HAUTE ÉCOLE
D'INGÉNIERIE ET DE GESTION
DU CANTON DE VAUD
www.heig-vd.ch

Department: Information and communication technologies
Faculty: Information technology and communication systems
Orientation: Software engineering

Bachelor Thesis

RF fingerprinting on NFC devices

Not confidential

Student: Luc Wachter

Project proposed by: Joël Conus
Kudelski Group SA
22-24, Route de Genève
1033 Cheseaux-sur-Lausanne

Teacher in charge: Prof. Alberto Dassatti

Academic year: 2019-2020

Yverdon-les-Bains, 29th May 2020

1 Scope statement

Scope statement for RF fingerprinting on NFC devices (Bachelor Thesis)

Project purpose: Develop a tool to identify NFC devices by analysing the RF spectrum of their transmitted signals

Duration: 450 hours (ends 31st July 2020)

Detailed description

RF fingerprinting is a technique that allows the identification of radio transmitters (such as IoT devices) by analysing the spectrum of their transmissions. This analysis can typically be performed using machine learning algorithms.

NFC technology is often used in access control and payment applications but many implementations are vulnerable to relay attacks with research and tools that facilitate such attacks being publicly available.

The goal of this project is to determine if RF fingerprinting could be used as an authentication technique against relay attacks.

The main steps of this project are the following:

- Build a simple lab setup with Software-Defined Radio (SDR) equipment to acquire signals between an NFC device and its reader
- Acquire RF spectrum data of various NFC devices
- Analyse the signals
- Classify the signals of the devices by using supervised machine learning classification techniques in order to differentiate trusted devices from attacker / relay devices
- Determine if this identification technique could be used as an authentication feature against relay attacks

As the receiving equipment (SDR) has an influence on the recorded signals, for this project we consider a single receiver to record the RF samples. Similarly, the lab setup should be built to provide an ideal low-noise & low-interference environment to simplify the analysis phase.

The expected deliverables are the following:

- A tool able to identify NFC devices by analysing the RF spectrum of their signals, at least in an ideal environment and with a small number of devices
- A detailed account of the steps taken and the setup used (as part of the report)
- An analysis of the results (as part of the report)

Collaboration with other researchers in this field is wished (EPFL, ElectroSense).

Table of contents

2 Introduction

2.1 Project description

RF fingerprinting is a technique that allows the identification of radio transmitters by extracting small imperfections in their spectrum. These imperfections are caused by tiny manufacturing differences in the devices' analog components. Using Software-Defined Radio (SDR) equipment, we can analyse this spectrum in order to extract the aforementioned differences and identify a device.

Such techniques can be used on any type of radio transmission: Bluetooth, BLE, WiFi, LTE, etc. This project aims to use RF fingerprinting on NFC devices. Indeed, NFC is often used in access control and payment applications but many implementations are vulnerable to relay attacks. Spoofing the imperfections in an emitter's radio spectrum is close to impossible at the present time, since it is essentially a hardware signature. This is why a technique like the one described here would be a valuable additional security layer.

The goal of this project is to determine if RF fingerprinting of NFC devices could be used as an authentication technique, in order to prevent relay attacks.

2.2 Context

This project is conducted in the context of my bachelor thesis at HEIG-VD.

- Department: Information and communication technologies
- Faculty: Information technology and communication systems
- Orientation: Software engineering

2.3 Document description

In this document,

3 State of the art

3.1 Taxonomy

It is certainly useful to start with a review of the different ways to categorize the features and algorithms used by researchers in the field. A good categorization of parameters allows us to define our needs precisely. It makes it easier to select the important things to consider.

3.1.1 Taxonomy for features

The features we select must allow us to identify a precise device among potentially very similar devices. We need what Delgado et al. [1] describe as a Physical Unclonable Function (PUF). PUFs are physical distortions that are unique to a specific system. They are another way of talking about fingerprints.

Xu et al. [2] propose three ways to categorize radio signal features:

- based on the specificity of the feature (from vendor specific to device specific),
- based on the layers (PHY, MAC, Network and higher),
- and based on the acquisition method (passive or active).

Whether we end up with a system that is able to identify many devices uniquely, or one that only tries to separate a specific device from the others, we will need device specific accuracy. We don't want to make relay attacks impossible only if the attacker doesn't use a device from the same vendor as the victim's device.

Moreover, features from the MAC and higher layers typically require in depth knowledge of the protocols in play. Not only that, but they also tend to be less specific than we would like (either vendor specific or depending on the type of device). This indicates we should probably focus on the physical (PHY) layer features, which rely on imperfections in the manufacturing process of the devices.

3.1.2 Taxonomy for fingerprinting algorithms

Riyaz et al. [3] provide a visual categorization of fingerprinting approaches. To summarize it, they first separate supervised from unsupervised learning. They then further categorize supervised approaches between similarity-based and classification techniques. **MAKE A SCHEMA?**

Include taxonomy from Xu et al. [2]

Because of the nature of our problem, a supervised classification approach seems most appropriate. Indeed, we can assume that we have access to the legitimate device, and to a population of illegitimate devices. This means we can label the gathered data and use it to train, for example, a Convolutional Neural Network (CNN). **WEAK**

Going the unsupervised route would imply a radically different approach. This is because unsupervised systems cannot by themselves discriminate a legitimate device from an illegitimate one. They don't have that information, since they work with unlabeled data. With this said, a system conceived like that would still be able to detect attempts of impersonation by keeping a dictionary of fingerprints and linked identifiers.

3.2 Acquisition

Not certain this is useful.

3.3 Features selection

In section 3.1.1 we discussed the different types of features that exist in radio signal data. We concluded that the features we are most interested in are from the physical layer.

transient phase

3.4 Machine learning

3.4.1 Comparing approaches

Several articles have compared the performance of different machine learning approaches.

...

3.4.2 Neural network architecture

...

What about [pre 4, page 2]?

I love[1].

- **Waveform domain techniques** [11], [14], [22], [24], [25] consider time and frequency representation as the basic blocks while **modulation domain techniques** [6] represent signals in terms of I/Q samples. - Waveform domain techniques are more flexible but more complex. Modulation domain techniques are better structured and well-behaved but require knowledge of the respective modulation scheme.

[1] pre 5, post.

4 Dataset creation

4.1 Initial setup

4.2 Upgraded setup

4.3 Inventory of devices

The content of the tags is harmonized to ensure the algorithm won't use the content as a feature to identify devices.

MEGA TABLE

4.4 Dataset description

5 Model conception

5.1 Model architecture

Decision on the metaparameters...

5.2 ...

6 Conclusion

Bibliography

1. DELGADO, Oscar; KECHTBAN, Louai; LUGAN, Sebastien, et al. Passive and active wireless device secure identification. *IEEE Access* [online]. 2020, pp. 1–1 [visited on 2020-05-08]. ISSN 2169-3536. Available from: https://ieeexplore.ieee.org/document/9082628/.
2. XU, Qiang; ZHENG, Rong; SAAD, Walid, et al. Device Fingerprinting in Wireless Networks: Challenges and Opportunities. *arXiv:1501.01367 [cs]* [online]. 2015 [visited on 2020-04-24]. Available from arXiv: 1501.01367.
3. RIYAZ, Shamnaz; SANKHE, Kunal; IOANNIDIS, Stratis, et al. Deep Learning Convolutional Neural Networks for Radio Identification. *IEEE Communications Magazine* [online]. 2018, vol. 56, no. 9, pp. 146–152 [visited on 2020-04-24]. ISSN 0163-6804 1558-1896. Available from: https://ieeexplore.ieee.org/document/8466371.
4. YOUSSEF, Khalid; BOUCHARD, Louis; HAIGH, Karen, et al. Machine Learning Approach to RF Transmitter Identification. 2017. Available also from: https://www.researchgate.net/publication/320890695_Machine_Learning_Approach_to_RF_Transmitter_Identification.
5. STANKOWICZ, James; ROBINSON, Josh; CARMACK, Joseph M., et al. Complex Neural Networks for Radio Frequency Fingerprinting. In: *2019 IEEE Western New York Image and Signal Processing Workshop (WNYISPW)* [online]. Rochester, NY, USA: IEEE, 2019, pp. 1–5 [visited on 2020-02-24]. ISBN 978-1-72814-352-1. Available from: https://ieeexplore.ieee.org/document/8923089.

List of Figures

List of Tables