纹链接在一起放入射频指纹库。最后则是一个识别的环节，把待识别设备的射频指纹提取出来然后将其与射频指纹库对比并得出结果 [2]。

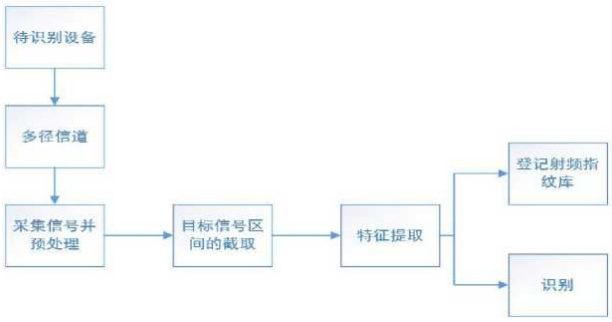

图 1 射频指纹识别系统框架

3 基于卷积神经网络的射频指纹识别

基于卷积神经网络的射频指纹识别，首先建立好模型后，导入数据集作为输入，并产生损失和准确率，然后通过调整不同的参数来更好的训练数据集，使其达到最佳的训练准确率和测试准确率。

3.1 数据集介绍

论文使用 16 个 X310USRP 无线电平台组成发射机，以固定的 USRPB210 作为接收器。发射机可发射 IEEE802.11aWiFi 信号帧，这些帧通过 MATLABWLAN 系统工具箱生成，生成的数据帧虽包含随机有效载荷，但都具有相同的地址字段。传播介质为空气，接收器以 5MS/s 的采样率对传入的 WiFi 信号进行采样，对 WiFi 信号的采样中心频率为 2.45GHz。图 2 展示了收集信号的过程。

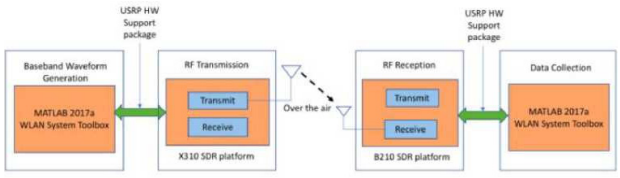

图 2 收集数据流程

该文章在一个反射较少的开放区域进行实验共采集了 11 组数据，为每台设备至少收集了 2000 万个样本，每组数据的区分在于发射机和接收器之间的距离，从相离 2 英尺开始每次增加 4 英尺，依此类推一直增加到 62 英尺。采集到的原始数据被分类到文件名为“xxft”的不同文件夹，其中 xx 表示以英尺为单位的发射机和接收机相隔的距离，每个数据都有一个扩展名为“.sigmf-data”的数据集文件以及一个扩展名为“.sigmf-meta”的元数据文件，元数据文件包含描述数据集的信息。图 3 表明了该实验所在的区域。论文采用的数据是该文章的第四组数据，发射机和接收机相离 14 英尺，所属的文件名为“14ft”。该组数据集一共

包含 78144 个数据样本，在进行数据随机分割之后选择其中 70329 个样本作为训练集并且在每一轮的训练中从中选择 7815 个样本作为验证集，其余的 7815 个样本作为测试集。

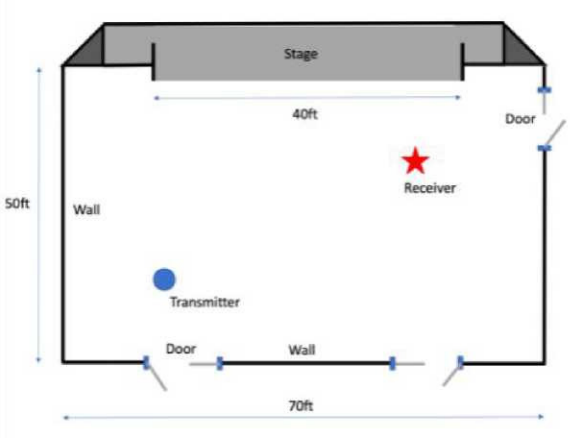

图 3 数据采集环境

3.2 实验设计框架

如图 4 所示，论文的设计框架分为五个部分，分别是：收集数据集、对数据进行降维归一化预处理、搭建神经网络、模型训练以及识别。在收集数据部分，论文采用的是来自甲骨文《通过卷积神经元算法进行优化的辐射辅助》的一组数据；在预处理部分，论文采用两种归一化方法，分别是：最小最大尺度算法和标准比例尺算法；在搭建神经网络部分，论文采用四层层次搭建，分别是两个卷积层和两个全连接层，第一个卷积层由 50 个滤波器组成，每个滤波器大小为 1×7，第二个卷积层同样由 50 个滤波器组成，每个滤波器大小为 2×7，第一个全连接层由 256 个神经元组成，第二个全连接层由 80 个神经元组成，且四层都使用“relu”激活函数，输出层使用“软最大值”分类函数 [3]。

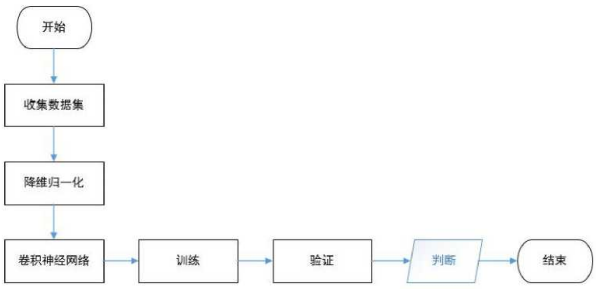

图 4 实验设计框架

3.3 实验结果

论文得出的培训和验证损失的结果如下：

CNN_14ft_training 和验证损失（见图 5）。

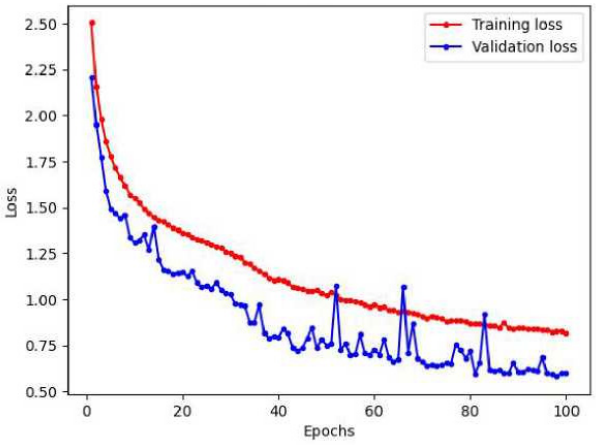

图 5 培训和验证损失

论文得出的培训和验证图如下：培训项目、验证 accCNN_14ft_training 和验证 accEpochs100。图 6 为培训和验证图。

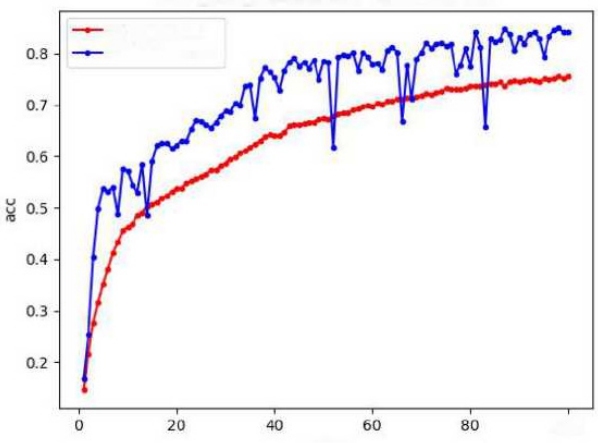

图 6 培训和验证图

如图 7 所示，论文得出训练准确率为：0.8709493949923536；得出的测试准确率为：0.8427383237516582。

列车损失：0.5254205553004563；

列车精度：0.8709493949923536；

测试损失：0.6019464090247224；

测试准确度：0.8427383237516582。

本节首先说明本次实验使用的数据集，在改变输入参数之后放入卷积神经网络进行训练，在此期间遇到很多问题，例如论文一开始将 batch_size 设置为 32，当神经网络训练到 67 轮时出现了 OOM 问题，内存使用完了。本节改变 batch-size 的大小为 16，解决了这个问题。最终成功得出训练准确率为 87%，测试准确率为 84.2%，当然本次实验还存在着很多不足，例如 epochs 的轮数较少，倘若训练到四五百轮，也许准确率能够达到 90% 以上，还有搭建的卷积神经网络规模不够大，这些问题也给出了后续研究的方向。

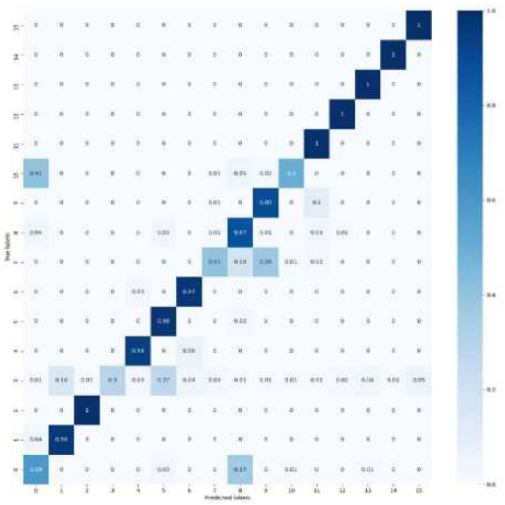

图 7 confusion matrix 图

4 结语

论文利用框架对收集到的数据集进行训练，使其达到较好的识别率。首先论文先研究了无线通信设备射频指纹提取与识别，接下来查阅 GitHub 开源、相关论文以及各大数据集网址后确定下论文所研究的数据集，该数据集由 16X310USRP 无线电平台作为发射机，产生的数据一共有 11 组，论文选取其中一组作为研究的数据集。然后选取基于 Kears 框架的 CNN 模型作为论文的深度学习算法并且介绍了该 CNN 模型所用到的相关算法。在训练过程中论文还不断的更改优化器、更改归一化算法和更改一些必要的参数来产生结果进行对比。论文虽然有了部分成果，但仍然存在很多问题，需要进一步地进行探讨，以下给出一些未来方向的展望：

①受限于知识的薄弱和高精度的设备，论文未能自己收集信号，而是寻找的相关数据集，后续应能够自己搭建环境来采集信号等一系列研究。

②论文只研究了 CNN 模型，后续应研究不同的深度学习模型，比较不同的深度学习模型下训练相同的数据集有什么不一样的区别。

③论文得出的实验准确率未能突破 90 大关，后续应继续优化神经网络训练模型，将其进一步提高准确率。

参考文献

[1] R Jones. Most Secret War[M]. London: Hamish Hamilton,1978.

[2] Z Li, W Xu, R Miller,et al. Trappe, Securing wireless systems via lower layer enforcements[J]. 5th ACM Workshop Wireless Secur.,2006(3):33-42.

[3] R Zhang, L Song, Z Han, et al. Physical layer security for two way relay communications with friendly jammers[J].IEEE Global Telecommun. Conf. (GLOBECOM),2010(6):1-6.