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 based 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 100 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.

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. However, in a wireless environment, the untethered nature of 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.

In this paper, we study an additional mechanism for 802.11 wireless device identification based on the notion of *radiometric identification* — a technique to establish the physical-layer identity of a transmitter based on its unique set of benign hardware imperfections that are manifested in the emitted signals as NIC-specific radio-frequency (RF) artifacts[1].

These artifacts are present in every transmitted frame and serve as signatures of the specific transmitter and can be used to establish transmitter identity when compared against a set of pre-recorded signatures.

Radiometric identification vs RF fingerprinting.

We use the term radiometric identification instead of the more commonly adopted term, radio-frequency fingerprinting because the latter has a wider meaning and usage than our specific application. The term RF fingerprinting, in general, refers to the process of classifying transmissions based on observed features of an RF signal. We can broadly classify RF features of a signal into (i) channel-specific features: that characterize the properties of the wireless channel and

*This research was performed under an appointment to the Department of Homeland Security (DHS) Scholarship and Fellowship Program, administered by the Oak Ridge Institute for Science and Education (ORISE) through an interagency agreement between the U.S. Department of Energy (DOE) and DHS. ORISE is managed by Oak Ridge Associated Universities (ORAU) under DOE contract number DE-AC05-06OR23100. All opinions expressed in this paper are the author's and do not necessarily reflect the policies and views of DHS, DOE, or ORAU/ORISE.

†This study was supported in part by the US National Science Foundation under grants CNS-0639434, CNS-0627589, CNS-0627102, CNS-0520152, and CNS-0747177.

‡This study was supported in part by the US National Science Foundation under grant CNS-0524475.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
MobiCom'08, September 14–19, 2008, San Francisco, California, USA.
Copyright 2008 ACM 978-1-60558-096-8/08/09 ...$5.00.

[1]We adapt the term radiometric identification from its biological analog, biometric identification — since the goal of both is to recognize individuals based on manifestations of intrinsic traits.

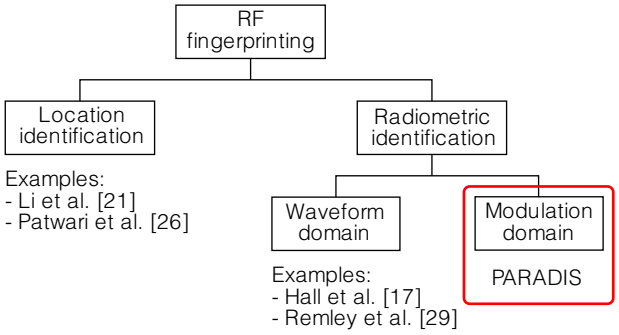

Figure 1: Radiometric identification and PARADIS

environment, and (ii) transmitter-specific features that characterize the wireless transmitter itself, and are independent of the channel between the transmitter and the receiver.

Radiometric identification requires us to ignore channel-specific features, such as channel impulse response, and to utilize only transmitter-specific features that help uniquely identify the specific wireless transmitter. Channel-specific features are used to uniquely identify the channel between the transmitter and the receiver and have been successfully adopted in robust location distinction (see work by Li, Miller and Trappe [21] as well as Patwari and Kasera [26]). Radiometric identification is, therefore, a special instance of RF fingerprinting.

1.1 Basics of radiometric identification

Radiometric identification works because of existence of benign hardware imperfections within transmitters, also known as transmitter impairments. These impairments are typically acquired at manufacture and assembly of analog components of transmitter RF-front-ends. We use the term "benign" because these impairments do not have an adverse impact on communications and fall within the quality tolerances specified by the 802.11 standards. Virtually all analog components on the NIC's transmit path, from interconnects to antennas, contribute to deviation in emitted signal with respect to the ideal. Some significant sources of these deviations are presented Figure 2. More information on this subject can be found in [1, 2, 3, 4, 27].

In a sense, despite superficial sameness of network interface cards 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 will greatly increase device cost. In fact, common technology standards, including 802.11, explicitly require different NICs to tolerate rather wide ranges of RF variations in received signals for seamless inter-operation. Therefore, the RF artifacts of the transmitter-specific hardware impairments can be used to establish the identity of the corresponding transmitter.

1.2 Our proposed approach: PARADIS

At the core of any radiometric identification technique are the specific artifacts that are used to distinguish transmitters. The optimal artifacts for the differentiation process depend on many factors, including hardware and transmitter design, and the underlying communication technology.

The notion of radiometric identification itself is not new. We found references to some of the earliest efforts in this direction dating back to the Vietnam War era, with military aircrafts attempting to distinguish between friendly and enemy radars. Since then, similar systems appear to have been deployed by cellular networks to establish authenticity of cellular transmitters in order to prevent fraud. Due to the commercial and military nature of such systems, only very limited implementation details are publicly available. Nevertheless, there is strong evidence that they use transient signal characteristics in the waveform domain for their identification goals [23].

Similar studies of radiometric identification for 802.11 transmitters in were carried out by Jeyanthi Hall, Michel Barbeau and others [5, 15, 16]. However, its performance, as well as fundamental difficulties associated with transient-based identification, led us to explore a new approach, one that utilizes protocol-specific properties of the PHY layer of the 802.11 standards, and operates in the modulation domain.

Our system, called Passive RAdiometric Device Identification System, or PARADIS, is uniquely suited to distinguish between 802.11 NICs and can achieve significantly improved identification accuracy when compared to schemes operating over transient signal characteristics. PARADIS uses a large feature space consisting of five distinct features from the modulation domain, namely, frequency error, magnitude error, phase error, I/Q offset, and SYNC correlation of the corresponding wireless frame.

We believe that PARADIS is the first radiometric identification system designed for 802.11 transmitters that is both (i) completely implemented and (ii) has been shown to accurately *distinguish between more than a hundred, identically manufactured, 802.11 NICs.*

1.3 PARADIS application and overview

PARADIS is a technique to differentiate between transmissions of different 802.11 NICsbased on physical layer information. The main application of this technique is, naturally, in wireless security. In our experiments, we show that PARADIS provides identification accuracy of over 99%. Misclassifications are rare but do occur largely due to RF interference. Even though this is far better than results reported in prior literature, we believe that no mechanism that relies of the physical layer information for identification can provide 100% accuracy, and so care needs to be taken how such a scheme is used in a security solution.

Skipping ahead, our evaluation revealed that PARADIS should not be used as a drop-in replacement for just any access control mechanism specific to wireless networks. For example, using PARADIS for per-frame transmitter identification is likely to lead to performance degradation of a small but not insignificant fraction of transmitters. In its current state, PARADIS may provide reliable per-frame identification in certain environments, but is not suitable for every WLAN.

Instead, PARADIS should be used a secondary security perimeter that detects breaches in the primary security perimeter, established using the usual cryptographic mechanisms. We explain this with an example, next.

PARADIS as a secondary security perimeter.

WLANs today perform access control for users in different ways. A simple technique for WLAN access control is 802.11

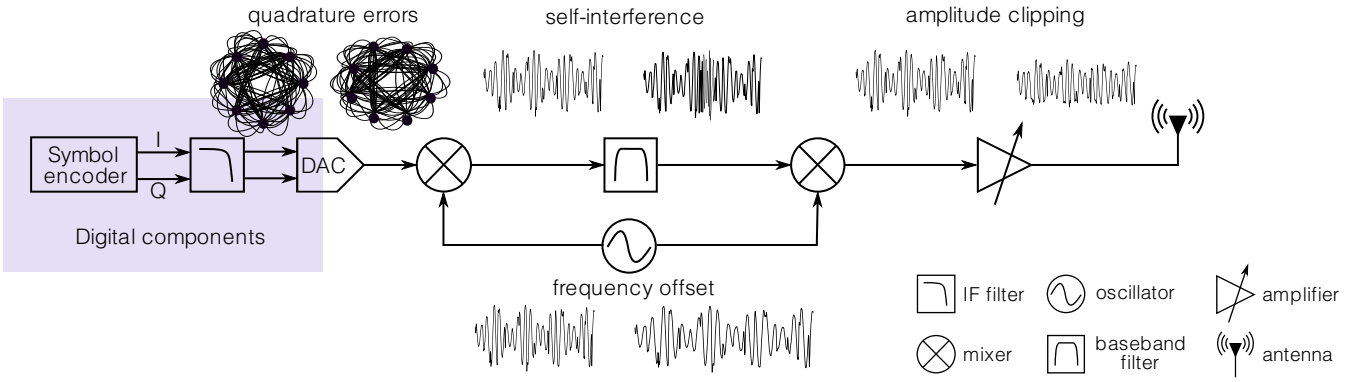

Figure 2: Common transmitter impairments and their sources

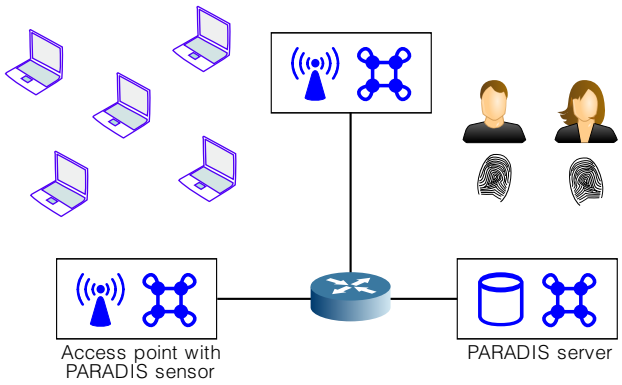

Figure 3: PARADIS schematic

MAC address based filtering, whereby traffic from only authorized NICs is permitted into the network, based on their MAC addresses. However, given the ease of spoofing MAC addresses, such access control mechanisms do not provide any real security. A more security conscious WLAN performs access control through use of data encryption using *secure key material* exchanged between an authentication server and the supplicant (wireless client). Appropriate access control in this scenario depends completely on the secrecy of the key material. If the WLAN administrators realize that the key material is compromised, then they can issue new key material for its users. However, detecting when certain key material is compromised is one of the biggest remaining security challenges. The role of PARADIS is in addressing this challenge, which we explain next through an example of an enterprise WLAN, shown in Figure 3.

As shown in the figure, PARADIS consists of a server, an associated radiometric fingerprint database, and a set of wireless sensors, that can demodulate 802.11 transmissions. In our current implementation, we use a vector signal analyzer as the sensor. Each time a new NIC is authorized into the network, the administrator builds its radiometric fingerprint, by analyzing a small set of 802.11 frames transmitted by this NIC, and stores it in the database. The device using this NIC is also configured with appropriate secret key material in the regular process.

Let us, now consider an unauthorized NIC (*Impostor*) has somehow acquired the secret key material of an authorized NIC (*Victim*) and is, therefore, able to authenticate itself with the network. When the *Impostor* sends one or more 802.11 wireless frames using these credentials, they will be captured by the PARADIS sensors (like all other transmitted frames) and will be sent to the server for analysis. If the radiometric fingerprint of these frames do not match any valid radiometric fingerprint in the database, the server will alert the administrators of possible compromise of such secure credentials. The administrators can, then revoke specific keys, and create new ones.

Clearly, such a security application requires high accuracy in radiometric identification, but can tolerate the occasional errors made by PARADIS. In addition, due to the defensive nature of access control, it is important to ensure that PARADIS errs on the side of caution, i.e., always generates alarms on suspicion of detecting an unauthorized NIC, while trying to minimize instances when it fails to identify some authorized NIC. Another important consideration for such a system is the existence of systematic errors. For example, despite of the low overall mis-classification rate, if an individual unauthorized NIC can almost always successfully masquerade as another individual authorized NIC, then PARADIS will fail to detect such occurrences. We will comment on these performance properties of PARADIS, in Section 5.

1.4 Feature highlight

The following are some of the salient features of PARADIS that make it particularly attractive for our proposed security application:

- *Simplicity:* One of the main advantages of PARADIS is its use of RF signal features in the modulation domain. These features are explicitly defined within the 802.11 communication standard itself, and also allows for compact representation and efficient comparison. Having examined a range of alternative features, we picked three features to serve as the radiometric identity of an 802.11 transmitter, namely the frequency error, I/Q offset, SYNC correlation, phase error, and magnitude error of the corresponding wireless frame.
- *Consistency:* In order to serve the radiometric identification purpose, the feature set need to be consistent across noisy wireless environments, multi-path effects, and diverse channel characteristics. We show experimentally that this is true for our set of chosen features.
- *Unforgeable:* The features used for radiometric identifica-

tion are based on minor hardware impairments. The cost and effort of creating a new NIC to mimic the hardware impairments of an existing NIC is likely to make forgery difficult.
- *Unescapable:* 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.

Other key contribution: Apart from the above contributions in the design of an effective radiometric identification system, our other major contribution is in the systematic evaluation of this scheme is a relatively large scale. We have completely implemented PARADIS and have demonstrated its capability on the ORBIT testbed at Rutgers University. In particular, we have showed that PARADIS is able to accurately distinguish more than the hundred, identically manufactured, 802.11 NICs that were used in the experiment. The environment of the ORBIT testbed itself was a quite challenging one, because there was significant interference and noise within the testbed itself from different wireless nodes. This is, most likely, the largest scale successful demonstration of accurate radiometric fingerprinting using 802.11 NICs.

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 [33]. 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 [20, 30, 23].

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. [33]. Similar issues are also explored in the field of biometrics [31, 34].

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 place in the toolkit of network administrators [13]. Typically such tools are used to identify computers running vulnerable software.

In the context of 802.11 devices, Wright [36], Guo et al. [14] 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. [11] 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. [25] 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. [19] proposed to fingerprint networked devices based on estimation of its clock skew using TCP and ICMP time stamps. Although, unlike the previous approaches, Kohno measured properties of the hardware, since measurements are based on what software reports, users still can interfere with identification by manipulating the time stamps or disabling them.

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. 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. [10] evaluated use of signal strength to model channel response. Patwari et al. [26] evaluated an approach that compared multipath-related properties of signals. Finally, Li et al. [21] leveraged ideas of location distinction and use of probing in the context of sensor authentication.

Note that location distinction approaches cannot recognize a transmitter that moved or remained inactive even for a short period of time.

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 the most intuitive and straightforward way of radiometric analysis was studied by Remley [29], who visually identified differences between signals of different 802.11 transmitters. Gerdes et al. [12] used physical layer signal processing to distinguished between wired Ethernet NICs.

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, 9, 15, 16, 28, 35]. 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 2.

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.

Technique	Type	Goal	Identity Model	Evaluation scale
Kohno et al. [19]	Software meas.	hardware id	clock skew variation	n/a
Franklin et al. [11]	Software meas.	device driver id	compliance with 802.11 standard	17 802.11 NICs
Faria et al. [10]	RF fingerprinting	location distinction	signal power attenuation	135 locations
Patwari et al. [26]	RF fingerprinting	location distinction	multipath channel response	44 locations
Hall [16]	RF fingerprinting	radiometric id	transient properties	30 802.11 NICs
Gerges et al. [12]	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

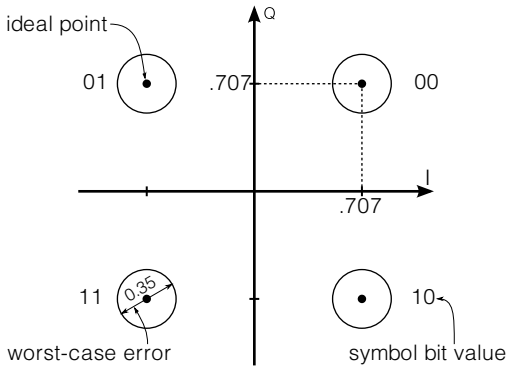

Figure 4: The 4 symbols of QPSK on I/Q plane

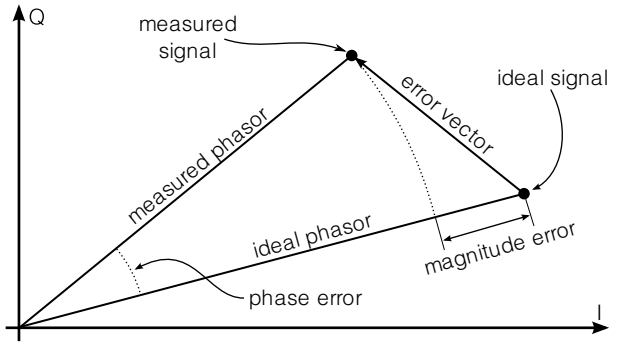

Figure 5: Modulation errors

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 [18], 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 finite 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.

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:* 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 [17]. 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.

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.

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 [26, 21] 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.

So far we have kept the notion of radiometric identity abstract. We will now give it some concreteness using the following example that we will refer back to in later sections. Suppose we have a NIC call it A. Whenever A transmits a frame, a PARADIS sensor, in our case a vector signal analyzer (VSA), computes its radiometric signature and makes it available for the identification software. For the sake of brevity, suppose a signature is defined using only frame's phase error, magnitude error and frequency error, all averaged over the entire frame. The VSA could, for example, demodulate A's frame and report that its phase error is 10 degrees, magnitude error is 0.3 of 802.11 I/Q unit, and frequency error is 40 kHz with respect to channel center. In practice, it is advantageous to normalize data before giving it to a classification algorithm. Therefore, in our implementation a signature of a frame actually was a 6-tuple of real values between 0 and 1, each corresponding to one of the metrics we used.

3.3 Classification algorithms

We have considered two rather different radiometric classifiers, one using support vector machines algorithm (SVM) [8] and the other using the k-nearest-neighbor (kNN) [22] 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 satisfactory, while 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.

Our classifiers used standard, well known kNN and SVM implementations. Therefore, rather than give technical implementation details, we will describe their simplified operation using the example from the previous section.

We will begin with teaching the PARADIS-kNN classifier the radiometric identity of NIC A. Suppose training set of A is made up of four frames with the following normalized signatures: $(.8, .1, .8), (.9, 0, .9), (.7, .2, .8), (.9, 0, .9)$. PARADIS-kNN discards half of the signatures as outliers, second and fourth in this case, and memorizes the per-component average of the rest. The resulting signature, in this case $(.75, .15, .8)$ is the model of NIC A.

Now, suppose PARADIS-kNN needs to identify the sender of a frame with signature $(.8, .2, .8)$. PARADIS-kNN computes a measure of similarity between the incoming signature and signatures of all the known NICs. Without loss of generality, say we actually consider dissimilarity, measured as the sum of absolute values of component-wise differences between signatures. In our example, dissimilarity between $(.8, .2, .8)$ and the stored signature of A, $(.75, .15, .8)$, is low, 0.1. If there is NIC B, whose stored signature is, say $(.2, .2, .5)$, it will have higher dissimilarity, 0.9. Assuming A is indeed least dissimilar, the incoming frame will be classified as having come from A.

Further, suppose that classification is done a group of three frames from the same transmitter that needs to be identified. PARADIS-kNN determines best match for all of the frames, and returns the most frequent identity.

PARADIS-SVM classifier was implemented using LIBSVM [8]. Once again, since PARADIS-SVM used typical SVM-based

classifier implementation, we will describe its operation using our running example.

Unlike PARADIS-kNN it is difficult to give intuition on operation of SVM-based classifiers without getting into SVM theory. During the training phase PARADIS-SVM computes a model based on the training data from the available NIC. To keep things simple, this model can be thought of as a number of real matrices.

At a very high level, signature classification can be though of as a series of matrix multiplications of the signature (which essentially is a real vector) and the internal matrices of the SVM model. The result of this multiplication is a measure of likelihood that the signature came from the NIC whose model was used in classification, making it possible to find the most likely source of the signature. Classification on groups of frames was performed in a manner similar to PARADIS-kNN.

Note that although PARADIS-SVM for convenience used a single monolithic model whose output was the most likely identity, the underlying support vector machine actually computed the likelihood of a match for all possible outcomes, or NIC identities. Therefore, conceptually, this model is equivalent to using a separate support vector machines for every known NIC and picking the best match as output, which is how a practical implementation would likely operate.

3.4 Sequence of operations

The operations of PARADIS can be partitioned into a training phase (for building radiometric signatures) and the on-line phase (for identifying unauthorized intruder NICs). We explain both in turn, briefly.

Learning phase: Before a new NIC in issued to a user for the first time, the administrator collects has it transmit a number of frames to make a training sample. PARADIS-converts those frames into a model for that NIC and stores it. We found training sets of 20 frames to be adequate. The training operation needs to be done once for every NIC.

On-line phase: A PARADIS sensor captures a frame from some NIC. It extracts the modulation domain metrics relevant to the classification algorithm, as well as the software-level identifier (such as, MAC address) and sends this information to the PARADIS server. In practice, we send the real tuple or normalized error metrics and the MAC address, as discussed in the earlier example.

If the PARADIS server performs classification on each wireless frame, it immediately determines the identity of the frame using the signature database. If the server performs classification on binned data, i.e. a small group of frames together, it waits until it collects the requisite number of frames from this transmitter (based on the software identifier), and performs the classification process by serving the corresponding bin of signatures as an input to the classification process. The classifier returns the best matching NIC from the database, or informs the administrators of the fact that no match has been found, or that the MAC address in the frame did not match the established identity.

To guard against impostors injecting frames with forged MAC address into the classification bins, the classifier may discard outliers before running classification on the bin. See description of PARADIS-kNN above for an example approach outlier detection using dissimilarity measures.

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 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 using a vector signal analyzer at the ORBIT indoor wireless testbed facility [24]. 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 the other nodes, all of which use identical Atheros NICs based on Atheros AR5212 baseband processor the AR2111 RF front-end.

Evaluation using NICs of the same model and manufacturer had its advantages and disadvantages. On the positive side, population homogeneity meant that most if not all of the NICs were likely to have been manufactured using the same techniques, time and facility, since many had consecutive MAC addresses. Therefore, homogeneous population

can be expected to be more difficult to classify, or worst-case scenario, compared to a population of NICs made using different processes by different manufacturers. 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) [32] 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, 8dBi 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 number of negative samples is likely to be very large, and cause FAR to be very low even for mediocre classifiers. It turns out that for both the SVM and kNN schemes, this metric was negligibly small.

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, if 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 explain next.

Note 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 effectively 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

Scheme	NIC population size	Bin size	Training fraction	Reported error rate	Equivalent performance of PARADIS-kNN	Equivalent performance of PARADIS-SVM
Franklin et. al. [11] [2]	17	8	5%	15%	0%	0%
Hall et. al. [16] [3]	30	10	33%	8%	0%	0%
PARADIS	**138**	**4**	**20%**	-	**3%**	**0.34%**

Table 2: Comparison of PARADIS and other studies. Reported error rates correspond to scenarios that could be emulated best with our dataset. Please refer to the original papers for details.

folds [7], resulting in testing set being at least twice and at most nine times as large as the training set. In evaluation of PARADIS we used 5-fold cross-validation, resulting in the test set being four times as large as the training set.

Eliminating invalid frames. Prior to evaluation we removed the invalid frames, i.e., 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 were not useful to us.

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 [18]. (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 [18]. 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 evaluate PARADIS on a dataset 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 2 compare PARADIS to works of Franklin et. al. [11], which employs 802.11 device driver fingerprinting for device identification, and Hall et. al. [16] that employ signal transients in the waveform domain for device identification. Under the simulated environments 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 2 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 full-blown evaluation, our best scheme, PARADIS-SVM had an error rate of 0.0034%, 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 of PARADIS-SVM.

Figure 6 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 7 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.

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

[3] transient-based transmitter distinction; data for 802.11 transmitters; stated error rate includes both classification failures and failures to extract signatures.

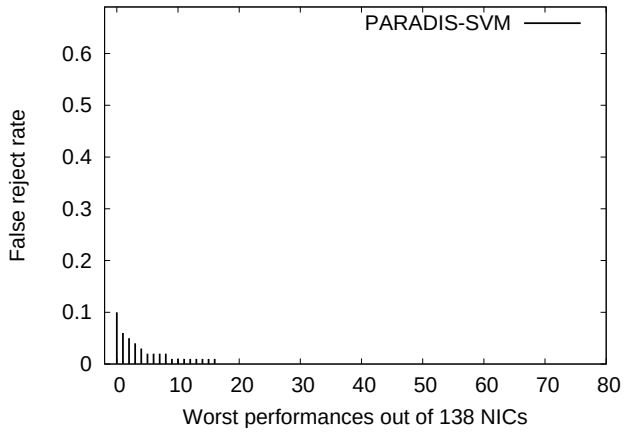

Figure 6: PARADIS-SVM: per-NIC FRR

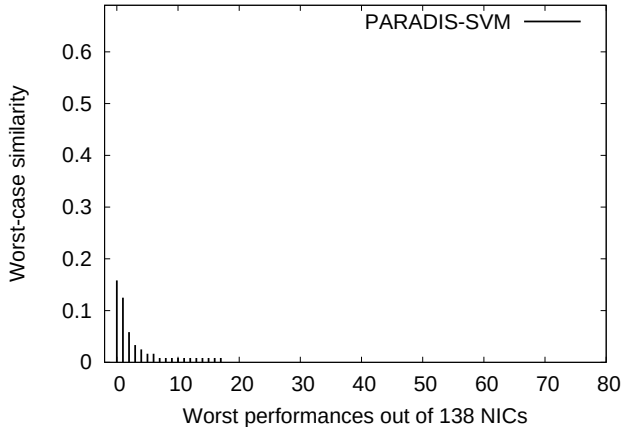

Figure 7: PARADIS-SVM: per-NIC similarity

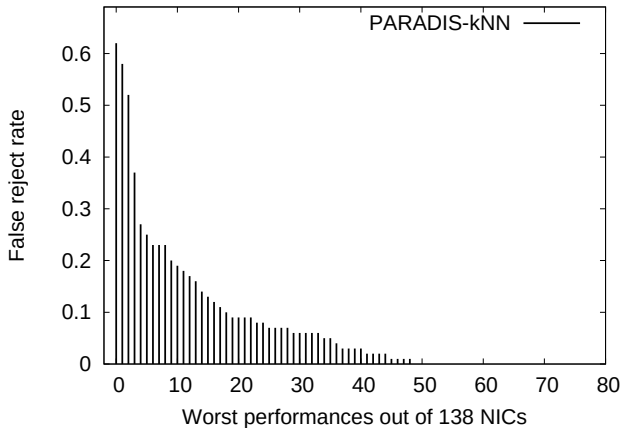

Figure 8: PARADIS-kNN: per-NIC FRR

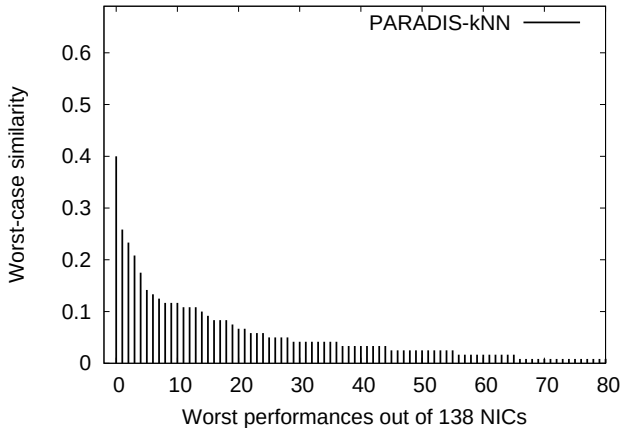

Figure 9: PARADIS-kNN: per-NIC similarity

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 8 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 9 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. 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: It is fairly obvious that PARADIS-SVM consistently outperforms PARADIS-kNN. However, PARADIS-kNN performs well enough 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 0.0034%, 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.

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

Table 3: Performance of travelling nodes

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 containes 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.

The nodes were identified by the last portion of their MAC address: 0d6396, 354534, 358563. Table 3 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 3 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 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 10. 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. This suggests that even 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

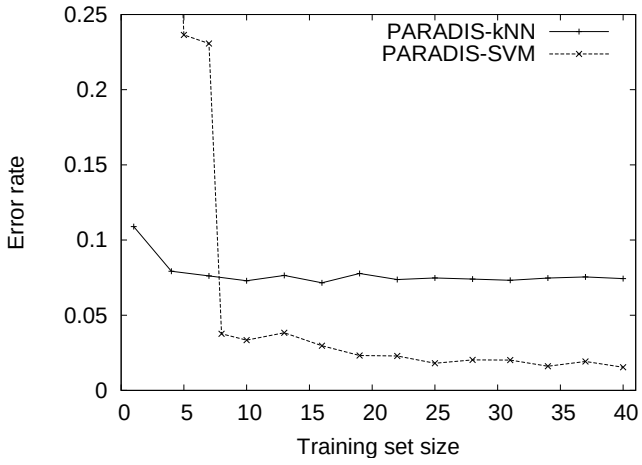

Figure 10: Effect of training set size on accuracy

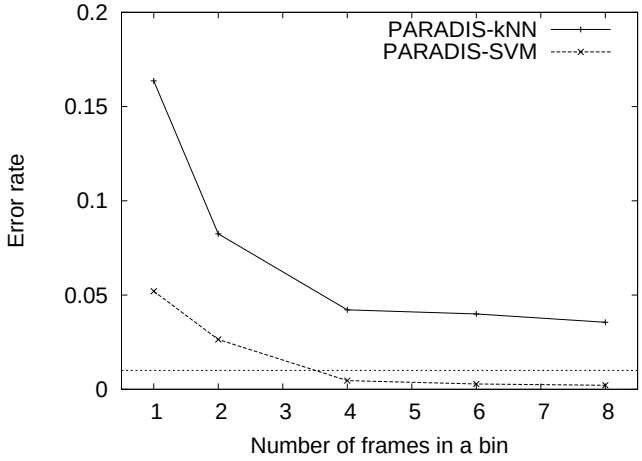

Figure 11: Effect of bin size on accuracy

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 11 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 a percentage point. 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 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, which are ag-

nostic of modulation, PARADIS defines a signal's signature in terms of structure imposed by the modulation scheme, thus greatly simplifying the problem. Specifically, signatures are expressed with respect to the ideal signal in terms of frequency, phase and magnitude errors, as well as I/Q offset, SYNC correlation and error vector magnitude.

Evaluation PARADIS involved over 130 identical IEEE 802.11 wireless NICs. 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 suggest that it could be especially useful for detection of compromised cryptographic material.

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] Carl Edward Rasmussen and Christopher K.I. Williams. *Gaussian Processes for Machine Learning*. The MIT Press, 2006.

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

[9] 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.

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

[11] 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.

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

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

[14] F. Guo and T. Chiueh. Sequence Number-Based MAC Address Spoof Detection. *Recent Advances in Intrusion Detection: 8th International Symposium, RAID 2005, Seattle, WA, USA, September 7-9, 2005: Revised Papers*, 2006.

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

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

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

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

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

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

[21] 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.

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

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

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

[25] 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.

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

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

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

[29] 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.

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

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

[32] 89600S series VXI-based Vector Signal Analyzer. Agilent technologies.

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

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

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

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