PROPOSAL

ML LTE Fingerprinter

Brad Williams, Chris Limson

GMU CYSE 640 Wireless Network Security - Moinul Hossain, PhD

October 2, 2024

Problem Statement

When a device attempts to connect to a base station, it needs to be authenticated and processed unless it has been authenticated before. With LTE, it increases the complexity for authentication due to the LTE standards that are necessary to implement. With technology becoming more advanced, there are an increasing number of ways to spoof a device's MAC Address, IP address or other identifying information.

With Radio Frequency (RF) Fingerprinting, where due to imperfections in the hardware of a device when it is created there is a specific signal impairments emitted by the device, we can decipher the specific device attempting to connect to the base station. Authenticating with a device's RF fingerprint could make it quicker and more secure. This is because there is currently no known way to spoof a RF fingerprint, thus allowing each device to be cataloged and identified off this attribute.

Related Work

With the introduction of radio fingerprinting in [2], we are able to use machine learning models and software defined radio frequencies to give each device, although similar in nature, a specific RF fingerprint.

It is mentioned in [2] that "No higher level decoding, feature engineering, or protocol knowledge is needed, further mitigating challenges of ID spoofing and coexistence of multiple protocols in a shared spectrum" which would allow the protocols to be further strengthened in other directions. Instead of having overlapping protocols for authentication, there is the possibility of a single authentication method based off of the RF fingerprint, further reducing the overhead cost of authentication in each protocol. One significant finding in this paper is that they have tested it at specific directions, notably ranging from 2 feet away to 50 feet away. It is also mentioned that the identification remains accurate until it drops off after 34 feet.

In [1], the authors use RF fingerprinting on a variety of devices from different manufacturers and find that detecting and authenticating a user based on their RF fingerprint is a valid control. With the authors implementing their scenario in 4G-LTE, it gives precedence for our development also centered in LTE. There is a brief mention in the future work of the paper about how degradation can occur over the course of the hardware's lifetime. This could potentially cause it to no longer have the same fingerprint, or possibly emulate a similar fingerprint of another device. Additionally, there is no note of intentional disruption of the hardware, potentially changing the fingerprint of the device.

Proposed New Solution

Our proposed solution is further development into data for convolutional neural networks. We hope to address the proposal in section 7, Research Challenges, of [2] where the authors suggest building a standard dataset benchmarking training time and classification accuracy. We plan on implementing that with cloud storage that is continually updated and potentially analyzed for information. Like the original researchers, we will continue to employ TensorFlow, the open-source software library for computing the convolutional neural network [3].

Additionally, we will look into addressing the challenge of hardware degradation. Since this will be a simulated model, we will be updating properties for each device and seeing how the deep learning model reacts to the new information. In the potentially additional interest of flagging malicious devices by only their fingerprints, as all other higher-level (logic, MAC addresses) can be altered. Proposing a solution where the base station identifies a connecting device based solely off their RF fingerprint. This is due to the increasing commonality of logic, IP addresses and MAC addresses being able to be altered. We hope to address the impact of distance affecting the success of identification by increasing the sensitivity of the detection hardware. While this is not always viable, since we are simulating it, we can change the sensitivity and make the receiver know what it is looking for. We could also potentially increase the power of the 'smartphones' that are simulated, to see how far the signal can be pushed in an ideal scenario.

Evaluation Plan

Our evaluation plan is to apply a large number of simulated devices to the entire system's algorithm, showing how the built dataset improves identification, as more devices contribute. This includes having more than one device try authenticating against the base station based off of their RF fingerprint.

A couple novel ideas for authentication and experimentations include; Having everything cataloged by the base station for a device be incorrect except the fingerprint, have everything cataloged be correct and the fingerprint be incorrect, and have everything be correct / incorrect for authentication. It is undecided how feasible some of these scenarios are. The effectiveness will be considered if the device is properly authenticated or not, with there being 2 success scenarios and 2 intended failure scenarios.

Timeline and Milestones

October 16 - Replicate up to the described proof-of-concept simulations with MATLAB Communications, WLAN and LTE Toolboxes

October 30 - Build cloud database, use GNU Radio's Signal Metadata Format (SigMF) to store signal data, TensorFlow CNN

November 13 - Test, Evaluate accuracy / speed / potential defensive process

November 27 - Writeup

December 4 - Present

Works Cited

[1] Abbas, S., Nasir, Q., Nouichi, D. *et al.* Improving security of the Internet of Things via RF fingerprinting based device identification system. *Neural Comput and Applications* (2021) 33:14753–14769 (2021). https://doi.org/10.1007/s00521-021-06115-2

[2] S. Riyaz, K. Sankhe, S. Ioannidis and K. Chowdhury, "Deep Learning Convolutional Neural Networks for Radio Identification," in *IEEE Communications Magazine*, vol. 56, no. 9, pp. 146-152, Sept. 2018, doi: 10.1109/MCOM.2018.1800153.

[3] F. Ertam and G. Aydın, "Data classification with deep learning using Tensorflow," 2017 International Conference on Computer Science and Engineering (UBMK), Antalya, Turkey, 2017, pp. 755-758, doi: 10.1109/UBMK.2017.8093521.