Experimental Evaluation of a Wireless Local Area Network (WLAN) Impersonator Detection System Using Deep Learning

Mir Lodro, Gabriele Gradoni, and Steve Greedy

Abstract – We present an experimental evaluation of a radio impersonator detection system using deep learning in an indoor environment. RF fingerprinting is used to detect a wireless local area network (WLAN) impersonator that transmits beacon frames to WLAN users with a known medium access control address. The RF impersonator detection was performed by training a neural network with real-time beacon frame data captured from the trusted routers in a multifloor building in a campus environment. In addition to the beacon frames from the trusted routers, the beacon frames were captured from an unknown router at five different locations. The neural network was trained using RF fingerprinting from the trusted routers and the unknown router. The confusion matrices about the classification of the trusted routers and the unknown router are presented. Finally, a software-defined radio (SDR) over-the-air (OTA) test was performed in the same multifloor indoor environment to detect router impersonator. Our SDR-based approach shows that the trained neural network can classify the WLAN router impersonator with high accuracy.

1. Introduction

Physical (PHY) layer security of wireless networks such as a wireless local area network (WLAN) has gained attention from researchers worldwide [1–3]. In a WLAN network, the confidentiality of the user data may be compromised if the malicious transmitter (Tx) acts as an unauthorized access point. In wireless networks, users have some degree of cooperation and information sharing among them. This makes it even more susceptible to data confidentiality if the malicious transmitter acts as a WLAN impersonator and enters the network and becomes part of the network. Therefore, deep learning–based security systems to identify potential attacks on networks and detection of RF impersonators is critically important. RF fingerprinting in conjunction with a medium access control (MAC) address can be an effective solution for the detection of impersonators. A variety of features in the PHY layer and the MAC layer can be used for RF fingerprinting. We exploit PHY layer features that are derived from the received beacon frames. PHY layer RF fingerprinting can be classified into two types: location dependent and location independent or hardware specific. Received signal strength (RSS) and more specifically channel state information (CSI) based RF fingerprinting are location-dependent methods. Two users from the transmitter at two locations can have different RSS and the CSI. The two colocated users may have the same path loss, and the RSS and the CSI values may be correlated [4, 5]. Other methods include channel frequency response PHY layer security methods [6, 7]. Location-independent RF features are hardware specific, and these are related to hardware imperfections introduced during manufacturing of the transmitters [8–10]. One approach is by identifying the transient part of the received signal that occurs in transmitters even produced by the same manufacturer. This behavior is addressed in [11]. Similarly, power amplifier imperfections can be used for identifying wireless transmitters. Work in [12] exploits carrier frequency offset and the phase offset for RF fingerprints for legitimate transmitters. Location-independent or hardware-specific RF fingerprinting can be divided into waveform based and modulation based. Waveform-based RF fingerprinting is readily available; however, it may require high sampling rates to capture the transient, especially when the transient occurs for a short duration. The modulation-based RF fingerprinting is well defined but may require knowledge of the modulation techniques. However, most of the modern communication networks, such as WLAN, use short preamble and long preamble (L-LTF) in the packet structure. Therefore, CSI-based RF fingerprinting in addition to other digital identifiers, such as a MAC address, can be more effective. An RF impersonation attack can happen on networks in which a malicious transmitter impersonates a legitimate router and connects with WLAN users. Other forms of digital identifiers, such as MAC addresses, IP addresses, and service set identifiers (SSID), are ineffective and can be easily breached. In addition to digital identifiers, RF fingerprinting is a more secure solution when a malicious router can be identified based on its RF fingerprint. Wireless transmitters such as WLAN routers create a unique RF signature or CSI at the receiver. A legitimate transmitter and an RF impersonator can be distinguished based on CSI at the receiver. The fixed WLAN router and static user create a stationary channel profile. In such a scenario, a legitimate transmitter and RF impersonator can be

Manuscript received 11 November 2022.

Mir Lodro, Gabriele Gradoni, and Steve Greedy are with Department of Electrical and Electronic Engineering University of Nottingham, Nottingham, NG7 2RD, United Kingdom; e-mail: mir.lodro2@nottingham.ac.uk.

This work was supported by the European Commission through the H2020 RISE-6G project (grant 101017011). The work of Gabriele Gradoni was supported by the Royal Society (grant INF\R2\192066).

identified using a deep learning network. The main objective of this article is to assess RF fingerprinting and train a deep learning network using data from a real-time router in the 5 GHz band. Section 2 is about the data set measurement from real WLAN routers, and it also explains architecture and the training of convolutional neural network (CNN). Section 3 focuses on a software-defined radio (SDR) experimental setup and the measurement scenarios in an indoor environment. Section 4 highlights RF impersonator detection results using real-time OTA tests. Section 5 concludes the article.

2. Data Set and CNN training

A deep learning network is trained using in-phase and quadrature (IQ) data and extracted L-LTF from a set of trusted known routers. The observer or the system collects orthogonal frequency division multiplexing (OFDM) non-high throughput (nonHT) beacon frames and uses L-LTF to identify the RF signature. The observer receives beacon frames and the decoded MAC address, and by using L-LTF, the observer measures the RF fingerprint. The CNN architecture consists of two convolutional and three fully connected layers. The first layer learns features independently from I and Q of the beacon frames. The first two layers use filter sizes of 1×7 and 2×7, respectively. The second layer learns the features by combining I and Q together. Between each convolutional layer leaky rectified linear unit (ReLU) activation function is applied. We also use maximum pooling to reduce spatial size and reduce amount of computations in the network. The maximum pooling layer is helpful, as it detects the features of down-sampled data more effectively. The last three fully connected layers behave as classifiers using extracted features from the previous two layers. The final layer uses softmax activation function to generate the probabilities of the input being in a particular class. An adaptive moment (ADAM) estimation optimizer with a minibatch size of 128 is used. The execution environment is set to graphics processing unit (GPU) so that it uses a NVIDIA P400 GPU. CNN is trained using beacon frames transmitted from a set of trusted known and unknown routers. The data set of 3600 nonHT beacon frames from each router is collected from six trusted WLAN routers in a multistory building environment on the campus. The MAC addresses of the known routers are used as labels. Also, the data set of 500 nonHT beacon frames is collected from an unknown router at five different locations. The data set was collected at channel 36 in the 5 GHz band at a RF frequency of 5.180 GHz, taking around 17 h to capture data from the six trusted routers. The MATLAB program R2022a was forced to wait for the router if it had jumped to another frequency, and it captured Rx data when it only returned to channel 36. The data set includes SSIDs of the routers, MAC addresses, and the L-LTF information. The MAC address of the unknown router is labeled as *unknown*. CNN is trained to decide

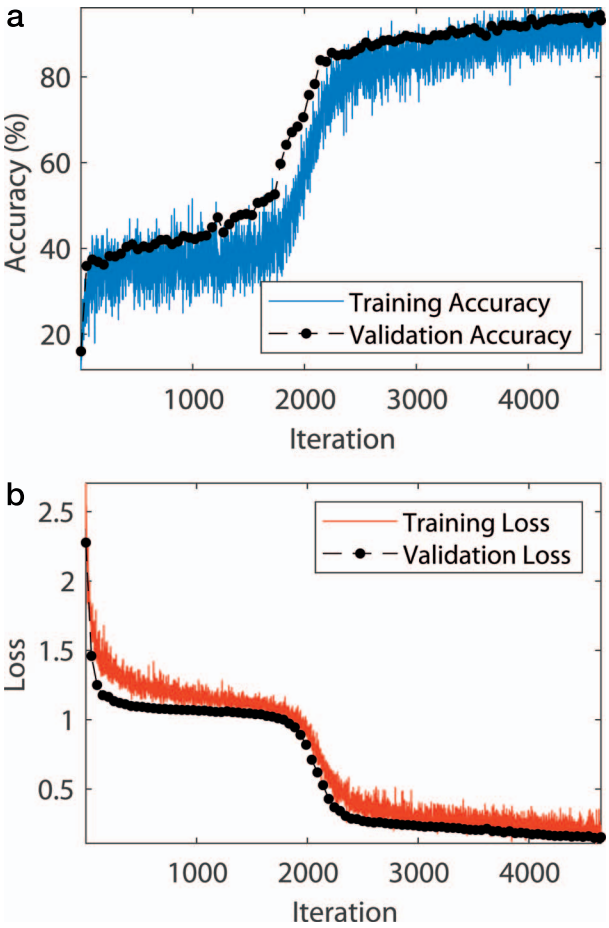

Figure 1. CNN training.

based on the MAC addresses and the RF fingerprinting to detect impersonators. The CNN was trained by combining data sets from six trusted routers and the data set from an unknown router measured at five different locations in an indoor office and corridor within the same building. The data set is split into training and the testing at 90% and 10%, respectively. Figure 1 shows the training progress of the CNN and training and validation of the accuracy and the loss. The CNN converges in about 22 epochs to 100% accuracy within the elapsed time of 6 min 4 s. The training was performed using a single small form factor NVIDIA P400 GPU. The NVIDIA P400 GPU has a memory of 2 Graphics double data rate 5 (GDDR5) and a memory bandwidth of 2 GB/s.

3. Experimental Setup

In our measurement setup, we used PlutoSDRs for capturing the real-time beacon frames from WLAN routers. PlutoSDR is a cost-effective SDR based on AD9363 RFIC and can perform transmit and receive operations with a minimum sampling rate of 65.1 ksps. We have used PlutoSDR in our previous works for testing the wireless communication systems in metal

Figure 2. Experimental setup in indoor office environment and the corridor environment.

enclosures [13] and for the reconfigurable intelligent surface–assisted modulation classification in the indoor environment [14]. PlutoSDRs were used to measure the real-time data from the set of trusted WLAN routers on the same floor and across the floor. A database

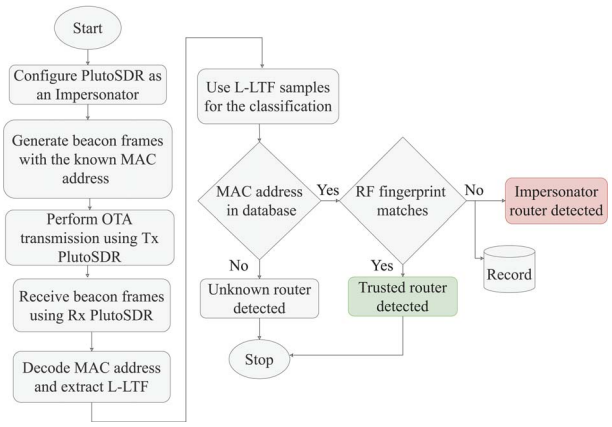

Figure 3. Measurement and router impersonator detection process.

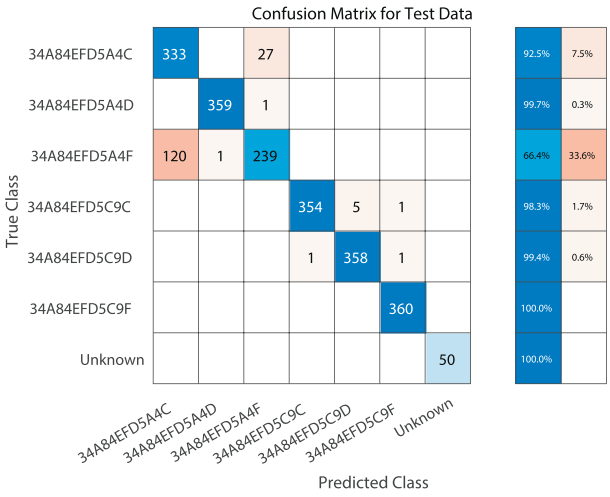

Figure 4. Confusion matrix of classification of known routers and unknown router.

consisting of real-time data and the MAC addresses of the trusted routers was created. WLAN routers that broadcast the OFDM nonHT beacon frames to show their presence in the 5 GHz band use IEEE 802.11 a/g/n/ac protocols [15]. The beacon frames were measured in an indoor office environment. The Rx PlutoSDR also known as the observer was stationary for the whole measurement duration when the known router beacon frames were captured. The data were also captured by an unknown router at five different locations. Two were in the corridor and three in office environments. For this, one PlutoSDR was configured as a transmitter and another PlutoSDR as a receiver that would collect data at five different locations. The unknown router data were collected at two different locations in the corridor and at three locations in the office environment (see Figures 2b and 2c). The data were collected from six trusted routers with known MAC addresses. The nonHT beacon signals were collected from these routers, and L-LTF was used to collect unique RF fingerprints of these routers. The routers and the observer were fixed, so the RF signature did not vary in time. Unknown router data were collected at five random locations in line of sight (LOS) and non-LOS. Therefore, it is a set of random RF fingerprints that is different from the known routers RF fingerprints. The observer (Rx PlutoSDR) receives beacon frames, decodes MAC addresses, extracts L-LTF, and uses this to classify the RF fingerprints of the source of beacon frames. The observer then makes the following decisions: 1) if the MAC address is in the database and the RF fingerprint matches, the system decides known router; 2) if the MAC address is not in the database, and the RF fingerprint does not match any of the known routers in the database, the observer declares the source an as unrecognized router; 3) if the address is in the database, but the RF fingerprint does not match, the system declares it as RF impersonator.

4. Results

We performed a real-time OTA test using Tx PlutoSDR and Rx PlutoSDR for the detection of an RF impersonator. We transmitted beacon frames using Tx PlutoSDR using MAC addresses of the trusted routers and captured beacon frames using another PlutoSDR acting as the observer. After the RF data capture, we performed CSI extraction, and the CSI was sent to the classifier for RF impersonator detection. We found that the classification accuracy of the classifier at higher SNR is greater than 99%. The classification accuracy was compromised when the OTA was performed at low SNRs.

5. Conclusion

We have experimentally demonstrated a WLAN impersonator detector using deep learning in an indoor environment in a multifloor building. The beacon frames were captured from six trusted WLAN routers operating at channel 36 in the 5 GHz industrial, scientific, and medical band. We found that the system successfully detects WLAN impersonators in the indoor environments. The classification accuracy for the unknown router is 100%, and the classification accuracy of WLAN routers is higher than 82.8%. The worst classification accuracy (67.5%) is shown only for one WLAN router. We verified the classification performance by doing OTA transmission with Tx and Rx PlutoSDRs and performed the RF fingerprinting measurement and WLAN impersonator classification.

6. References

1. J. Lv, D. Man, W. Yang, X. Du, and M. Yu, "Robust WLAN-Based Indoor Intrusion Detection Using PHY Layer Information," *IEEE Access*, **6**, December 2017, pp. 30117-30127.
2. S. Chen, S. Wang, X. Xu, L. Jiao, and H. Wen, "Physical Layer Security Authentication Based Wireless Industrial Communication System for Spoofing Detection," IEEE Conference on Computer Communications Workshops, New York, NY, USA, May 2–5, 2022, pp. 1-2.
3. D. S. Karas, A.-A. A. Boulogeorgos, G. K. Karagiannidis, and A. Nallanathan, "Physical Layer Security in the Presence of Interference," *IEEE Wireless Communications Letters*, **6**, 6, December 2017, pp. 802-805.
4. J. K. Tugnait and H. Kim, "A Channel-Based Hypothesis Testing Approach to Enhance User Authentication in Wireless Networks," 2010 Second International Conference on Communication Systems and Networks, Bangalore, India, January 5–9, 2010, pp. 1-9.
5. L. Y. Paul and B. M. Sadler, "MIMO Authentication via Deliberate Fingerprinting at the Physical Layer," *IEEE Transactions on Information Forensics and Security*, **6**, 3, September 2011, pp. 606-615.
6. L. Xiao, L. Greenstein, N. Mandayam, and W. Trappe, "Fingerprints in the Ether: Using the Physical Layer for Wireless Authentication," 2007 IEEE International Conference on Communications, Glasgow, United Kingdom, June 24–28, 2007, pp. 4646-4651.
7. L. Xiao, L. Greenstein, N. Mandayam, and W. Trappe, "A Physical-Layer Technique to Enhance Authentication for Mobile Terminals," 2008 IEEE International Conference on Communications, Beijing, China, May 19–23, 2008, pp. 1520-1524.
8. V. Brik, S. Banerjee, M. Gruteser, and S. Oh, "Wireless Device Identification With Radiometric Signatures," 14th ACM International Conference on Mobile Computing and Networking, San Francisco, CA, USA, September 14–19, 2008, pp. 116-127.
9. A. C. Polak, S. Dolatshahi, and D. L. Goeckel, "Identifying Wireless Users via Transmitter Imperfections," *IEEE Journal on Selected Areas in Communications*, **29**, 7, August 2011, pp. 1469-1479.
10. S. Dolatshahi, A. Polak, and D. L. Goeckel, "Identification of Wireless Users via Power Amplifier Imperfections," 2010 Conference Record of the Forty Fourth Asilomar Conference on Signals, Systems and Computers, Pacific Grove, CA, USA, November 7–10, 2010, pp. 1553-1557.
11. J. Hall, M. Barbeau, and E. Kranakis, "Enhancing Intrusion Detection in Wireless Networks Using Radio Frequency Fingerprinting," Communications, Internet, and Information Technology 2004, St. Thomas, US Virgin Islands, USA, November 22–24, 2004, pp. 46-56.
12. N. T. Nguyen, G. Zheng, Z. Han, and R. Zheng, "Device Fingerprinting to Enhance Wireless Security Using Nonparametric Bayesian Method," 2011 Proceedings IEEE INFOCOM, Shanghai, China, April 10–15, 2011, pp. 1404-1412.
13. M. Lodro, G. Gradoni, J.-B. Gros, S. Greedy, and G. Lerosey, "Reconfigurable Intelligent Surface-Assisted Bluetooth Low Energy Link in Metal Enclosure," *Frontiers in Communications and Networks*, **2**, October 2021, p. 733637.
14. M. Lodro, H. Taghvaee, J.-B. Gros, S. Greedy, G. Lerosey, et al., "Reconfigurable Intelligent Surface-Assisted Classification of Modulations Using Deep Learning," 2022 3rd URSI Atlantic and Asia Pacific Radio Science Meeting, Gran Canaria, Spain, May 30–June 4, 2022, pp. 1-4.
15. Hanif Rahbari, Marwan Krunz, "Exploiting Frame Preamble Waveforms to Support New Physical-Layer Functions in OFDM-based 802.11 Systems" IEEE Transactions on Wireless Communications, Vol. 16, N0.6, June 2017.