# Wireless Device Identification with Radiometric Signatures

Vladimir Brik [*†], Suman Banerjee [†]
University of Wisconsin at Madison
Department of Computer Sciences
1210 W. Dayton St., Madison, WI 53706
{vladimir,suman}@cs.wisc.edu

Marco Gruteser [‡], Sangho Oh [‡]
Rutgers University
WINLAB
671 Route 1 South, North Brunswick, NJ 07310
{gruteser,sangho}@winlab.rutgers.edu

## ABSTRACT

We design, implement, and evaluate a technique to identify the source network interface card (NIC) of an IEEE 802.11 frame through passive radio-frequency analysis. This technique, called PARADIS, leverages minute imperfections of transmitter hardware that are acquired at manufacture and are present even in otherwise identical NICs. These imperfections are transmitter-specific and manifest themselves as artifacts of the emitted signals. In PARADIS, we measure differentiating artifacts of individual wireless frames in the modulation domain, apply suitable machine-learning classification tools to achieve significantly higher degrees of NIC identification accuracy than prior best known schemes.

We experimentally demonstrate effectiveness of PARADIS in differentiating between more than 130 identical 802.11 NICs with accuracy in excess of 99%. Our results also show that the accuracy of PARADIS is resilient against ambient noise and fluctuations of the wireless channel.

Although our implementation deals exclusively with IEEE 802.11, the approach itself is general and will work with any digital modulation scheme.

## Categories and Subject Descriptors

C.2.1 [**Computer-communication networks**]: Network Architecture and Design—*Wireless communication*; K.6.5 [**Management of computing and informations systems**]: Security and Protection—*Authentication, Unauthorized access*

## General Terms

Design, Experimentation, Measurement, Security

## Keywords

802.11, wireless, security, PHY, RF, fingerprinting, signature, radiometric, authentication, security token

## 1. INTRODUCTION

Device identity management is, perhaps, one of the most significant challenges in any network security solution. Since the source MAC address in a frame is easy to forge, administrators need other mechanisms to identity the source of frames within their networks. In a wired network, switches provide the capability to distinguish traffic based on the incoming port, each mapped to a single Ethernet jack in the wall. In contrast, the untethered nature of wireless communication makes similar identification of a frame's source difficult. To overcome this hurdle, 802.11 WLAN administrators rely on various cryptographic mechanisms for wireless device identity management and access control. However, the usefulness of such cryptographic schemes is, sometimes, impacted by various challenges of key distribution, and in particular by robustly detecting and revoking compromised keys. In this paper we introduce an approach and a prototype to accurately identify the source network interface card (NIC) of a wireless frame, and propose the use of this approach in various identity management and security applications.

Our approach is based on the notion of *radiometric identity*: minor variations in analog hardware of transmitters are manifested as idiosyncratic artifacts in their emitted signals and thus can be used to identify a signal's device-of-origin. As the name implies, biometric and radiometric identities are similar, inasmuch as both define identity as the set of observable characteristics that are the manifestations of the individual's constitution.

### 1.1 Radiometric identification

We define and use the term radiometric identification instead of the more commonly adopted term, radio-frequency
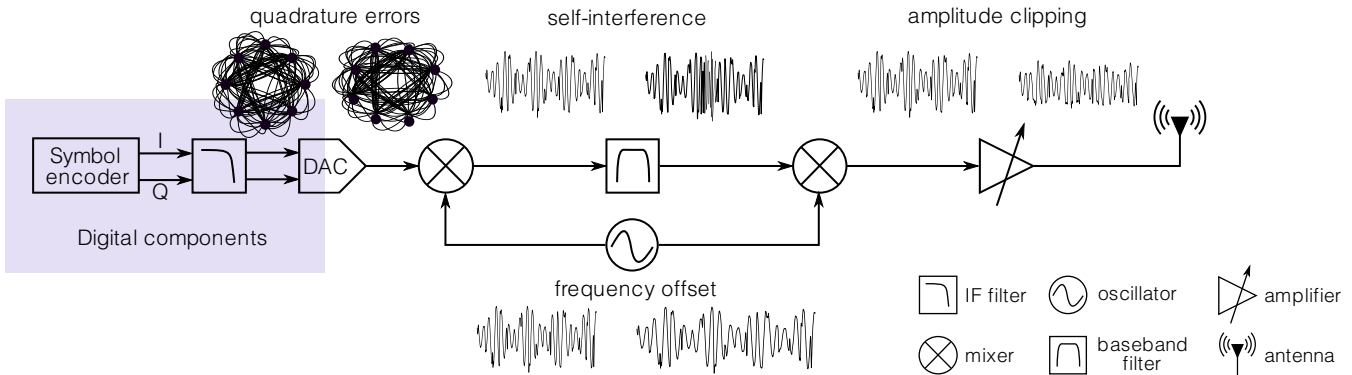
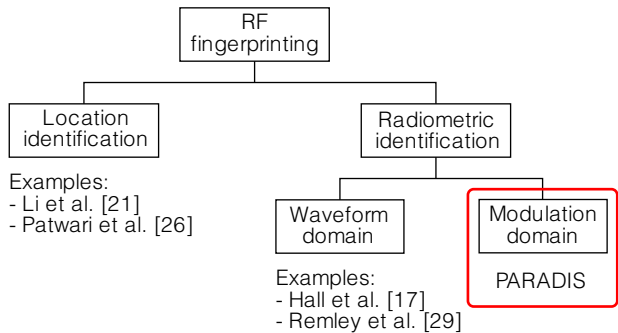**Figure 1: Common transmitter impairments and their sources**



**Figure 2: Radiometric identification and PARADIS**

(RF) fingerprinting because the latter has a wider meaning and usage than our intent. The term RF fingerprinting, in general, refers to various PHY layer classification approaches of RF signals. We broadly classify RF features into: (i) channel-specific ones, e.g., channel impulse response, that characterize the wireless channel; and (ii) transmitter-specific ones that are independent of the channel, e.g., signal encoding.

Since channel-specific features uniquely identify the channel between the transmitter and the receiver, they have been successfully adopted in robust location distinction [23, 28]. In contrast for radiometric identification, we need to ignore such channel-specific features and focus on the transmitter-specific ones.

We classify possible radiometric techniques based on the fundamental differences in how they treat signals as those operating either in the *waveform domain* or the *modulation domain* as shown in Figure 2. Waveform domain techniques use signal samples from time and frequency domains as the basic blocks of representation, which allows the most flexibility at the cost of complexity. Our approach falls under modulation domain techniques that represent signals at the most basic level in terms of I/Q samples, whose interpretation depends on the underlying modulation scheme. Therefore, signals in modulation domain are more structured and better behaved, but require knowledge of the modulation scheme being used.

## 1.2 Transmitter individuality and PARADIS

Radiometric identification is possible because of benign

hardware imperfections inherent to the analog components of a NIC's transmit path, which exist due to normal variations in physical properties of such components. These imperfections are also called impairments because they cause the device's emissions to differ from the theoretically ideal output. Figure 1 presents a typical transmitter design and illustrates likely causes of common impairments [1, 2, 3, 4, 29].

In a sense, despite superficial sameness of NICs, even when constructed using the same manufacturing and packaging processes, no two are identical. While it may be possible to eliminate these hardware imperfections through more precise manufacturing and quality control, doing so can greatly increase costs. Manufacturers allow such impairments in their devices because they are benign in nature. In particular, a wireless transmitter, even with such minor impairments, continue to operate well within the tolerances specified by the corresponding communication standard, such as IEEE 802.11. This work will demonstrate that a transmitter's radiometric identity, or the unified effect of its impairments, can be observed in its radio transmissions and can be used to discern between different 802.11 NICs.

The first radiometric identification systems were developed to distinguish between friendly and enemy radars during the Vietnam War era. Since then, similar systems have been deployed by cellular networks to prevent access from unauthorized phones [32, 35]. However, due to the commercial and military nature of such systems few implementation details are available. Nevertheless, evidence suggests that such system are based on radiometric identification in waveform domain using signal transients [25], which are minute waveform structures that do not server a protocol function.

Use of transients for radiometric identification of 802.11 transmitters was studied by Barbeau, Hall and others [5, 17, 18]. However, using transients for identification appears to be difficult, as is indicated by imperfect performance of existing schemes even in modestly-sized evaluations. Instead, we leverage understanding of the PHY layer to bypass the challenges that hinder device identification using transients, and attain significantly higher accuracy.

In particular, we introduce an approach, called PARADIS, which stands for Passive RAdiometric Device Identification System. PARADIS quantifies radiometric identity of a transmitter on per-frame basis by comparing an observed signal to the ideal in the modulation domain. Radiometric identity of a frame in PARADIS is defined to be a real vector,
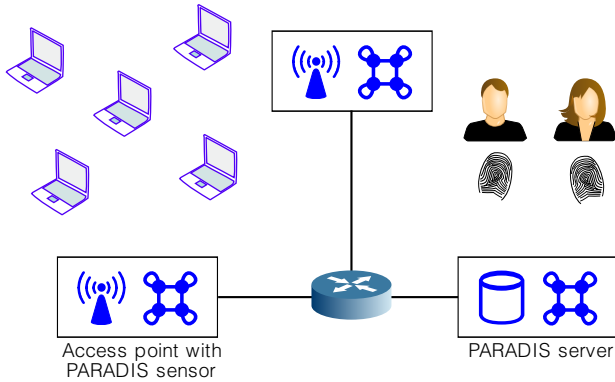
**Figure 3: PARADIS schematic**

whose elements correspond to the normalized magnitude of frequency, magnitude and phase errors, I/Q origin offset and SYNC correlation of the frame in question.

Note that although experimental section of this work deals exclusively with IEEE 802.11, PARADIS as a radiometric identification approach is more general and will work with any communications standard that uses digital modulation. This is because the metrics listed above are defined for any digital communication standard.

## 1.3 Applications of PARADIS

PARADIS is intended for applications in wireless security. Although capable of overall identification accuracy exceeding 99%, which by far outperforms prior literature, PARADIS is not perfectly accurate. Furthermore, we believe that no identification mechanism that relies solely on physical layer measurements of intercepted analog signals can deliver perfect accuracy because of environmental noise. Consequently, we envision PARADIS as a secondary security perimeter that detects breaches in the primary perimeter established, say, using the usual cryptographic mechanisms. We explain this with an example, next.

A standard wireless network access control mechanism involves authentication using secure cryptographic keys. The security of such approach is entirely dependent on the key being private. Should an unauthorized party come in possession of the key and gain network access, there is no direct way for the administrator to detect that they key has been compromised. This example illustrates one the biggest security challenges today: detection of compromised keys.

Figure 3 illustrates how PARADIS could be used to address this challenge. A PARADIS system consists of one or more sensors co-located with access points, and a server containing known radiometric identities. Before a new NIC is allowed to access the network, the administrator measures and records its radiometric identity.

If an unauthorized entity gains network access using a stolen a secret key, its traffic will be received by a PARADIS sensor. The sensor will compute the transmitter signature and send it for identification to the PARADIS server. If the radiometric signature does not match a NIC authorized to use the secret key in question, PARADIS will alert the administrator of possible security breach. If the breach is confirmed, the administrator may, for example, revoke the compromised key.

## 1.4 Salient features of PARADIS

The following are the key features of PARADIS:

- *Simplicity:* By operating in the modulation domain, PARADIS takes advantage on the structure of signals imposed by the modulation scheme. This makes it much easier to identify transmitter-specific properties of signals, and reformulates the problem is such a way that makes identification intuitive to understand and straightforward to implement.

- *Resilience:* Physical modulations schemes, e.g. OFDM, are designed explicitly to protect encoded data against adverse channel conditions. Therefore, signal representation using symbols in modulation domain is more stable and resilient to effects of noise that distorts raw waveforms. Because its input is more stable, PARADIS performs well in the presence of noise.

- *Robustness:* The features used for radiometric identification are based on minor hardware impairments. The cost and effort of creating a device to mimic hardware impairments is likely to make forgery difficult. Furthermore, circumvention of a PARADIS sensor will likely require far more expensive hardware than the sensor itself.

- *Ubiquity:* The entire identification scheme is passive in nature. Therefore, no single NIC that is attempting to communicate with the WLAN infrastructure can evade the radiometric identity test except by staying silent. Furthermore, PARADIS is completely transparent to users and does not require any modifications to client-side hardware and software.

Our evaluation demonstrates that PARADIS is able to *accurately distinguish more than one hundred thirty identically manufactured 802.11 NICs.* The evaluation was performed in realistic conditions with significant environmental noise. Furthermore, the results suggest resilience to location-dependent multipath effects, and show that PARADIS is able to recognize devices using signatures obtained over five months earlier.

## 2. RELATED WORK

First wireless transmitter identification systems were developed as early as the 1960s for military aircrafts to differentiate between friendly and enemy radars [35]. However, it is not clear whether such systems were effective and practical enough for the military to use for day-to-day operation in the field [6]. Nevertheless, similar transmitter identification systems have since been developed and used in the context of cellular networks [22, 25, 32].

A large body of literature is dedicated to the general issues of design, implementation and operation that are relevant to many kinds of identification systems, whether they identify radars, cellphones, people, or 802.11 transmitters. A comprehensive overview of high-level issues in the context of transmitter identification is presented by Talbot et al. [35], while a biometric perspective can be found in [33, 36].

## 2.1 Software-based fingerprinting

Multiple efforts have addressed the issue of distinguishing network nodes based on differences in software configuration. Many applications that, for example, determine the version of a node's operating systems have already established their

| Technique | Type | Goal | Identity Model | Evaluation scale |
|---|---|---|---|---|
| Kohno et al. [21] | Software meas. | hardware id | clock skew variation | n/a |
| Franklin et al. [13] | Software meas. | device driver id | compliance with 802.11 standard | 17 802.11 NICs |
| Faria et al. [12] | RF fingerprinting | location distinction | signal power attenuation | 135 locations |
| Patwari et al. [28] | RF fingerprinting | location distinction | multipath channel response | 44 locations |
| Hall [18] | RF fingerprinting | radiometric id | transient properties | 30 802.11 NICs |
| Gerges et al. [14] | RF fingerprinting | radiometric id | waveform accuracy | 16 Ethernet NICs |
| *PARADIS* | *RF fingerprinting* | *radiometric id* | *modulation accuracy* | *138 802.11 NICs* |

**Table 1: Comparison of PARADIS with related ideas**

place in the toolkit of network administrators [15]. Typically such tools are used to identify computers running vulnerable software.

In the context of IEEE 802.11 devices, Wright [38], Guo et al. [16] and others have discussed approaches to detect presence of multiple 802.11 devices using the same MAC address using analysis of frame sequence numbers. In a similar vein of research, Franklin et al. [13] developed a technique to identify devices based on differences in MAC layer behavior that depends on the combination of the chipset, firmware and the device driver being used. Specifically, this approach was based on observing differences in implementations of 802.11 protocol's MAC functionality. Pang et al. [27] studied techniques to identify users based on the patterns of their wireless traffic, such as packet sizes and and network destinations. Since all of these approaches are software-based, they can be circumvented by changing computer configuration or behavior.

Finally, Kohno et al. [21] proposed fingerprinting of networked devices based on estimation of its clock skew using TCP and ICMP time stamps. Unlike the previous approaches, here identity is based on properties of hardware. However, users can still interfere with identification using software means. For example, time stamps could be altered or disabled altogether.

## 2.2 Location distinction

Techniques for differentiating transmitter locations also differentiate transmitters, assuming they remain active and do not move. Location distinction systems effectively group frames by transmitter but lack the ability to actually make an identification or recognize a previously seen device that moved or sat silent for some time. However, this functionality could be useful to PARADIS as a mechanism to allow allow radiometric identification on groups of frames guaranteed to be from the same transmitter, without having to rely on MAC addresses.

Location distinction and secure localization techniques are based on distinguishing channel response, or environmental effects on signals that are unique to transmitter-receiver pairs. Faria et al. [12] evaluated use of signal strength to model channel response. Patwari et al. [28] evaluated an approach that compared multipath-related properties of signals. Finally, Li et al. [23] leveraged ideas of location distinction and use of probing for sensor authentication.

## 2.3 Radiometric identification

The approach to identification we took with PARADIS falls into the category of techniques that base the notion of identity on idiosyncratic hardware properties of transmitters. Perhaps the most intuitive and straightforward way of radiometric analysis was studied by Remley [31], who visu-

ally identified differences between signals of different 802.11 transmitters. Gerdes et al. [14] used physical layer signal processing to distinguished between wired Ethernet NICs based on fragments of their signals.

By far the most researched type of radiometric identification deals with the so-called signal transients. A transient is a brief radio emission produced while the power output of an RF amplifier goes from idle to the level required for data communication. The nature of transients is such that they are difficult to detect and there is no obvious correct way to succinctly describe them. The latter property is reflected in the amount of literature of the subject [5, 10, 17, 18, 30, 37]. These works study various techniques related to transient detection, data processing and machine learning. PARADIS outperforms transient-based techniques as will be discussed in the Section 5 and Table 3.

Table 1 is a summary of the most relevant related work, including data on scale of evaluation.

## 3. TECHNICAL DETAILS OF PARADIS

PARADIS stands out from other radiometric identification approaches because it leverages understanding of 802.11 PHY layer to reduce the complexity of the underlying problem.

For example, consider conversion of a transient waveform to a compact representation suitable to be input of an identification algorithm. The factors that determine a transient's shape are poorly understood, perhaps because they are of limited use, since, lasting under $2\mu s$ in 802.11 [20], they cannot serve a protocol function. Therefore, transients have to be treated essentially as arbitrary waveforms, and finding a compact representation that will be effective for all possible transients involves heuristics and guesswork. In contrast, modulation, by definition, gives a waveform a well-defined structure of limited complexity, making operations on it far easier.

Another potential advantage to working in modulation domain is the straightforward integration of PARADIS sensors with normal 802.11 hardware. Conceptually, the main difference between a regular 802.11 receiver and a PARADIS sensor is that the former only outputs data payload and discards intermediate calculations of the demodulation process. A PARADIS sensor, on the other hand, produces both demodulated payload and a subset of the intermediate results that can be used in software for identification. Therefore, a PARADIS sensor is essentially a specially instrumented, more precise version of a regular receiver.

Prior to presenting a detailed description of PARADIS, we give a quick overview of modulation mechanisms, and then explain specifics of the radiometric identification process.
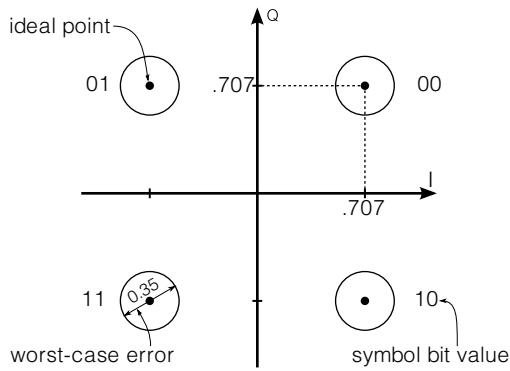
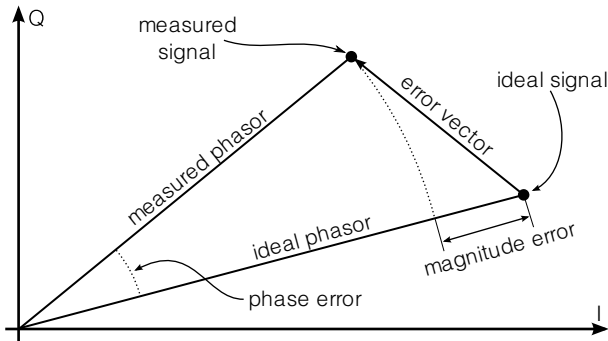**Figure 4: The 4 symbols of QPSK on I/Q plane**



**Figure 5: Modulation errors**

## 3.1 Background

The physical layer of the IEEE 802.11 standards use different I/Q modulation techniques for carrying bits across the wireless channel. As the name I/Q signifies, data is encoded using two independent carrier components, or sub-carriers. These sub-carriers are called in-phase (I) and quadrature(Q), because they are separated in phase by $\pi/2$.

Symbols of an I/Q modulation scheme are can be defined using a constellation diagram where different symbols are represented as points in I/Q space, or modulation domain. Depending on modulation scheme, a single symbol can encode multiple data bits. For example, in QPSK modulation each symbol encodes two data bits, as shown in Figure 4.

To give another example, consider sending a bit sequence 0010 using QPSK. First, the transmitter modulates the carrier wave to correspond to I/Q value of (.707, .707) to send the first two bits (00), and then transitions the carrier to (.707, -.707) to send the next two bits (10).

### 3.1.1 Errors in modulation domain

The transmitted RF signal experiences distortions due to, for example, hardware impairments, channel characteristics, and noise at the receiver. The goal of the receiving NIC is to determine the transmitted symbols despite all the RF distortions that exhibit themselves as errors in the modulation domain. Modulation errors are typically measured by comparing phasors, or vectors corresponding to the in-phase and quadrature values of a signal at the instant in time when a symbol was detected. In our context, phasors could be thought of as vector representations of symbols.

Relevant error metrics are explained below and illustrated in Figure 5.

- *Phase error:* the angle between the ideal and measured phasor.

- *Magnitude error:* the difference in magnitudes of the ideal and measured phasor.

- *Error vector magnitude (EVM):* the magnitude of the vector difference between the ideal and measured phasor.

Error metrics of just one symbol in a frame are not useful to us directly. Therefore, unless stated otherwise, we will use the above terms to describe the average errors across all symbols in a frame, rather than a specific symbol. In contrast, the following error metrics are only defined for an entire frame.

- *I/Q origin offset:* the distance between the origin of the ideal I/Q plane and the origin of the observed symbols.

- *Frequency error:* the difference between the ideal and observed carrier frequency. This is the amount by which the receiver's frequency had to be adjusted from channel center to achieve carrier lock.

- *SYNC correlation:* the correlation of I/Q values from an observed and the ideal SYNC, which is a short signal that precedes encoded data and is used to synchronize the transmitter and the receiver.

The IEEE 802.11 standards specify error tolerance for these metrics with respect to the ideal signal. For example, in QPSK modulation, the symbol error vector magnitude tolerance is 0.35. Similarly, the frame frequency error tolerance for the IEEE 802.11a standards is $\pm$ 20 ppm [19]. Hence, for 802.11a channel 36 centered at 5.180 GHz, valid frames need to have center frequency in the 207.2 KHz band around the channel center.

Note that although error tolerances may be specific to IEEE 802.11, the error metric we defined are general, and can be used with any standard that uses I/Q modulation, not just 802.11.

## 3.2 Features of radiometric identity

In our experiments we found that, overall, the following metrics could be used to establish radiometric identity. Ordered from most to least effective: (i) frequency error, (ii) SYNC correlation, (iii) I/Q offset, (iv) magnitude error, and (v), phase error. Best results were achieved when all the metrics were combined together.

Not surprisingly, we observed that classification accuracy can be improved if it is performed on multiple frames, rather than just one. This could be explained informally as follows. Distortions in a metric that are caused by transmitter hardware impairments should manifest themselves consistently across multiple frames from the same transmitter, while distortions caused by channel-specific and noise-related effects are likely to have a more random structure. Therefore, if we calculate statistical averages of the distortions of a metric, we expect to amplify the features caused by hardware impairments, while reducing the effects of the wireless channel and ambient noise.

| Error type | unit | reference | range | definition |
|---|---|---|---|---|
| frequency | Hz | 2142 MHz | ±60.3 | $\pm 25 ppm\ f_c$ |
| phase | ° | ideal symb | ±10 | $asin(E_{max})$ |
| magnitude | n/a | ideal symb | ±0.17 | $\pm E_{max}$ |
| EVM | n/a | ideal symb | [0, .35] | upto $2E_{max}$ |
| I/Q offset | n/a | ideal origin | [0, 0.17] | upto $E_{max}$ |
| SYNC | % | max corr. | [0, 1] | correlation |

$f_c$ – channel frequency    $E_{max}$ – max I/Q error

**Table 2: Modulation error metrics**

For the purposes of the evaluation we assumed that frames from different transmitters cannot be mistakenly grouped into a single bin. In practice, perhaps the most reliable way to achieve this is to combine PARADIS with a secure location distinction system [28, 23] that would ensure that all frames in a group came from the same location. Such enhancement may not significantly increase hardware cost, since the same hardware PARADIS uses can also be used for location distinction.

## 3.3 Radiometric signatures in practice

So far we have kept the notion of radiometric identity fairly abstract. We will now explain how signatures computed and represented in practice.

When a PARADIS sensor receives a wireless frame, the sensor analyzes the frame's modulation accuracy and reports the metrics we have chosen to use in signatures. For practical reasons, it is advantageous to scale and normalize data before applying a classification algorithm. Therefore, in our implementation a signature of a frame is a real vector whose five elements correspond to our metrics, and range in value from 0 to 1.

Normalization is performed according to the valid range of a given error metric, which is ultimately defined by the communication standard. If a metric exceeds this range, the corresponding frame is considered invalid. Table 3.3 summarizes valid ranges of modulation accuracy metrics for IEEE 802.11b 2Mbps QPSK modulation on channel 1. Note that EVM is not used by PARADIS and is included for completeness only. See Figure 4 and Figure 5 for intuition on derivation of valid ranges.

## 3.4 Operational details

At its core PARADIS performs identification using a classifier that converts signatures to NIC MAC addresses. We implemented two rather different radiometric classifiers, one using support vector machines algorithm (SVM) [7, 9, 11] and the other using the $k$-nearest-neighbor (kNN) [24] algorithm. Our evaluation showed the SVM algorithm to be more effective than kNN, though over 5 times slower. The main reason for inclusion of the kNN results is that the algorithm is so simple that its results can give insight into the structure of the underlying data. For example, kNN does no data pre-processing, while SVM maps its input onto a higher-dimensional space. For clarity, we will refer to our implementations of these algorithms as PARADIS-kNN and PARADIS-SVM, respectively.

It is worth pointing out that it was not our goal to find the best-performing algorithm for radiometric classification. Since PARADIS-kNN performed well, and PARADIS-SVM came close to perfect accuracy, for the purposes of this work,

we did not see the need to evaluate neither alternative algorithms, nor alternative classifier implementations.

From the implementation standpoint, PARADIS-kNN and PARADIS-SVM are typical data-agnostic classifiers that took one or more real vectors corresponding to signatures and identified corresponding NIC as an integer between 1 and the number of NICs represented in the training set. Rather than repeat the details of implementation that can be found in references [7, 9, 11, 24], we will illustrate operation of PARADIS on an example. Those familiar with the basics of kNN and SVM algorithms may want to skip the corresponding sections below.

**Training stage.** Before PARADIS can identify a wireless NIC its signature needs to be recorded. To do this, the WLAN administrator simply has the NIC transmit frames for a short period of time. We found that in practice 20 frames are a sufficient training set.

For simplicity, suppose that NIC $\mathcal{A}$ broadcasts four frames with the following radiometric signatures, as defined in Section 3.3:

$$\mathcal{A}_1 = [0.1, 0.9, 0.3, 0.7, 0.6] \qquad \mathcal{A}_3 = [0.1, 0.8, 0.1, 0.6, 0.4]$$
$$\mathcal{A}_2 = [0.2, 0.7, 0.1, 0.6, 0.4] \qquad \mathcal{A}_4 = [0.1, 0.9, 0.9, 0.7, 0.6]$$

**PARADIS-kNN:** In order to train, this classifier requires two kinds of functionality: signature manipulation using matrix operations, and a distance function that computes a scalar measure of similarity between two signatures. We used the $l_1$ or Manhattan distance, which is the sum of absolute values of component-wise differences between two signatures.

During training PARADIS-kNN first discards outliers one-by-one until only half of the training signatures remain. If there are $n$ signatures, an outlier is defined as the signature farthest away from the average:

$$\arg \max_{\mathcal{A}_i} l_1 \left( \frac{1}{n} \sum_j \mathcal{A}_j - \mathcal{A}_i \right)$$

In our example, $\mathcal{A}_4$ and $\mathcal{A}_1$ are the outliers:

$$\arg \max_{\mathcal{A}_1, \mathcal{A}_2, \mathcal{A}_3, \mathcal{A}_4} l_1 \left( \frac{1}{4} \left( \mathcal{A}_1 + \mathcal{A}_2 + \mathcal{A}_3 + \mathcal{A}_4 \right) - \mathcal{A}_i \right) = \mathcal{A}_4$$

$$\arg \max_{\mathcal{A}_1, \mathcal{A}_2, \mathcal{A}_3,} l_1 \left( \frac{1}{3} \left( \mathcal{A}_1 + \mathcal{A}_2 + \mathcal{A}_3 \right) - \mathcal{A}_i \right) = \mathcal{A}_1$$

The remaining training signatures, $\mathcal{A}_1$ and $\mathcal{A}_2$, constitute the kNN model for NIC $\mathcal{A}$ that later will be used for transmitter identification.

**PARADIS-SVM:** The exact training process can be found in [7, 9, 11] and is too involved to describe here. Instead, we will give a simplified view of the underlying SVM model.

At the most basic level the training process in PARADIS-SVM involved calculation of a special matrix $W_\mathcal{A}$ such that $W_\mathcal{A} \cdot \mathcal{S} - b$ is positive whenever signature $\mathcal{S}$ is from a frame sent by NIC $\mathcal{A}$, and negative otherwise. PARADIS-SVM classifier can be thought of as consisting of multiple layers of groups of matrices like that. To learn the signature of $\mathcal{A}$, PARADIS-SVM would construct a number of matrices that, given a signature, would return a measure of similarity between it and known signatures of $\mathcal{A}$. Then, another layer of matrices would be computed so that given a signature,

the model would return the likelihood of it originating from every known NIC.

**Identification stage.** We envision PARADIS sensors as being integrated with wireless access points. This way, every frame in the WLAN will be available to PARADIS. In the simplest scenario, a PARADIS sensor simply demodulates all wireless frames, extracts radiometric signatures and converts them to vectors, which are then sent to a PARADIS server for identification along with the MAC address of the frame. The server may perform identification immediately, or delay identification until it accumulates multiple signatures known to be from the same transmitter. We will assume that identification is always performed on groups, or bins, that could have one or more signatures. To guard against impostors injecting frames with forged MAC address into the classification bins, the server may choose to discard outliers before running classification on the bin. See description of PARADIS-kNN earlier for an example approach to outlier detection using dissimilarity measures.

**PARADIS-kNN:** Every incoming signature in a bin is compared to all the signatures learned during training by using the $l_1$ function and best match is identified based on the similarity values. For example, the similarity between signatures $\mathcal{A}_1$ and $\mathcal{A}_2$ is 0.8, and between $\mathcal{A}_3$ and $\mathcal{A}_4$ is 1.2. The return value is the identity that appears most frequently among the computed best matches. In case of a tie, the identity with the greatest cumulative similarity is chosen.

**PARADIS-SVM:** Every incoming signature in a bin is processed through the underlying SVM classifier, which can be thought of as a series of matrix multiplications. The model returns the best-matching identity, along with a measure of similarity of the match. The return value is the identity that appears most frequently among the computed best matches. In case of a tie, the identity with the greatest cumulative similarity is chosen.

Once the classifier matches the frames to a known NIC, its MAC address is compared to frames' source address to establish whether it is genuine. Alternatively, the classifier may conclude that the frame did not come from a known NIC, for example, if the similarity of the best match is poor. This information is then passed on to a security application to make a policy-dependent action, such as raising alarm or instructing access points to reject impostor's frames.

## 4. DISCUSSION

There are two issues related to the design of PARADIS that warrant further discussion. In this section, we, first, discuss the cost advantage that PARADIS enjoys over potential attackers that can form the long term basis for its security properties. We then, briefly mention the privacy implications of radiometric identification, in general, and of PARADIS, in particular.

### 4.1 Security foundations

Effectiveness of any physical security system is a function of the cost gap between defensive and offensive systems. In case of PARADIS, its effectiveness and success will be based on the cost gap between monitoring infrastructure and the attacker's cost to circumvent it. We present an intuitive explanation why a system such as PARADIS is always likely to have a cost advantage over attackers.

In order for an attacker to steal victim's identity, the attacker NIC needs to *precisely* learn the signature of a victim, and then *accurately* mimic this signature in its transmissions. It implies that the attacker's transceiver needs to have two properties: (i) the receive path of the attacker needs to be precise in estimating the victim's signature, and (ii) the transmit path of the attacker needs to be accurate in mimicking the victim. In contrast, the PARADIS sensors need to be precise in its receive path alone. Since the cost of a manufacturing process grows with accuracy and precision needed, the cost advantage above, clearly, lies with PARADIS. We believe this cost advantage is fundamental and can be exploited by PARADIS to provide high quality security.

### 4.2 Privacy

A system such as PARADIS can naturally increase privacy risks. Since many wireless devices, such as laptops, cellphones with integrated 802.11 interfaces, etc., are uniquely associated with a single individual, the ability to identify and correlate transmissions from such a device will compromise user privacy. In fact, we believe that the high-level of accuracy achieved by PARADIS make such privacy concerns even more important. Understanding the privacy implications and attempting to mitigate them will be an important direction of future work.

## 5. EXPERIMENTAL RESULTS

We implemented and evaluated PARADIS at the ORBIT indoor wireless testbed facility [26]. ORBIT's main wireless network testbed consists of 400 nodes with two 802.11 wireless interfaces each. The nodes are suspended from the ceiling in a room with area of a little over 400 square meters with three narrow pillars supporting the ceiling. The nodes' external "rubber duck" antennas form a square grid where adjacent antennas are separated by 1 meter. We collected data from a subset of nodes, all of which use identical Atheros NICs based on Atheros AR5212 baseband processor the AR2111 RF front-end.

In our evaluation we have chosen to focus on NICs of the same model. We believe that this is the most challenging classification scenario since all the NICs are likely to have been made of the same components at the same facility. Indeed, many of our NICs had consecutive serial numbers. Conversely, we expect that NICs of different brands, or even different models, will be easier to distinguish since their design, and perhaps component specifications are different. Although it is possible that transmitters by other manufacturers could be more resistant to classification, we note that all the data we used was compliant with 802.11 standard's accuracy requirements.

### 5.1 Data collection process

The data collection process was as follows. The data for this work was collected over about one week in the August of 2007 and another week in January of 2008. Exact collection set-up varied somewhat between the collection sessions. The variations were dictated by practical concerns, such as availability of hardware, and having to move our equipment. The available ORBIT nodes with Atheros NICs were configured as 802.11b access points on channel 1. We used Agilent 89641S vector signal analyzer (VSA) [34] as the PARADIS

sensor to capture the wireless frames sent out by the different nodes and to extract the modulation metrics of interest.

Over the collection period, the VSA used a 6 dBi omnidirectional antenna, 8 dBi patch antenna, and, for a few days, an 18 dB low-noise amplifier. Antenna orientation and location also changed by a few meters between sessions but maintaining line-of-sight with all the nodes. All the data was collected from nodes between 5 and 25 meters away from VSA's antenna. RF noise conditions fluctuated as well, depending on the level of activity of other wireless networks in the vicinity.

Success of the identification process across such changes indicate robustness of PARADIS to variations in channel characteristics.

## 5.2 Evaluation Methodology

Effective evaluation of overall classifier performance is not a trivial task. In part the difficulty has to do with the many application-dependent factors affecting performance. For example, some applications may have little training data available, while other require low computational cost.

Below are the definitions and discussion of the prominent concepts employed in evaluation of PARADIS that, we hope, will show that our results are unbiased and generalizable. We first discuss our metrics of performance, and then other methodology details of this evaluation.

**Average error rate.** Average error rate is the same as the average misclassification rate over the entire dataset. It is the ratio of all misclassified samples to the total number of samples in the dataset.

We will only use average error rate as a rough measure of overall classifier performance. One of the reasons we do not use this metric more is because it does not differentiate between false positives and false negatives. In identification systems, however, depending on application, one may be far more important than the other.

**False accept rate (FAR).** FAR is also known as the false positive rate (false accept rate is a term more common in the domain of biometrics). We define it for a given NIC as the ratio of false positives to the number of negative examples.

Although FAR is a useful metric in some cases, in a radiometric identification system, a strong focus on this metric can be misleading. This is because in a typical dataset with samples from many NICs, for a given NIC, the negative examples consist of the data of all the other NICs, which is likely to be relatively large, and cause FAR to be very low even for mediocre classifiers. For both the SVM and kNN schemes, this metric was negligibly small and will not be reported.

**Worst-case similarity.** One of the important performance measures that error rates do not capture is how uniformly the misclassifications are distributed across the population. In particular, it is important that no NIC is consistently misidentified as another one, since otherwise that NIC would be able to masquerade as someone else. To quantify this aspect of performance we use the measure of a NIC's worst-case similarity in the following way. Given a NIC, the victim, we find among the other NICs the one with the greatest fraction of frames misclassified as the victim. This is the worst-case, or most dangerous impostor, and the fraction of

its frames misclassified as the victim is the worst-case similarity of the victim. For example, suppose a NIC has worst-case similarity of 0.5, it means there is another transmitter in the population half of whose frames are misclassified as coming from the NIC.

**False reject rate (FRR).** False reject rate is also a term commonly used in biometrics and is equivalent to false negative rate. We calculate FRR on per-NIC basis as the number of false negatives divided by the total number of samples from the NIC. FRR estimates the likelihood of a NIC's frame being incorrectly identified as someone else's.

We choose to pay a lot of attention to this metric and this choice requires justification, as it may be seen as unorthodox. Consider the example application of detection of compromised keys. For PARADIS to be truly useful for this application, it has to have two properties: reliable detection of identity theft, and keeping false alarms to an absolute minimum, especially since actual security compromises are expected to be rare. The first property is addressed by worst-case similarity, the second one by false reject rate, as we will explain next.

Note that there is a subtle issue with terminology. A false positive, or false alarm, from the point of view of an administrator happens when PARADIS rejects identity of a legitimate user, that is, a false negative from the point of view of PARADIS. Therefore, in our case false reject rate corresponds to the likelihood of PARADIS wasting administrator's time, and must be minimized.

**$k$-fold cross-validation.** $k$-fold cross-validation is a technique to reduce random performance artifacts by evaluating and averaging performance over multiple subsets of the dataset. A common way to implement $k$-fold cross validation is to divide the dataset into $k$ disjoint subsets, then evaluate the model $k$ times each time using different subset as training set, and all remaining subsets as the testing set. This way the testing set is always $k - 1$ times larger than the training set.

There are other ways to validate performance. One of the reasons we chose $k$-fold cross-validation is that it restricts the size of the training set, thus ensuring that the model does not simply memorize the data, or overlearn.

Overlearning describes the phenomenon when a model's performance is specific to the dataset. In other words, it happens when a model in effect memorizes a specific dataset, and its performance does not generalize to other datasets. While overlearning is a complex issue, the easiest way to overlearn is by making the training set too large.

If training and testing sets are obtained using $k$-fold cross-validation, the rule of thumb is to use between 3 and 10 folds [8], resulting in testing set being at least twice and at most nine times as large as the training set. In our evaluation of PARADIS was done using 5-fold cross-validation, the resulting test sets were four times as large as the corresponding training sets.

**Eliminating invalid frames.** Prior to evaluation we removed invalid frames from the dataset. Specifically, the frames that were correctly decoded by the VSA, but actually were distorted to such an extent that a standards-compliant 802.11 receiver should have rejected them. The VSA, being a more sophisticated device than a typical NIC was able to demodulate those frames, but since such frames were in fact invalid, they will be rejected by a 802.11-compliant re-

ceiver. Hence, we decided to ignore all such frames in our identification processes.

Therefore, any data that did not correspond to a valid 802.11 frame was discarded. Specifically, in order to be included in our dataset a frame had meet the following three requirements. (i) The frames had to be a beacon frame with a correct 802.11 checksum. (ii) Frame's frequency error was within ±25 ppm of the channel's center frequency (± 60.3 kHz for channel 1 at 2.412 GHz), as required by Section 18.4.7.4 on transmit center frequency tolerance of the 802.11b standard [20]. (iii) Frame's error vector magnitude values were below 0.35, as required by the Section 18.4.7.8 on transmit modulation accuracy of the 802.11b standard [20]. About 4% of the collected data was discarded using these criteria mostly due to failed checksums.

## 5.3 Evaluation

We evaluated PARADIS by first establishing optimal model parameters and then examining the overall performance. In our presentation, however, we will reverse these steps lest the overall message gets lost in model tuning data. We will first present the overall performance of optimal PARADIS models, and then justify the parameters we chose.

For the overall performance results, we have used training set size of 20 frames, and groups of four frames as classifier input. The optimal features set were found to be, in order of positive effect on performance: for PARADIS-SVM: frequency error, SYNC correlation, I/Q offset, magnitude and phase errors; for PARADIS-kNN: frequency error, SYNC correlation, I/Q offset. The overall performance evaluation dataset consisted of data from 138 NICs.

### 5.3.1 Overall performance

We first present the overall performance of PARADIS in the context of related approaches from recent literature. A direct comparison of performance between different systems is difficult due to differences in the evaluation methodology and even design goals. To make the comparison more meaningful, we chose to evaluate PARADIS on datasets altered to emulate complexity of the datasets in the other studies.

We modeled dataset complexity and evaluation environment along the following three dimensions: (i) population size, or the number of NICs used in the evaluation; (ii) bin size, or the number of measurements (frames) the model required to render a decision; (iii) training fraction, or the fraction of the dataset used for training. We then adjusted our dataset to match the values of these parameters used in other works, evaluated PARADIS on the altered dataset, and compared its performance to that reported in the corresponding related work.

The results in Table 3 compare the performance of PARADIS to studies of 802.11 device driver fingerprinting by Franklin et al. [13], and the work of Hall et. al. [18]. Note that Hall et. al. employ signal transients in the waveform domain for device identification, while PARADIS uses mod-

ulation domain analysis. We compared the performance of PARADIS to these approaches by emulating the conditions reported in the original studies of these approaches. We observed that under comparable conditions, error rate of PARADIS was close to 0, compared to 15% and 8% of the original studies.

These numbers cast our scheme in a positive light, however, we realize that there are many aspects of classifier performance and they cannot be completely described by a single number, nor can be captured without a head-to-head comparison. Nevertheless, we believe that Table 3 is strong evidence that PARADIS would outperform other leading approaches to device identification in a head-to-head comparison on the same dataset.

*In a much larger population of 138 NICs used in our evaluation, our best scheme, PARADIS-SVM had an error rate of 0.34%, while PARADIS-kNN had an error rate of about 3%.*

### 5.3.2 Details of large scale evaluation

In this section we turn out attention to the performance of individual NICs. When using PARADIS-SVM, very few NICs experienced any misclassification. False reject rate accounted for virtually entire misclassification rate.

Figure 6(a) shows that there was one NIC with a 10% false reject rate, and just 16 other NICs with a non-zero false reject rate, 11 of which were under 2%.

Similarly, Figure 6(b) shows that among the 17 NICs that were victims of impersonation, albeit unintentional, one had a 17% worst-case similarity, and 14 NICs were under 1%.

Two NICs did relatively poorly with respect to both metrics: one NIC had FRR of 7% and similarity of 4%, another had FRR of 5% and similarity of 1%. Otherwise there did not seem to be a correlation between a NIC's FRR and worst-case similarity.

We can make several observations based on this data. First, *no NIC was able to masquerade as another.* The least fortunate NIC had similarity of 17%, meaning that only about 1 out 6 frames sent by the perpetrator were mistaken as coming from the victim, making it unlikely to avoid detection. At the same time, similarity was under 5% for 99% of the population, most of which was not affected at all. Another observation is that fewer than 5% of the 138 transmitters in our dataset accounted for 75% of all misclassifications.

We now consider performance of PARADIS-kNN. The average false reject rate for PARADIS-kNN was 3%. However, Figure 7(a) shows that almost two thirds of the population did not experience any false rejects, while false reject rate of just a few NICs towers that of the majority. Interestingly, these are not the same NICs that caused trouble with PARADIS-SVM, raising the possibility of combining the kNN and SVM approaches to reduce the number of inherently difficult to classify NICs. Still, a little over 10% of the population experienced false reject rate of 10% or more, and about 20% of the population had false reject rate above 5%. However, 80% of the population still only saw 5% or lower false reject rate.

Figure 7(b) shows the per-NIC similarity distribution observed under PARADIS-kNN. As with figure on false reject rate, the distribution is skewed, but to a lesser degree.

---

[1]Reported error rates correspond to scenarios that could be emulated best with our dataset. Please refer to the original papers for details.

[2]Designed for transmitter distinction based on driver software; average accuracy over all test sets.

[3]Transient-based 802.11 transmitter distinction. The error rate includes classification failures and failures to detect signatures.

| Approach | NIC population size | Bin size | Training fraction | Reported error rate[1] [2] | Equivalent performance of | |
| --- | --- | --- | --- | --- | --- | --- |
| | | | | | PARADIS-kNN | PARADIS-SVM |
| Franklin et. al. [13][2] | 17 | 8 | 5% | 15% | 0% | 0% |
| Hall et. al. [18][3] | 30 | 10 | 33% | 8% | 0% | 0% |
| PARADIS | **138** | **4** | **20%** | - | **3%** | **0.34%** |

**Table 3: Comparison of PARADIS and other studies**

(a) PARADIS-SVM: per-NIC FRR
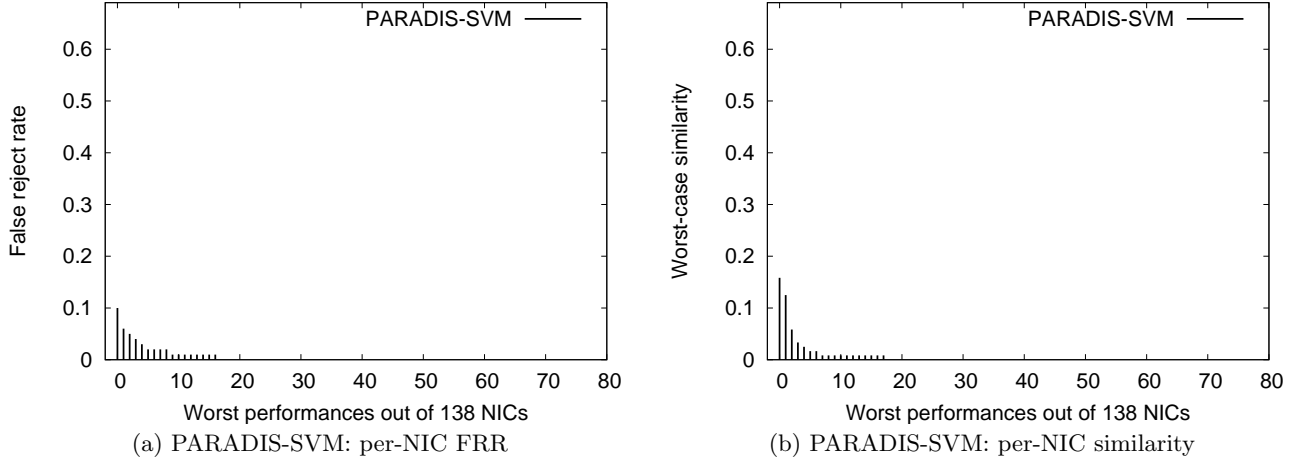
(b) PARADIS-SVM: per-NIC similarity

**Figure 6: PARADIS-SVM: Summary of performance**

About 90% of all NICs had worst case similarity under 10%, meaning that on average fewer that 1 out of 10 frames sent by an impostor would be misclassified. Average similarity was 3%.

**Overall performance summary:** PARADIS-SVM consistently outperforms PARADIS-kNN, whose performance, nevertheless, is sufficient to be useful in certain applications, especially involving constrained computation resource, or implementation in hardware.

To summarize, in evaluation involving 138 identical NICs, PARADIS-SVM had an overall error rate of 0.34%, with only a few NICs having non-zero FRR or worst-case similarity.

Since PARADIS-SVM was implemented using generic SVM library at default parameter settings, this suggests that the classification problem did not stretch the capability of the model to its limits. In other words, PARADIS-SVM may still perform well with noisier data, whether due to interference or less precise hardware.

We now comment on other aspects of our evaluation.

### 5.3.3 Location insensitivity

In order to estimate the effect of the multipath phenomena on performance we collected data from three "travelling nodes" that had the same hardware configuration as the other nodes used in our dataset. The travelling nodes were carted together to three locations from which we recorded their beacons using a stationary VSA. All the locations were inside WINLAB office space that contains 25 cubicles. The three collection locations were chosen randomly, had line-of-sight to the VSA's antenna, and were between 10 and 50 feet away from it. At each of three spots about 100 frames were collected from each node.

| NIC id | PARADIS-SVM | | PARADIS-kNN | |
| --- | --- | --- | --- | --- |
| | similarity | FRR | similarity | FRR |
| 0d6396 | 0 | 0 | 0 | 0 |
| 354534 | $1.6 \times 10^{-5}$ | 0 | $1.6 \times 10^{-5}$ | $7.2 \times 10^{-4}$ |
| 358563 | 0 | 0 | $3.2 \times 10^{-5}$ | $9.0 \times 10^{-5}$ |

**Table 4: Performance of travelling nodes**

The nodes were identified by the last portion of their MAC address: $0d6396, 354534, 358563$. Table 4 lists the misclassifications that involved the travelling nodes when combined with the data of the 138 stationary nodes, averaged across the 5 folds of cross-validation.

The travelling nodes did not appear to present a significant challenge to PARADIS-SVM. Although a further investigation is warranted, data in Table 4 suggests that, as expected, PARADIS signatures are not greatly affected by the location of the transmitter.

### 5.3.4 Parameter tuning

We now discuss the empirically established parameters of the optimal models. Note that performance numbers reported in the following sections are not meant to quantify optimal performance of PARADIS. The tuning models were suboptimal in at least one of their parameters and in some cases required modified dataset to explore extreme cases.

**Training set size.** We evaluated the effect of training set size on effectiveness of PARADIS by performing classification on datasets using varying training set sizes while keeping bin size at 1 frame. In order to maintain constant problem complexity, every evaluation had the same ratio of train to test set sizes, the same number of NICs with the
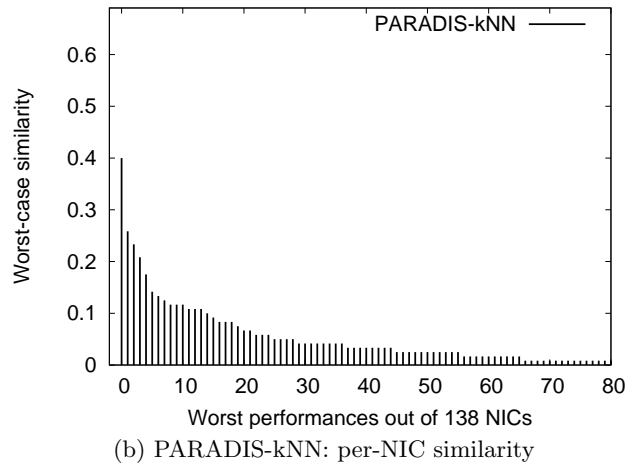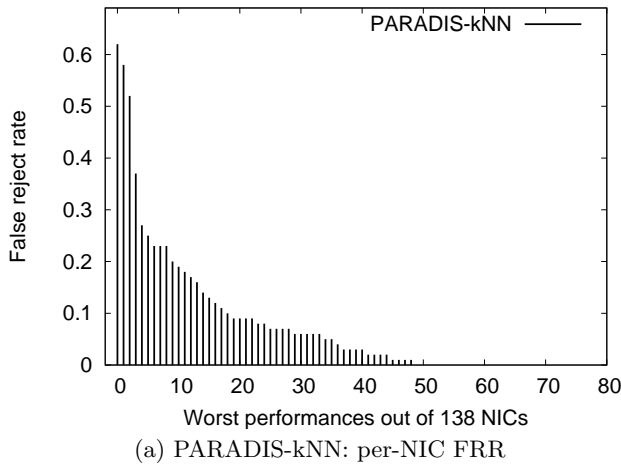
(a) PARADIS-kNN: per-NIC FRR



(b) PARADIS-kNN: per-NIC similarity

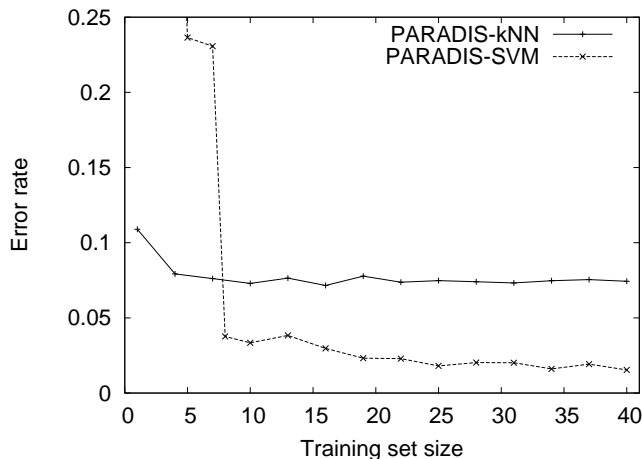**Figure 7: Summary of performance**



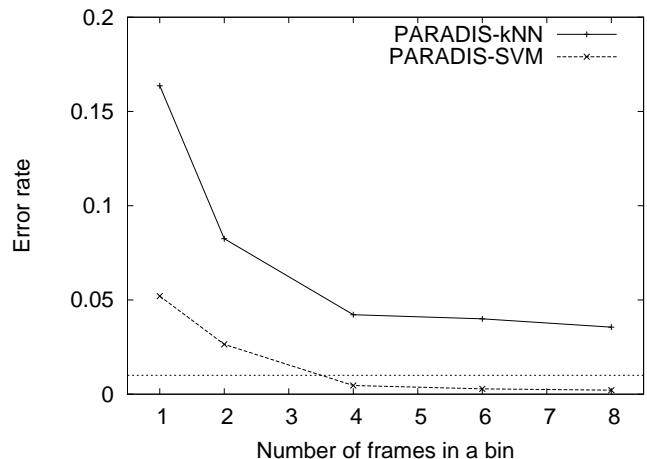**Figure 8: Effect of training set size on accuracy**



**Figure 9: Effect of bin size on accuracy**

same number of points from every NIC. 57 NICs had sufficient number of data points to be used in such evaluation. Note that 138 NICs were used in previous evaluations.

We make two observations on the effect of training set size on accuracy, as shown in Figure 8. First, PARADIS-SVM requires a larger training set size, at least 7 frames, to be effective. This is due to the fact that the SVM training process requires the training set to be split in two subsets for model tuning, thereby increasing training set size requirements.

The more interesting observation is that once effectiveness stabilized, neither implementation of PARADIS appeared to significantly benefit from increased training set size. For examples, accuracy of PARADIS-SVM increases only by about 2.5% between training set sizes of 7 and 40. Therefore, just a few frames capture the radiometric identity of a NIC.

**Tuning bin size.** Relatively few applications require per-frame identification. At the same time, it is often possible to infer which frames originated at the same NIC. For example, 802.11 acknowledgements and retransmissions can conveniently boost the number frames that are highly likely to have come from the same transmitter. In this section

we present the relationship between bin size and accuracy. Training set size was 10 frames, which is smaller than that used for earlier evaluations.

Figure 9 illustrates the effect of binning on average misclassification rate and shows the error rate level of 1% with a dotted line. Beyond bin size of 4 performance of PARADIS-SVM becomes nearly perfect, with error rate falling below half percent. On the other hand, performance of PARADIS-kNN had its best performance of 3% with the bin size of 8.

## 6. CONCLUSION

In this paper we have addressed the fundamental issue of wireless NIC identification at the physical layer. We designed, implemented and evaluated PARADIS, a technique that identifies wireless transmitter devices based on minor artifacts in their emissions that are produced by idiosyncratic hardware properties of individual NICs. Unlike the previous state-of-art techniques, PARADIS defines a signal's signature in terms of structure imposed by the modulation scheme, thus greatly simplifying the problem.

Evaluation of PARADIS involved over 130 identical IEEE

802.11 wireless NICs with two sets of data collected over 5 months apart. Our technique by far outperformed the other state-of-art techniques, and proved to be capable of achieving accuracy in excess of 99%. Further, our evaluation demonstrated that PARADIS is resilient to mobility, varying noise conditions and hardware aging.

Our evaluation of the strengths and weaknesses of PARADIS suggests that it could be especially useful for detection of compromised cryptographic keys, which in general are hard to detect. Further, PARADIS could be used in the future in the context of enforcing regulations on open access or shared spectrum bands. For example, it could be used to collect the evident identify of interfering devices during investigations as evidence against a specific transmitter. Or, the identity of an interfering transmitter could be quickly established by comparing it with a database of registered transmitters in the region.

# 7. REFERENCES

[1] Agilent Technologies. Agilent 802.11a/g Manufacturing Test Application Note : A Guide to Getting Started. *Application note 1308-3.*

[2] Agilent Technologies. Making 802.11g transmitter measurements. *Application note 1380-4.*

[3] Agilent Technologies. RF Testing of WLAN products. *Application note 1380-1.*

[4] Agilent Technologies. Testing and Troubleshooting Digital RF Communications Transmitter Designs. *Application note 1313.*

[5] M. Barbeau, J. Hall, and E. Kranakis. Detecting Impersonation Attacks in Future Wireless and Mobile Networks. *MADNES*, 2006.

[6] David K. Barton and Sergey A. Leonov. *Radar Technology Encyclopedia.* Artech House, 1998. See entry on Target Recongnition and Identification.

[7] B.E. Boser, I.M. Guyon, and V.N. Vapnik. A training algorithm for optimal margin classifiers. *Proceedings of the fifth annual workshop on Computational learning theory*, pages 144–152, 1992.

[8] Carl Rasmussen and Christopher Williams. *Gaussian Processes for Machine Learning.* The MIT Press, 2006.

[9] Chih-Chung Chang and Chih-Jen Lin. *LIBSVM: a library for support vector machines*, 2001. http://www.csie.ntu.edu.tw/~cjlin/libsvm.

[10] H.C. Choe, C.E. Poole, A.M. Yu, and H.H. Szu. Novel identification of intercepted signals from unknown radio transmitters. *SPIE*, 2491:504, 2003.

[11] R.E. Fan, P.H. Chen, and C.J. Lin. Working Set Selection Using Second Order Information for Training Support Vector Machines. *The Journal of Machine Learning Research*, 6:1889–1918, 2005.

[12] D.B. Faria and D.R. Cheriton. Detecting identity-based attacks in wireless networks using signalprints. *ACM WiSe*, pages 43–52, 2006.

[13] J. Franklin, D. McCoy, P. Tabriz, V. Neagoe, J. Van Randwyk, and D. Sicker. Passive Data Link Layer 802.11 Wireless Device Driver Fingerprinting. *Usenix Security Symposium*, 2006.

[14] R. Gerdes, T. Daniels, M. Mina, and S. Russell. Device Identification via Analog Signal Fingerprinting: A Matched Filter Approach. *NDSS*, 2006.

[15] Gordon Lyon. Nmap network mapper. http://nmap.org.

[16] F. Guo and T. Chiueh. Sequence Number-Based MAC Address Spoof Detection. *Recent Advances in Intrusion Detection: 8th International Symposium, RAID 2005*, 2006.

[17] J. Hall, M. Barbeau, and E. Kranakis. Radio frequency fingerprinting for intrusion detection in wirless networks. *Defendable and Secure Computing*, 2005.

[18] Jeyanthi Hall. *Detection of rogue devices in wireless networks.* PhD thesis, 2006.

[19] IEEE Standards Association. IEEE Std 802.11a. http://standards.ieee.org/getieee802/download/802.11a-1999.pdf.

[20] IEEE Standards Association. IEEE Std 802.11b. http://standards.ieee.org/getieee802/download/802.11b-1999.pdf.

[21] T. Kohno, A. Broido, and KC Claffy. Remote physical device fingerprinting. *Dependable and Secure Computing*, 2(2):93–108, 2005.

[22] L.E. Langley. Specific emitter identification (SEI) and classical parameter fusion technology. *WESCON*, 1993.

[23] Z. Li, W. Xu, R. Miller, and W. Trappe. Securing wireless systems via lower layer enforcements. *Proceedings of the 5th ACM workshop on Wireless security*, pages 33–42, 2006.

[24] Tom M. Mitchell. *Machine Learning.* McGraw-Hill, 1997.

[25] Motron Electronics. TxID Transmitter FingerPrinter. http://www.motron.com/TransmitterID.html.

[26] ORBIT wireless research laboratory. WINLAB, Rutgers University. http://www.orbit-lab.org/.

[27] Jeffrey Pang, Ben Greenstein, Ramakrishna Gummadi, Srinivasan Seshan, and David Wetherall. 802.11 user fingerprinting. In *MobiCom '07: Proceedings of the 13th annual ACM international conference on Mobile computing and networking*, pages 99–110, New York, NY, USA, 2007. ACM.

[28] N. Patwari and S.K. Kasera. Robust location distinction using temporal link signatures. *ACM MOBICOM*, pages 111–122, 2007.

[29] Philip Stepanek and William M. Kilpatrick. Modeling Uncertainties in a Measuring Receiver. *Agilent Technologies.*

[30] K.B. Rasmussen and S. Capkun. Implications of Radio Fingerprinting on the Security of Sensor Networks. *Proceedings of IEEE SecureComm*, 2007.

[31] K.A. Remley, C.A. Grosvenor, R.T. Johnk, D.R. Novotny, P.D. Hale, M.D. McKinley, A. Karygiannis, and E. Antonakakis. Electromagnetic Signatures of WLAN Cards and Network Security. *ISSPIT*, 2005.

[32] M.J. Riezenman. Cellular security: better, but foes still lurk. *Spectrum, IEEE*, 2000.

[33] A. Ross and A. Jain. Information fusion in biometrics. *Pattern Recognition Letters*, 24(13):2115–2125, 2003.

[34] 89600S series VXI-based Vector Signal Analyzer. Agilent technologies.

[35] K.I. Talbot, P.R. Duley, and M.H. Hyatt. Specific Emitter Identification and Verification. *Technology Review*, 2003.

[36] P. Tuyls and J. Goseling. Capacity and Examples of Template-Protecting Biometric Authentication Systems. *ECCV*, 2004.

[37] O. Ureten and N. Serinken. Wireless security through RF fingerprinting. *Electrical and Computer Engineering, Canadian Journal of*, 32(1):27–33, 2007.

[38] J. Wright. Detecting Wireless LAN MAC Address Spoofing. *White Paper, January*, 2003.