2022 19th Annual IEEE International Conference on Sensing, Communication, and Networking (SECON) | 978-1-6654-8643-9/22/$31.00 ©2022 IEEE | DOI: 10.1109/SECON55815.2022.9918173

TeRFF: Temperature-aware Radio Frequency Fingerprinting for Smartphones

Xiaolin Gu*, Wenjia Wu†, Naixuan Guo†, Wei He*, Aibo Song†, Ming Yang†, Zhen Ling†, Junzhou Luo†
*School of Cyber Science and Engineering, Southeast University, Nanjing, China
†School of Computer Science and Engineering, Southeast University, Nanjing, China
Email: {xiaolin_gu, wjwu, guonaixuan, hw2021, absong, yangming2002, zhenling, jluo}@seu.edu.cn

***Abstract*—In recent years, radio frequency (RF) fingerprinting has attracted more and more attention. Many different types of RF fingerprints have been proposed, such as carrier frequency offset (CFO), sampling frequency offset and error vector magnitude. Among them, the CFO fingerprint is recognized as a promising RF fingerprint. However, for commonly used smartphones, we find that its CFO fingerprint is unstable, because the temperature of crystal oscillator varies greatly and large fluctuations of temperature significantly affect its CFO fingerprint. Therefore, the solutions of CFO-based fingerprinting will no longer be effective for smartphones if the temperature of crystal oscillator is not involved. To this end, we propose a more reliable and applicable CFO-based fingerprinting approach called temperature-aware radio frequency fingerprinting (TeRFF). First, we construct a dataset by extracting crystal oscillator's temperature and the corresponding CFO value on multiple smartphones over a period. In the dataset, the extracted temperature values constitute a set of temperature values, and each registered temperature value corresponds to a group of CFO samples. On this basis, we train multiple Naive Bayes models, each tagged with a registered temperature value. Moreover, since there are many temperature values which are not in the temperature set, we design a CFO estimation method to estimate the CFO fingerprint at the unregistered temperature. Finally, the experimental results demonstrate that our proposed solution TeRFF makes the CFO fingerprinting still effective for smartphone identification, and its performance is better than other existing RF fingerprinting schemes.**

***Index Terms*—Smartphones, Radio frequency fingerprinting, Carrier frequency offset, Crystal oscillator's temperature**

I. Introduction

As the Wi-Fi technique develops gradually, the number of Wi-Fi devices is increasing in recent years, and the global Wi-Fi market size is projected to grow from USD 9.4 billion in 2020 to USD 25.2 billion by 2026 [1]. In the case of large-scale Wi-Fi devices, the authentication becomes more and more significant for the security of Wi-Fi devices. Conventional Wi-Fi device authentication is based on IP address, MAC address and pre-shared secret information [2]. But these information are easily forged by malicious people for attacking the internal network. Motivated by the above, noncryptographic solutions based on device identification are proposed recently. For device identification, Radio Frequency (RF) fingerprinting is a reliable and secure approach because RF fingerprints have device-related characteristics in Wi-Fi devices. For attackers, it is very difficult to tamper with hardware imperfections in commercial Wi-Fi devices. Therefore RF fingerprinting has attracted the attention of many researchers.

RF fingerprinting can be equivalent to a multi-classification problem. Currently, RF fingerprinting for Wi-Fi devices can be divided into deep learning-based and handcrafted feature-based methods separately. Deep learning-based method can automatically extract effective and underlying features from raw or simple preprocessed I/Q signal without relying on the knowledge of wireless communication. In addition, the modulation mode of transmitted signal can not be known in advance. CNN [3]–[5] and MLP [6] have been used in Wi-Fi device identification. However deep learning model may be degraded largely by wireless channel [4]. Therefore the deep learning-based method is not applicable in the practical scenarios. On the other hand, the identification methods based on handcrafted features depend on expert knowledge and these handcrafted features can indicate specific imperfections in wireless devices. These handcrafted features can mainly be divided into transient [7], [8] and modulation [2], [9], [10] features. Transient features can be extracted in the transient stage of the transmitter, which presents a satisfactory performance. However, the acquisition of these features needs expensive RF receivers with high sampling rates in the order of gigabytes. This makes the method impossible to deploy on a large scale. By contrast, modulation features like carrier frequency offset (CFO), sampling frequency offset and error vector magnitude originate from modulation errors and these can be extracted by software-defined receiver like USRP N210 in the medium price range. Among the modulation features, CFO fingerprint is considered to be very effective in [9], [11].

Hence, CFO fingerprint is an effective feature in RF fingerprinting. However, it is found that most existing works assume that the temperature of crystal oscillator in the Wi-Fi device is relatively stable. In reality, the smartphone is also a type of Wi-Fi device, but crystal oscillator's temperature in the smartphone has a large variance because it is easily affected by its power consumption. Besides, temperature is one of the most important factors which affect the crystal oscillator in a RF front-end [12]. The following mistakes possibly occur, which are shown in Fig. 1: the same smartphone has different fingerprints, and different smartphones have similar fingerprints. We use the symbol f to represent a function from the crystal oscillator set $\mathbf{O}$ to CFO set $\mathbf{C}$. However, inverse function f^{-1} doesn't exist. To be specific, CFO y has two corresponding values in the crystal oscillator set $\mathbf{O}$: crystal oscillator O_1 and O_2. As a result, it will lead to the mistake

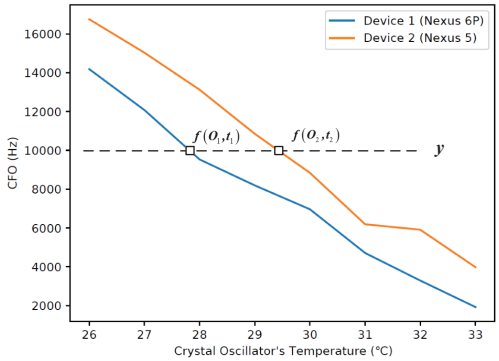


Fig. 1. The CFO of the same device decreases monotonically with temperature, and different devices may have the same CFO at different temperatures

of smartphone identification. Through further analysis, we find that temperature of the crystal oscillator in smartphone's Wi-Fi chipset is necessary to be taken into consideration. It is very hard for us to extract temperature-independent features from the smartphone because modeling high-integrated smartphones accurately is very difficult. Hence, we still leverage the CFO, which is recognized as a promising RF fingerprint, and propose an enhanced CFO-based fingerprinting approach TeRFF. First, we construct a dataset which contains extracted crystal oscillator's temperature and the corresponding CFO values over a period. These registered temperature values constitute a set of temperature values. Then, several Naive Bayes models are trained according to the stability of CFO at each temperature and each model is tagged with the registered temperature. In addition, temperature values in the dataset can not cover all temperatures that the smartphone can experience because the time spent on training sample collection usually is limited. Therefore we design a CFO estimation method for smartphone identification when the smartphone is at an unregistered temperature. Our contributions are listed as follows:

- We find that the CFO of Wi-Fi signals is unstable for a smartphone, and analyze the reasons for this unstability. Through further experiments, we can obtain that the temperature of crystal oscillator on smartphones varies greatly, which will significantly affect the stability of CFO.
- We propose an enhanced CFO-based fingerprinting approach called TeRFF. Firstly, we build the dataset through extracting the CFO of Wi-Fi signals and the corresponding crystal oscillator's temperatures over multiple smartphones, so that there are a large number of CFO samples for each registered temperature. Then, we train a classifier for smartphone identification, which includes a series of Naive Bayes models and each of them is tagged with a registered temperature. Moreover, we design a CFO estimation method to estimate the CFO fingerprint at the unregistered temperature.
- We conduct smartphone identification in different temperatures among 15 smartphones, which include seven kinds of the device model. It demonstrates that the accuracy of our proposed method can reach nearly 80%. In addition, the performance of our proposed method is stable at different times and locations. Finally the average accuracy of our method is 8.9% and 48.1% higher than that of other two state-of-art methods respectively in the smartphone identification.

II. BACKGROUND

In this section, we introduce the principle of CFO generation and then present the CFO extraction for the smartphone .

A. Principle of CFO Generation

CFO reflects the slight deviation between the carrier frequency in a transmitter and a receiver separately. As shown in Fig. 2, at the transmitter, a digital baseband signal is converted to an analog baseband signal $x(t)$ by a digital to analog converter (DAC) module, and then the $x(t)$ further is converted to a carrier signal which has a RF frequency f_{t_x}:

$$s_{t_x}(t) = x(t)e^{j2\pi f_{t_x} t} \tag{1}$$

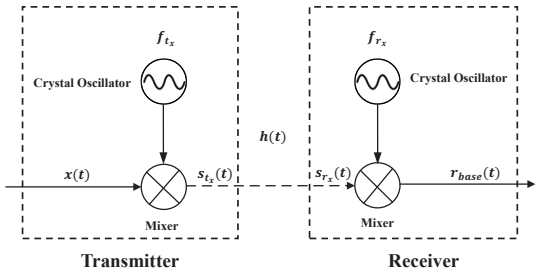


Fig. 2. The baseband signal is converted to the RF signal through an up-converter. After channel fading, the radio frequency signal is converted to the baseband signal through a down-converter. CFO occurs when the frequency of the crystal oscillator in the receiver does not synchronize with that of the crystal oscillator in the transmitter.

The RF signal $s_{t_x}(t)$ multiplied by channel effect $h(t)$ and arriving at the receiver is $s_{r_x}(t)$. To simplify the description, we take the channel with single propagation path as an example and ignore the weak noise $n(t)$, i.e.,

$$s_{r_x}(t) = h(t)s_{t_x}(t) + n(t) \approx h(t)s_{t_x}(t) \tag{2}$$

When the signal $s_{r_x}(t)$ is received by the antennas in the receiver and under downconversion, the analog baseband signal $r_{base}(t)$ is sampled by a analog to digital converter (DAC) module as a digital baseband signal for further processing.

$$\begin{aligned} r_{base}(t) = s_{r_x}(t)e^{-j2\pi f_{r_x} t} &= s_{t_x}(t) \times h(t) \times e^{-j2\pi f_{r_x} t} \\ &= x(t) \times h(t) \times e^{-j2\pi (f_{t_x} - f_{r_x})t} \\ &= x(t) \times h(t) \times e^{j2\pi \Delta f t} \end{aligned} \tag{3}$$

Obviously, Δf indicates CFO in Equation 3. CFO indicates the difference between carrier frequency generated from the transmitter and receiver separately. Furthermore, the carrier

frequency is strongly related to the crystal oscillator, which is an essential component of the smartphone. It generates an electrical signal with a constant frequency. According to different usages such as time tracking, up-conversion, down-conversion and so on, the frequency multiplier leverages the oscillator as a basic frequency source to generate various frequencies at different levels. However, many external factors will affect the stability of the crystal oscillator. The temperature is a significant factor which can affect the internal frequency of the crystal oscillator considerably [12]. In detail, the internal crystal in the crystal oscillator has a frequency-temperature characteristic represented by the following equation:

$$\frac{\Delta f}{f_0} = A_1 (T - T_0) + A_2 (T - T_0)^2 + A_3 (T - T_0)^3 \quad (4)$$

where f_0 is a reference frequency of the crystal when it is at temperature T_0. $\frac{\Delta f}{f_0}$ refers to relative frequency changes with the variation of temperature. Besides, the coefficients A_1, A_2 and A_3 are constants which have a very close relationship with physical properties of the crystal.

B. CFO Extraction in Smartphones

The common smartphones support 802.11a/g/n/ac protocols. These protocols leverage OFDM technology to enhance data rate [13]. A Wi-Fi OFDM frame has a common structure among several 802.11 protocols above. As shown in Fig. 3, the OFDM frame includes four fields which are Short Training Field (STF), Long Training Field (LTF), Signal Field (SIG) and Data Field respectively.

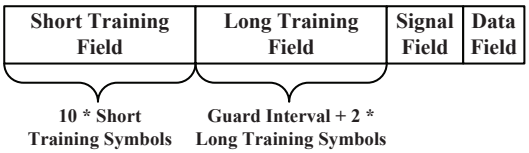


Fig. 3. Wi-Fi OFDM Frame Structure

1) Packet Detection: When a smartphone receives RF signal, it converts high radio frequency signal to baseband digital signal by a RF front end. Furthermore, it is necessary to determine whether OFDM frames occur in the received signal samples. To be specific, STF is represented by 12 subcarriers in the frequency domain. From the perspective of time domain, STF is a periodical signal which contains 10 repeated short training symbols. As a result, signal autocorrelation can indicate whether it is periodical. If the autocorrelation is periodical, an OFDM frame is detected. By the way, the coarse CFO α_i can be estimated quickly by the short training symbols i and $i+1$ for the next processing step. It can be calculated as the following equation:

$$\alpha_i = \frac{1}{16} arg\Big(\sum_{n=(i-1)*16}^{N_s-1+(i-1)*16} x[n] \times x^*[n+16] \Big) \quad (5)$$

where arg indicates the phase of signal, N_s is the length of window in STF and $x^*[n]$ represents complex conjugate of $x[n]$.

2) Symbol Alignment: After the packet detection and coarse CFO correction, symbol alignment needs to be carried out using LTF. LTF is represented by 53 subcarriers in frequency domain. From the view of time domain, LTF consists of 2 repeated training long symbols and a guard interval. At a sample rate of 20 MHz, LTF can be sampled as 160 discrete sampling points. The peak of cross-correlation between unknown signal and LTF reference signal corresponds to the accurate start point of data field in an OFDM frame.

3) Average CFO Estimation: Although coarse CFO is estimated in the period of packet detection, it has a large random error. To correct the random error, all short training symbols can be used to calculate a more precise value $\bar{\alpha}$:

$$\bar{\alpha} = \frac{1}{9} \sum_{n=0}^{8} \Big(\frac{1}{16} arg\Big(\sum_{k=0}^{N_s-1} x[16 \times n + k] \times x^*[16 \times (n+1) + k] \Big) \Big) \quad (6)$$

In order to simplify the description, the CFO in the following sections indicates the average CFO $\bar{\alpha}$ calculated by Equation 6.

III. Smartphone Identification

We first introduce the scenario of RF fingerprinting for device authentication, then the general process of model training and smartphone identification is presented. In addition, we focus on the CFO-based classification at registered temperatures and CFO estimation at unregistered temperatures.

A. System Model

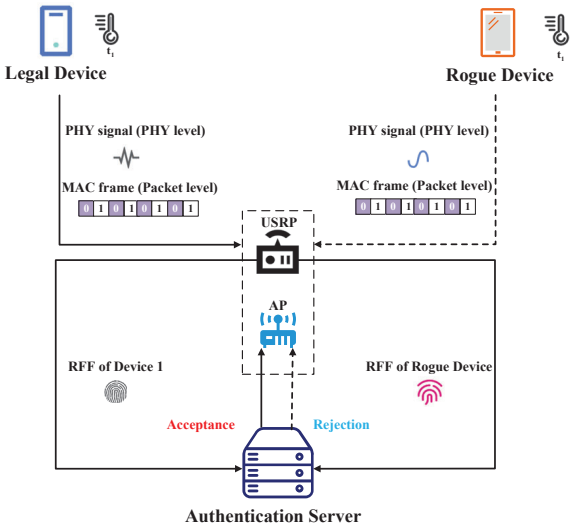


Fig. 4. Scenario of RF fingerprinting for smartphone in device authentication

RF fingerprinting for smartphone is used in device authentication. In the device authentication, we assume that attackers have an ability to masquerade as a legal user at the packet level but they do not forge RF signal at the PHY level. As illustrated in Fig. 4, the malicious attackers attempt to forge strong identifiers like MAC address and encryption/decryption

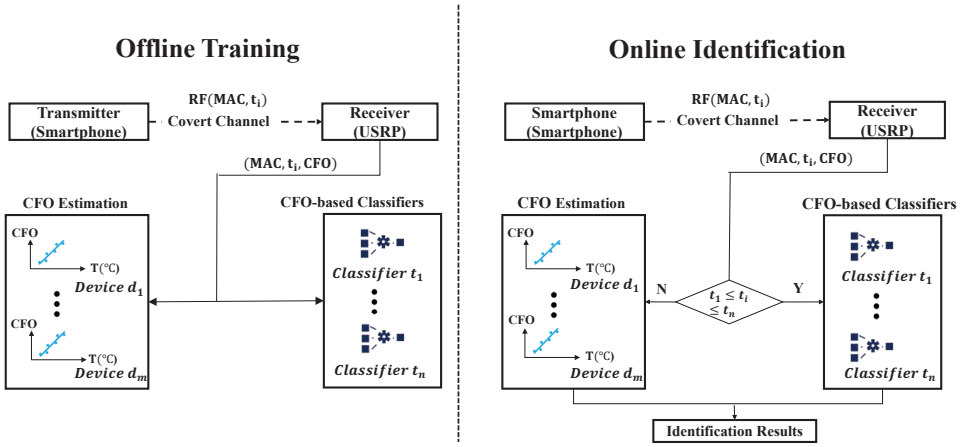


Fig. 5. TeRFF framework

keys as legitimate devices to enter the internal network behind the Wi-Fi Access Point (AP). Meanwhile, a software-defined radio device like USRP captures Wi-Fi signal transmitted from the smartphone and then extracts RF fingerprint. Furthermore, the RF fingerprint is sent to the authentication server and its corresponding device is identified by the pre-trained classifiers. If the strong identifier declared by the device matches the result identified by the classifier, the authentication server sends a message to AP which accepts the device's request to access the internal network. Otherwise, the authentication server sends a message which rejects the device's request.

B. TeRFF Framework

In order to protect the security of other devices inside the network, our TeRFF framework is divided into two stages, which are offline training for classifiers and online identification for smartphones as shown in Fig. 5:

- **Offline Training**: Software-defined radio is used to capture RF signal which modulates the MAC address and implicitly hides the temperature in the link layer. In addition, when the receiver captures RF signal, it can extract the CFO with Equation 6. Next, a CFO-based Naive Bayes Classifier which is tagged by t_i is constructed for training model. At the same time, the triple tuple (MAC, t_i, CFO) is used to conduct curve fitting for CFO estimation.
- **Online Identification**: Similar to the operation in the offline training, the demodulation and decoding are used to obtain the MAC address and temperature in the Wi-Fi packets. The receiver also extracts CFO with Equation 6. For the smartphone identification, if the temperature in the triple tuple is between minimum temperature t_{min} and maximum temperature t_{max} in the set of registered temperatures, the CFO-based Naive Bayes Classifier t_i is selected for smartphone identification. Otherwise, temperature t_i is used to estimate the CFO of all devices by their corresponding fitting curve. Then the device which has an estimated CFO value closest to the captured CFO is considered as the identification result.

In our proposed method, smartphones are required to report their own temperatures to the authentication server. Since temperature is closely related to the smartphone, the leakage of temperature will increase the success rate of forgery attack. To hide the secret about temperature, we design a simple and effective convert channel. Because an ICMP message is very common and inconspicuous in the Wi-Fi traffic, it is suitable as a carrier of covert information. In this paper, the length of the ICMP message is chosen to indicate device's temperature.

C. The Acquisition of Crystal Oscillator's Temperature

Obviously, we need to obtain the temperature of the crystal oscillator in Wi-Fi modules for next steps. According to experimental analysis, it is found that there are not any temperature sensors to measure the crystal oscillator directly. In addition, due to the limited size of the smartphone, it is difficult to measure the temperature of the crystal oscillator with extra electrical components. Fortunately, there are several temperature sensors in the smartphone, which are mainly used to monitor the working condition of the smartphone and control the heating degree of the smartphone. In order to measure the temperature of crystal oscillator, temperature sensors of the smartphone can be chosen as alternative sensors. According to the physical property of the crystal oscillator, the growth of its temperature is caused by radiation and conduction of the heat which comes from the CPU and external environment so that the temperature of the crystal oscillator varies more slowly than the CPU. In addition, the power of the CPU is greater than that of the crystal oscillator when the CPU load is high [14]. It means that temperature of the crystal oscillator will be lower than that of the CPU. According to the prior analysis, the selection of temperature sensors can be chosen based on the following rules:

1) The temperature of the crystal oscillator in the Wi-Fi chipset varies slowly.

2) The temperature of the crystal oscillator is much lower than that of the CPU when the smartphone runs a compute-intensive application.

According to the rules above, we can choose an alternative temperature sensor in the smartphone to represent the temperature of the crystal oscillator approximately. To be specific, the resolution of the temperature sensor achieves single digits so that the temperature in our paper is a discrete integer. To collect RF samples at high temperature within 40 to 60°C, we make the smartphone actively run computationally intensive programs and send packets to AP at the same time. By contrast, the state of crystal oscillator at low temperature is not stable which will decrease the number of RF samples at low temperature. To speed up the collection of RF samples at low temperature within 20 to 40°C, we leverage the PID algorithm [15] which controls the packet sending rate to automatically keep the temperature of crystal oscillator constant. For the purpose of making training samples balanced, a fast collection of the smartphone's RF fingerprints at high and low temperature can be accomplished. In addition, the temperatures in training period are added to temperature set S_{temp}. We call the temperature in S_{temp} as registered temperature and other temperature as unregistered temperature.

D. Temperature-aware Classification Based on CFO

The CFO of every frame is calculated from STS. It is found that the CFO is in Gaussian distribution. A simple proof is given as follows:

Property 1. *The distribution of $\bar{\alpha}$ is Gaussian Distribution.*

Proof. It is assumed that every coarse CFO random variable in STS is independent and identically distributed. In addition, α_{c_n} represents the coarse CFO calculated by two short OFDM symbols n and $n+1$. According to central limit theorem (CLT) [16], the variable α_{c_n} is Gaussian distributed separately and $\bar{\alpha}$ is calculated as follows:

$$\bar{\alpha} = \frac{\sum_{n=0}^{8} \alpha_{c_n}}{9}, \quad \alpha_{c_n} \sim N(\mu_n, {\sigma_n}^2) \tag{7}$$

Depending on the Lévy-Cramér theorem [17], it is proved that CFO $\bar{\alpha}$ is also Gaussian distributed.

$$\bar{\alpha} \sim N(\frac{\sum_{n=0}^{8} \mu_n}{9}, \frac{\sum_{n=0}^{8} {\sigma_n}^2}{81}) \tag{8}$$

□

To reduce computational cost of training models, Naive Bayes classifier is used to make a classification of Wi-Fi devices.

Since CFO is a continuous variable, it is reasonable that probability density function of CFO is leveraged to represent the probability of CFO's occurrence approximately. In the training period, the mean μ_i^j and variance $(\sigma_i^j)^2$ can be estimated from a amount of CFO sampling points of smartphone d_i with temperature t_j. In addition, the temperature in the smartphone is a concrete variable. Hence, when the device is at temperature t_j, the prior probability of CFO α is shown as below:

$$P^j(C = \alpha | D = d_i) = \frac{1}{\sqrt{2\pi}\sigma_i^j} e^{-(\alpha - \mu_i^j)^2 / 2(\sigma_i^j)^2} \tag{9}$$

When a CFO α_c is extracted by the steps above in Section II, it is used to calculate a posterior probability of the smartphone d_i with temperature t_j:

$$P^j(D = d_i \mid C = \alpha_c) = \frac{P^j(C = \alpha_c \mid D = d_i) P^j(D = d_i)}{P^j(C = \alpha_c)} \tag{10}$$

Then, an estimated device $\hat{d}$ with the largest posterior probability is searched in the device set S_d. It is assumed that $P^j(D)$ is in uniform distribution, and $P^j(C)$ is not related to a device itself. In other words, a comparison between posterior probabilities is equivalent to a comparison between prior probabilities. As a result, the identified device is determined as below:

$$\begin{aligned} \hat{d} &= \underset{d \in S_d}{argmax}\, P^j(D = d \mid C = \alpha_c) \\ &= \frac{P^j(D = d)}{P^j(C = \alpha_c)} \underset{d \in S_d}{argmax}\, P^j(C = \alpha_c \mid D = d) \\ &= \underset{d \in S_d}{argmax}\, P^j(C = \alpha_c \mid D = d) \end{aligned} \tag{11}$$

E. Estimation of CFO at Unregistered Temperatures

In the stage of fingerprint collection, an ideal case is that CFOs at any temperature reached by the smartphone are all captured. However, the temperature of the smartphone is affected by a working condition of the smartphone and the ambient temperature. It is obvious that the working condition can be adjusted but ambient temperature can not be controlled manually. Therefore it is very difficult for smartphones to traverse all possible temperatures in the stage of sample collection.

Fortunately, it is found that the CFO of each smartphone will decrease monotonously with the temperature in practical experiments. Therefore a potential relationship between CFO and temperature can be used for the estimation of smartphone fingerprints. To be specific, the CFO at unregistered temperatures can be roughly estimated through regression analysis.

The process of TeRFF is shown in Algorithm 1. In detail, it is assumed that the fingerprint dataset $FP = \left\{ \left(\mu_i^j, \sigma_i^j \right) \mid d_i \in \boldsymbol{D}, t_j \in \boldsymbol{T_r} \right\}$, where $\boldsymbol{D}$ and $\boldsymbol{T_r}$ represents the set of smartphones and registered temperatures separately. The appropriate function type in candidate function type set $\boldsymbol{F} = \{f_1, f_2, ..., f_{N_F}\}$ is needed to use for CFO estimation in the set of unregistered temperatures $\boldsymbol{T_u}$. The registered temperature set $\boldsymbol{T_r}$ is divided into two parts. In one part of registered temperature set $\boldsymbol{T_{r_1}}$, the mean value of CFO of the device d_i is leveraged to estimate parameters θ_i^k of function f_k. In addition, the mean value of CFO in the other part of registered temperature set $\boldsymbol{T_{r_2}}$ is used for selection of the regression function. Firstly, the parameters θ_i^j of function f_j of device d_i are estimated by the least square method as follows:

$$\theta_i^j = \min \sum_{t_k \in \boldsymbol{T_{c_1}}} \left(\mu_i^k - f_j(t_k, \theta)\right)^2 \tag{12}$$

In the next process, the best function type in $\boldsymbol{F}$ is chosen to improve the performance of CFO estimation. Mean Absolute Error (MAE) is a common kind of evaluation metric in estimation, which is resistant to outliers of CFO sampling points in the fingerprint dataset. In addition, $\boldsymbol{X_{d_i,t_j}}$ indicates the real CFO set of the device d_i at the temperature t_j. Therefore the best function type is decided by the following expression:

$$\begin{aligned} f_{best} = \min & \sum_{d_i \in \boldsymbol{D}} \sum_{t_j \in \boldsymbol{T_{r_2}}} \sum_{x \in \boldsymbol{X_{d_i,t_j}}} \left| f_k(\theta_i^k, t_j) - x \right| \\ s.t. \quad & f_k \in \boldsymbol{F} \end{aligned} \tag{13}$$

In this paper, function type set $\boldsymbol{F}$ includes linear polynomial, second-degree polynomial and third-degree polynomial. Then the linear polynomial and corresponding parameter are chosen to estimate the CFO at unregistered temperatures.

Algorithm 1 The process of TeRFF

Input: fingerprint dataset FP, CFO sample x, temperature t, device set $D\{d_0, d_1, ..., d_n\}$, best function type f, registered temperature set T_r, unregistered temperature set T_o, parameter set Θ of function type f, index searching function $f' : \mathbb{Z} \to \mathbb{N}$

Output: identified device d_i

```
if t ∈ T_r then
    d ← d_0, p ← 0, m ← f'(t)
    for μ_i^m, σ_i^m in FP do
        if p > pdf(μ_i^m, σ_i^m, x) then
            p ← max(p, pdf(μ_i^m, σ_i^m, x)), d ← d_i
        end if
    end for
else if t ∈ T_u then
    dis ← +∞
    for θ_i in Θ do
        if |f(θ_i, t) − x| < dis then
            dis = |f(θ_i, t) − x|
            d ← d_i
        end if
    end for
end if
```

After determining the parameters and type of function, we can estimate the mean value of the device's CFO at any unregistered temperature among devices which can be used as the reference fingerprints. Compared with the fingerprints collected directly at the registered temperatures, the estimated CFO achieves less accuracy. Because the estimated function only calculates a mean value of CFO which can not describe the CFO comprehensively. Therefore it is more reliable to leverage the Gaussian distribution of CFO than to use a mean value of CFO for smartphone identification. However, practical experiments below have confirmed that these estimated CFOs are still effective in smartphone identification.

IV. Evaluation

A. Experiment Setup

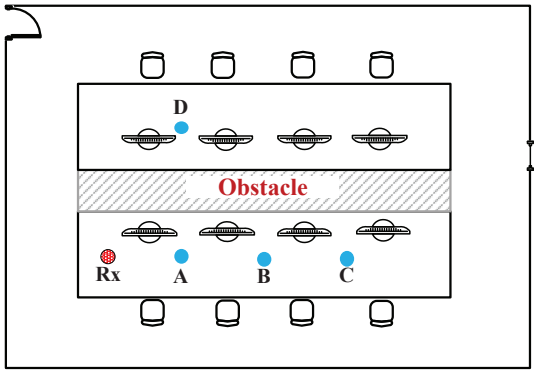


Fig. 6. This is our experimental scenario where point Rx is USRP N210. Besides, point A, B, C and D indicate four deployment locations of smartphones.

The experimental environment is an office in our university, which is full of electromagnetic interference. Besides, it is a typical indoor environment. The layout of the office is shown in Fig. 6. The model types and quantities of smartphones are shown in Table I, including multiple models of smartphones and multiple smartphones under one model. Each alternative sensor for extracting crystal oscillator's temperature in a smartphone is also listed in Table I. In addition, the front-end of the receiver is USRP N210 platform, and gr-ieee802-11 [18] project based on GNU Radio is modified to extract CFO from Wi-Fi signals.

TABLE I
THE LIST OF SMARTPHONES AND THEIR ALTERNATIVE SENSORS

Device Model	Alternative Sensor	Device Number
Nexus 5	pa_therm0	9
Huawei Nexus 6P	pa_therm1	1
Nexus 5X	pa_therm0	1
Huawei P9	pa_0	1
Vivo X9	pa_therm0	1
OnePlus 3	pa_therm0	1
Huawei Honor 7	pa_0	1

Given a RF sample from smartphones, accuracy refers to the proportion of correctly identified RF samples in all RF samples.

$$Accuracy = \frac{N_{right}}{N_{all}} \tag{14}$$

B. Effects of Temperature

We investigate the impact of whether considering temperature is necessary. The RF samples of all smartphones are captured at registered temperatures between 26°C and 33°C. In detail, the number of RF samples emitted by each smartphone at each registered temperature is 3000. The entire samples are split into training and test sets in 2:1 allocation ratio. To simplify the experiment, we assume a common case: the period of training sample collection is short and the temperature of the smartphone varies little. Therefore, the classifier is

trained by training samples at a certain temperature. Firstly, we train eight classification models in eight temperatures separately. When the temperature of device is considered, the corresponding model can be selected for smartphone identification. On the other hand, when smartphone's temperature is not under consideration, the temperature corresponding to testing samples is possibly not equal to the temperature of Naive Bayes model so that the performance of smartphone identification will degrade. As it is shown in Fig. 7, it is found that the identification method with temperature (i.e. TeRFF) is obviously greater than that without temperature (i.e. Naive Bayes approach). The experiment illustrates that temperature is one of the most significant factors which can influence smartphone identification. Therefore the identified smartphone would report its own temperature to the authentication server for a better performance.

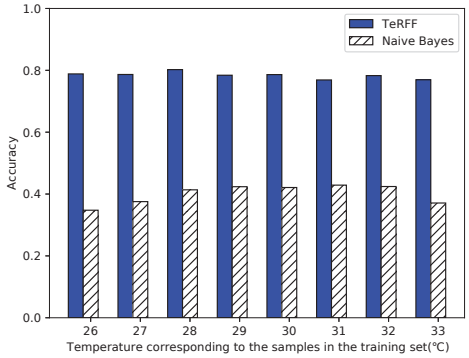
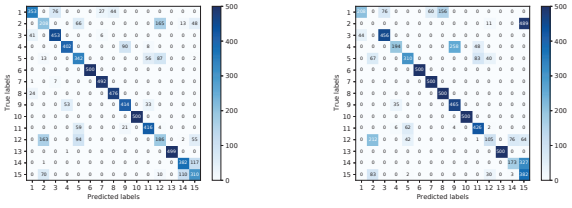


Fig. 7. Accuracy of TeRFF

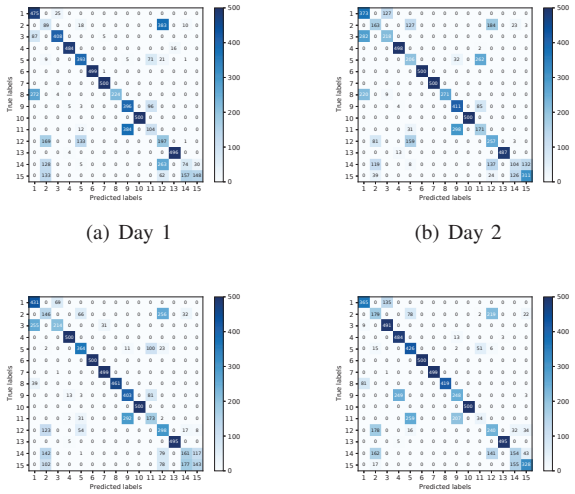


(a) Real CFO (b) Estimated CFO

Fig. 8. Effectiveness of estimated CFO

C. Effectiveness of Estimated CFO

We validate the effectiveness of estimated CFO at unregistered temperatures. As Fig. 8 shows, the accuracy of smartphone identification at unregistered temperatures achieves at 69.59%. Compared to the CFO-based method at registered temperatures whose accuracy reaches 79.10%, that at unregistered temperatures has a lower performance. However, it is obvious that the estimated CFO is also effective in smartphone identification when the smartphone's temperature is an unregistered temperature.



(a) Day 1 (b) Day 2

(c) Day 3 (d) Day 4

Fig. 9. Four-period(day) experimental results

D. Stability over Time

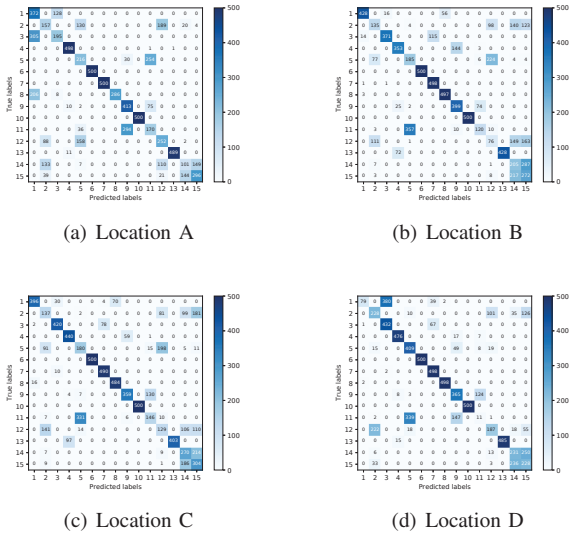
We collect RF samples in four days between December 2021 and January 2022. Furthermore, 500 samples are collected per day. The performance of smartphone identification in different days is shown in Fig. 9. It is found that the accuracy of smartphone identification on four days is 66.49%, 66.27%, 70.51% and 71.49% separately. It is shown that the accuracy can remain stable and can achieve a tolerable accuracy.

E. Effects of Different Locations

To further verify the effectiveness of our proposed method, we test whether the location change of smartphones will degrade the accuracy of smartphone identification. To be specific, location A is the farthest from the USRP receiver, location B is the next, and location C is the closest. The distance between the two close locations above is approximately 0.5 meters. Besides, We further put the smartphones in location D to investigate the impact of the obstacle on the performance of the smartphone identification method. As illustrated in Fig. 10, the accuracy of smartphone identification in location A to C are 65.93%, 66.23% and 68.77% which are very close. In addition, we also find that the obstacle doesn't degrade the performance of our proposed method which can achieve 68.36% in location D. Therefore our proposed method is robust to locations.

F. Multiple Consecutive Samples in the Network Traffic

We also explore whether utilizing multiple consecutive samples in a network traffic flow improves the performance of smartphone identification. The consecutive samples are captured in two days. The identification result of a traffic flow is determined by the majority of identification results in multiple consecutive samples. It is found that the low dimension of RF fingerprints limits the performance of smartphone identification. To be specific, the CFO of device 14 and



(a) Location A (b) Location B

(c) Location C (d) Location D

Fig. 10. Experimental results in different locations

15 can not be distinguished easily. Therefore, the accuracy of smartphone identification only increases from 68.39% to 69.80%, which does not show a noticeable performance boost.

G. Comparison with Baseline RF Fingerprinting

Finally, we compare our proposed method with state-of-art methods including LTF-based [6] and power spectral density(PSD)-based [10] methods. We randomly collect two sets of raw I/Q data as the training set and test set respectively. Furthermore, we use our proposed method and other two methods to identify smartphones. In Fig. 11, it is found that our proposed method can achieve the highest accuracy among three methods. We infer the performance of the other methods is easily degraded by wireless channel. To be specific, LTF-based method assumes the channel frequency response in a frame always is constant. However, it is only an approximation for demodulation and decoding of Wi-Fi signal, while the semi-steady features between two LTFs are easily overwhelmed by the channel variation within the duration of a frame. Besides, the PSD-based method only utilizes frequency-domain characteristic of RF signal which can not reflect the RF front-end impairments.

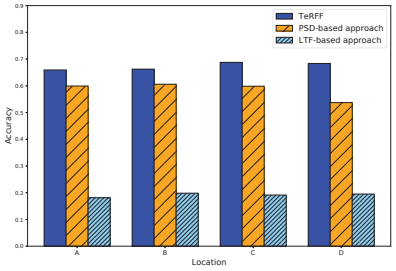


Fig. 11. Comparison of TeRFF and existing smartphone identification methods

V. Related Work

A. Device Fingerprinting

Device Fingerprinting is a significant approach which is often used in smartphone identification and authentication. Because it is more resistant to forgery attacks than traditional strong identity-based methods, many researchers have devoted their many efforts into identifying devices by constructing unique fingerprints from software or hardware. In terms of the software-based fingerprints, they are extracted from network traffic in the application layer. To be specific, the combination of distinct context in website browsers, chipsets, firmwares and drivers can be used as device fingerprints. In addition, the dynamic features such as traffic pattern and timing interval in probe request are also used to identify devices uniquely. However, these fingerprints above can be modified easily by malicious attackers. Therefore, hardware-based fingerprints have been proposed in recent years. The time of execution of readily available instructions in API function [19] and magnetic induction signal emitted from the CPU chips and affiliated power supply circuits [20] represent physical characteristic of CPU in devices. Besides, the attributes of lines on the surface of printed physical objects are determined by the feeder, positioner and hot end, which can be used in smartphone identification [21]. In addition, clock skews from 802.11 beacon or probe response frames can be used as a kind of device fingerprints [22]. In addition, the RF fingerprints in RF front-end of Zigbee [23] and Lora [24] devices can be extracted for smartphone identification during peer-to-peer network communication.

B. Temperature Effects on the Circuit

The electrical properties of a circuit are susceptible to variation in the temperature, which poses potential threats to the normal operation of the circuit. It is difficult for manufacturers to completely eliminate the negative effects caused by variation of environmental temperature on electronic components in the circuit. Therefore temperature compensation is studied for circuits' stability. A prototype CMOS power amplifier with temperature compensation is designed for suppressing the variation of gain [25]. Besides, frequency of a 7-MHz clock oscillator in a 0.25-μm process keeps constant by a voltage control [26]. Although variation of temperature deteriorates the performance of circuits, this seemly negative physical property also can promote some applications in an amazing way. Some researchers have exploited the temperature-sensitive characteristics of circuits that were not designed for temperature sensing to act as a low-precision thermometer. A Gaussian process model is designed to regress the relationship between phases and temperature, which can estimate the temperature with 5°C error [27]. Besides, for more accuracy of temperature measurement, the other environmental effects are canceled out by using a pair of RFID tags with different sizes [28]. In addition, internal electrical components of RFID circuit are used to measure the temperature without other electronic components' assistance [29].

VI. Conclusion

In this paper, we propose an enhanced CFO-based fingerprinting solution TeRFF to overcome the challenge that the CFO fingerprint of smartphone is unstable under different crystal oscillator's temperatures. First, we establish a dataset containing crystal oscillator's temperature and the corresponding CFO value on multiple smartphones over a period. To be specific, a set of temperature values is combined with the extracted temperature values, and when a smartphone is at the registered temperature, its corresponding CFO values are stored in the dataset. On this basis, we train multiple Naive Bayes models, each tagged with a registered temperature value. Moreover, in order to solve the problem that the temperature values are not in the temperature set during the smartphone identification, we design a CFO estimation method to estimate the CFO fingerprint at the unregistered temperature. Finally, it is demonstrated by actual experiments that our proposed solution TeRFF makes the CFO fingerprinting still effective for smartphone identification, and TeRFF has a better performance than other existing RF fingerprinting schemes.

VII. Acknowledgment

This work was partially supported by the National Key R&D Program of China (No. 2018YFB2100300); the National Natural Science Foundation of China (Nos. 62072102, 62132009, 62072103, 62022024, and 61972088); Jiangsu Provincial Natural Science Foundation for Excellent Young Scholars (No. BK20190060); Jiangsu Provincial Key Laboratory of Network and Information Security (No. BM2003201); the Key Laboratory of Computer Network and Information Integration of the Ministry of Education of China (No. 93K-9), the Fundamental Research Funds for the Central Universities (No. 2242022k30029).

References

[1] MarketsandMarkets. Wi-Fi Market Report. https://www.marketsandmarkets.com/Market-Reports/global-wi-fi-market-994.html, 2021.

[2] V. Brik, S. Banerjee, M. Gruteser, and S. Oh. Wireless Device Identification with Radiometric Signatures. In *Proceedings of the 14th ACM international conference on Mobile computing and networking*, pages 116–127, 2008.

[3] M. K. Fadul, D. R. Reising, and M Sartipi. Identification of OFDM-based Radios under Rayleigh Fading using RF-DNA and Deep Learning. *IEEE Access*, 9:17100–17113, 2021.

[4] A. Al-Shawabka, F. Restuccia, S. D'Oro, T. Jian, B. C. Rendon, N. Soltani, J. Dy, S. Ioannidis, K. Chowdhury, and T. Melodia. Exposing the fingerprint: Dissecting the Impact of the Wireless Channel on Radio Fingerprinting. In *IEEE INFOCOM 2020-IEEE Conference on Computer Communications*, pages 646–655. IEEE, 2020.

[5] K. Sankhe, M. Belgiovine, F. Zhou, L. Angioloni, F. Restuccia, S. D'Oro, T. Melodia, S. Ioannidis, and K. Chowdhury. No Radio Left Behind: Radio Fingerprinting Through Deep Learning of Physical-Layer Hardware Impairments. *IEEE Transactions on Cognitive Communications and Networking*, 6(1):165–178, 2019.

[6] G. Li, J. Yu, Y. Xing, and A. Hu. Location-invariant Physical Layer Identification Approach for WiFi Devices. *IEEE Access*, 7:106974–106986, 2019.

[7] J. Hall, M. Barbeau, and E. Kranakis. Radio Frequency Fingerprinting for Intrusion Detection in Wireless Networks. *IEEE Transactions on Defendable and Secure Computing*, 12:1–35, 2005.

[8] M. Köse, S. Taşcioğlu, and Z. Telatar. RF Fingerprinting of IoT Devices Based on Transient Energy Spectrum. *IEEE Access*, 7:18715–18726, 2019.

[9] T. D. Vo-Huu, T. D. Vo-Huu, and G. Noubir. Fingerprinting Wi-Fi Devices using Software Defined Radios. In *Proceedings of the 9th ACM Conference on Security & Privacy in Wireless and Mobile Networks*, pages 3–14, 2016.

[10] W. C. Suski II, M. A. Temple, M. J. Mendenhall, and R. F. Mills. Using Spectral Fingerprints to Improve Wireless Network Security. In *IEEE GLOBECOM 2008-2008 IEEE Global Telecommunications Conference*, pages 1–5. IEEE, 2008.

[11] T. Zheng, Z. Sun, and K. Ren. FID: Function Modeling-based Data-Independent and Channel-Robust Physical-Layer Identification. In *IEEE INFOCOM 2019-IEEE Conference on Computer Communications*, pages 199–207. IEEE, 2019.

[12] M. Frerking. *Crystal Oscillator Design and Temperature Compensation*. Springer Science & Business Media, 2012.

[13] Wikipedia. Orthogonal Frequency-division Multiplexing. https://en.wikipedia.org/wiki/Orthogonal_frequency-division_multiplexing, 2022.

[14] A. Carroll and G. Heiser. An Analysis of Power Consumption in a Smartphone. In *USENIX Annual Technical Conference*. USENIX Association, 2010.

[15] K. H. Ang, G. Chong, and Y. Li. PID control system analysis, design, and technology. *IEEE transactions on control systems technology*, 13(4):559–576, 2005.

[16] H. Fischer. *A History of the Central Limit Theorem: From Classical to Modern Probability Theory*. Springer Science & Business Media, 2010.

[17] K. B. Athreya and S. N. Lahiri. *Measure theory and probability theory*, volume 19. Springer, 2006.

[18] B. Bloessl, M. Segata, C. Sommer, and F. Dressler. An IEEE 802.11 a/g/p OFDM Receiver for GNU Radio. In *Proceedings of the second workshop on Software radio implementation forum*, pages 9–16, 2013.

[19] I. Sanchez-Rola, I. Santos, and D. Balzarotti. Clock Around the Clock: Time-Based Device Fingerprinting. In *Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security (CCS'18)*, pages 1502–1514, 2018.

[20] Y. Cheng, X. Ji, J. Zhang, W. Xu, and Y. Chen. DeMiCPU: Device Fingerprinting with Magnetic Signals Radiated by CPU. In *Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security*, pages 1149–1170, 2019.

[21] Z. Li, S. Rathore, A, C. Song, S. Wei, Y. Wang, and W. Xu. PrinTracker: Fingerprinting 3D printers using Commodity Scanners. In *Proceedings of the 2018 ACM sigsac conference on computer and communications security*, pages 1306–1323, 2018.

[22] S. Jana and K. Kasera, S. On Fast and Accurate Detection of Unauthorized Wireless Access Points Using Clock Skews. *IEEE transactions on Mobile Computing*, 9(3):449–462, 2009.

[23] J. Yu, A. Hu, F. Zhou, Y. Xing, Y. Yu, G. Li, and L. Peng. Radio Frequency Fingerprint Identification Based on Denoising Autoencoders. In *2019 International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*, pages 1–6. IEEE, 2019.

[24] G. Shen, J. Zhang, A. Marshall, L. Peng, and X. Wang. Radio Frequency Fingerprint Identification for LoRa Using Spectrogram and CNN. In *IEEE INFOCOM 2021-IEEE Conference on Computer Communications*, pages 1–10. IEEE, 2021.

[25] T. Yoshida, K. Takano, C. Li, M. Motoyoshi, K. Katayama, S. Amakawa, and M. Fujishima. CMOS Power Amplifier with Temperature Compensation for 79 GHz Radar System. In *2013 Asia-Pacific Microwave Conference Proceedings (APMC)*, pages 239–241. IEEE, 2013.

[26] K. Sundaresan, P. E. Allen, and F. Ayazi. Process and temperature compensation in a 7-MHz CMOS clock oscillator. *IEEE Journal of Solid-State Circuits*, 41(2):433–442, 2006.

[27] X. Wang, J. Zhang, Z. Yu, E. Mao, S. C. Periaswamy, and J. Patton. RFThermometer: A Temperature Estimation System with Commercial UHF RFID tags. In *ICC 2019-2019 IEEE International Conference on Communications (ICC)*, pages 1–6. IEEE, 2019.

[28] S. Pradhan and L. Qiu. RTSense: Passive RFID based Temperature Sensing. In *Proceedings of the 18th Conference on Embedded Networked Sensor Systems*, pages 42–55, 2020.

[29] X. Chen, J. Liu, F. Xiao, S. Chen, and L. Chen. Thermotag: Item-level Temperature Sensing with a Passive RFID Tag. In *Proceedings of the 19th Annual International Conference on Mobile Systems, Applications, and Services*, pages 163–174, 2021.