Uncovering the Portability Limitation of Deep Learning-Based Wireless Device Fingerprints

Bechir Hamdaoui and Abdurrahman Elmaghbub
School of Electrical Engineering and Computer Science, Oregon State University, Corvallis, Oregon, USA
{hamdaoui,elmaghab}@oregonstate.edu

Abstract—Recent device fingerprinting approaches rely on deep learning to extract device-specific features solely from raw RF signals to identify, classify and authenticate wireless devices. One widely known issue lies in the inability of these approaches to maintain good performances when the training data and testing data are collected under varying deployment domains. For example, when the learning model is trained on data collected from one receiver but tested on data collected from a different receiver, the performance degrades substantially compared to when both training and testing data are collected using the same receiver. The same also happens when considering other varying domains, like channel condition and protocol configuration. In this paper, we begin by explaining, through testbed experiments, the challenges these fingerprinting techniques face when it comes to domain portability. We will then present some ideas on how to go about addressing these challenges so as to make deep learning-based device fingerprinting more resilient to domain variability.

Index Terms—Device fingerprinting portability, deep learning, domain-adaptation, hardware impairments.

I. INTRODUCTION

Recently, there has been considerable interest in adopting deep learning-enabled device fingerprinting in automated network authentication mechanisms for emerging large-scale wireless devices (e.g., 6G, IoT, vehicular, etc.) [1], [2]. In essence, device fingerprinting relies on deep learning techniques to extract device-specific features and signatures, solely from raw RF (radio frequency) signals, that can be used to uniquely identify devices. Deep learning is capable of automatically processing raw RF signals and exploiting random distortions in the received RF signals that are caused by the devices’ hardware impairments to extract signatures that are unique to the devices [3]–[5]. These hardware impairments occur during manufacturing of the transmitters’ various hardware components, including the local oscillator, the power amplifier, the mixer, and others. More detailed explanation and analysis of various hardware impairments can be found in [5], [6].

II. FINGERPRINT PORTABILITY CHALLENGES

Deep learning-based device fingerprinting approaches have indeed achieved good performance results when both training and testing data are collected under the same channel conditions and network deployment configurations. However, they perform poorly when the data used during training and that used during testing are collected under varied network deployment setting, where a deployment setting could refer to a channel condition, a hardware receiver, a protocol configuration, a physical location, etc. Each of these are often referred to as domain, and the problem being raised is referred to as fingerprint portability or as domain-adaptation in the machine learning community. We will next demonstrate how the limited portability of fingerprints can impact device fingerprinting accuracy, and we do so using testbed experiments.

A. Testbed and Data Collection Setup

To explain these challenges, we used our IoT fingerprinting testbed [7] to run several experiments under different varied domains, by training and testing the deep learning models on data collected on different days, using different receivers, and/or under different protocol configurations. The IoT testbed, shown in Fig. 1, consists of 60 IoT/Pycom transmitters and 2 USRP B210 receivers, and is used to collect signal data under multiple different deployment settings, including indoors, outdoors and others. Each Pycom device was connected to a dedicated antenna and was configured to send LoRa packets at 125KHz bandwidth in the 915MHz frequency, with a spreading factor set to 7, a preamble set to 8, and a transmit power level set to 20dBm. For more detailed description of the testbed, experimental scenarios, and LoRa protocol specifications, see [5], [7].

B. Results Illustrating Fingerprint Portability Challenges

Fig. 2 demonstrates the impact of the poor fingerprint portability on the achieved accuracy of the learning models under three different varied domains: channel variability, receiver hardware variability, and protocol configuration variability. Specifically, Fig. 2a reveals the sensitivity of fingerprinting accuracy to channel variability, i.e., when training data and testing data are captured on different days. Fig. 2b shows its sensitivity to receiver hardware variability, i.e., when training data and testing data are collected using different
We now highlight two techniques we proposed to (potentially) match one another.

**Domain portability challenges when changing domain**

- **Time:** The domain being channel condition, receiver hardware, or configuration. The figures clearly show the significant degradation in accuracy due to a change in the domain, whether the domain is time, receiver, or protocol configuration variability, i.e., when training data is collected when these devices use a different protocol configuration, and that fingerprinting performs well only when the training and the testing domains (day, receiver, or protocol) match one another.

**III. LIMITATIONS OF PRIOR SOLUTION ATTEMPTS**

There have been some mad efforts aimed at addressing the fingerprint portability challenges, ranging from adding artificial impairments to increase device separability [1], to data augmentation to generalize the model to adapt to domain variation [2], to modifying transmitted signals to compensate for channel condition variations [9], to using MIMO capability to overcome channel fading [10]. The key issues with these proposed data augmentation approaches is that they require significant amounts of new labeled data, as well as re-training of the models with these new data. The other proposed approaches require modification at the transmitter side and are vulnerable to wireless channel conditions, as some require that receivers send their feedback to the transmitters in order for the transmitters to be able to modify their signals.

**IV. POTENTIAL SOLUTION IDEAS**

We now highlight two techniques we proposed to (potentially) address the domain portability issues of device fingerprints.

- **Tweak** [8], a recently proposed technique that leverages metric learning to achieve accurate device identification in both closed-set and open-set scenarios. **Tweak** significantly reduces the training time as well the amount of data needed to do so. It achieves portability through a calibration process that is not computationally intensive and hence is employable in resource-constrained devices, and is accomplished with a limited amount of labeled training data from the target domain. Our experimental results using our IoT testbed described in Section II-A show that **Tweak** increases the resiliency of fingerprinting to changes in receiver hardware, wireless channel conditions, and the configuration of the transmitters’ LoRa protocol.

- **WideScan** [5], [6], a technique we proposed to leverage signal distortions that occur due to the hardware impairments in the spectrum band surrounding the original bandwidth of the signal to uniquely identify devices. **WideScan** considers extracting features from both the in-band and the out-of-band spectrum information to increase device separability, thus increasing the resiliency of fingerprinting to domain variations. Our experimental results show that **WideScan** improves the accuracy of fingerprinting approaches.

**V. CONCLUDING REMARKS**

This paper sheds some light on the fingerprints portability challenges, goes over some of the limitations of prior works, and covers a couple of proposed approaches. From our own research experience, it is clear that so much work still needs to be done to address the domain adaption and fingerprints portability problems in the wireless classification field.

**REFERENCES**

[1] B. Hamdaoui and A. Elmaghabb, “Deep-learning-based device fingerprinting for increased lora-iot security: Sensitivity to network deployment changes,” *IEEE network*, vol. 36, no. 3, pp. 204–210, 2022.

[2] S. Rajendran and Z. Sun, “RF impairment model-based IoT physical-layer identification for enhanced domain generalization,” *IEEE Transactions on Information Forensics and Security*, vol. 17, pp. 1285–1299, 2022.

[3] B. Hamdaoui, A. Elmaghabb, and S. Mejri, “Deep neural network feature designs for RF data-driven wireless device classification,” *IEEE Network*, vol. 35, no. 3, pp. 191–197, 2020.

[4] K. Sankhe, M. Belgiovine, F. Zhou, S. Riyaz, S. Ioannidis, and K. Chowdhury, “Oracle: Optimized radio classification through convolutional neural networks,” in *Proceedings of the IEEE Conference on Computer Communications*, 2019, pp. 370–378.

[5] A. Elmaghabb and B. Hamdaoui, “Lora device fingerprinting in the wild: Disclosing rf data-driven fingerprint sensitivity to deployment variability,” *IEEE Access*, vol. 9, pp. 142 893–142 909, 2021.

[6] A. Elmaghabb, B. Hamdaoui, and A. Natrajan, “WideScan: Exploiting out-of-band distortion for device classification using deep learning,” in *Proceedings of the IEEE Global Communications Conference*, 2020.

[7] A. Elmaghabb and B. Hamdaoui, “Comprehensive RF dataset collection and release: Deep learning-based RF fingerprinting use case,” in *Proceeding of IEEE INFOCOM 2021-IEEE Conference on Computer Communications*, 2021.

[8] J. Gaskin, B. Hamdaoui, and W.-K. Wong, “Tweak: Towards portable deep learning models for domain-agnostic LoRa device authentication,” in *2022 IEEE conference on communications and network security (CNS)*. IEEE, 2022.

[9] S. D’Oro, F. Restuccia, and T. Melodia, “Can you fix my neural network? real-time adaptive waveform synthesis for resilient wireless signal classification,” in *IEEE INFOCOM 2021-IEEE Conference on Computer Communications*. IEEE, 2021, pp. 1–10.

[10] N. Basho, B. Hamdaoui, and K. Sivanesan, “Leveraging mimo transmit diversity for channel-agnostic device identification,” in *ICC 2022-IEEE International Conference on Communications*. IEEE, 2022.
This figure "abdurrahman.jpg" is available in "jpg" format from:

http://arxiv.org/ps/2211.07687v1
This figure "bechir.png" is available in "png" format from:

http://arxiv.org/ps/2211.07687v1
This figure "bullet.png" is available in "png" format from:

http://arxiv.org/ps/2211.07687v1
This figure "logo.png" is available in "png" format from:

http://arxiv.org/ps/2211.07687v1
This figure "notaglinelogo.png" is available in "png" format from:

http://arxiv.org/ps/2211.07687v1