9-15-2016

Physical Layer Defenses Against Primary User Emulation Attacks

Joan A. Betances

PHYSICAL LAYER DEFENSES AGAINST PRIMARY USER EMULATION ATTACKS

DISSERTATION

Joan Addison Betances, Major, USAF

AFIT-ENG-DS-16-S-005

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY

AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

AFIT-ENG-DS-16-S-005

PHYSICAL LAYER DEFENSES AGAINST PRIMARY USER EMULATION ATTACKS

DISSERTATION

Presented to the Faculty

Graduate School of Engineering and Management

Air Force Institute of Technology

Air University

Air Education and Training Command

in Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

Joan Addison Betances, B.S.C.S., B.S.E.E., M.S.C.E.

Major, USAF

September 2016

AFIT-ENG-DS-16-S-005

PHYSICAL LAYER DEFENSES AGAINST PRIMARY USER EMULATION ATTACKS

DISSERTATION

Joan Addison Betances, B.S.C.S., B.S.E.E., M.S.C.E.
Major, USAF

Committee Membership:

Kenneth M. Hopkinson, PhD
Chairman

Major Mark D. Silvius, PhD
Member

Robert F. Mills, PhD
Member

Michael R. Grimaila, PhD
Member

Adedji B. Badiru, PhD
Dean, Graduate School of Engineering and Management

AFIT-ENG-DS-16-S-005

Abstract

Cognitive Radio (CR) is a promising technology that works by detecting unused parts of the spectrum and automatically reconfiguring the communication system's parameters in order to operate in the available communication channels while minimizing interference. CR enables efficient use of the Radio Frequency (RF) spectrum by generating waveforms that can coexist with existing users in licensed spectrum bands. Spectrum sensing is one of the most important components of CR systems because it provides awareness of its operating environment, as well as detecting the presence of primary (licensed) users of the spectrum.

Current CR spectrum sensing research efforts tend to focus on the development of new mechanisms to detect Primary User (PU) or improve existing ones. However, previous researchers have identified that a Primary User Emulation Attack (PUEA) can disrupt the operation of a CR system by significantly reducing the spectrum available to unlicensed users. This dissertation presents three methods to counteract PUEAs: Radio Frequency Distinct Native Attribute (RF-DNA), Constellation-Based Distinct Native Attribute (CB-DNA), and signal watermarking.

RF-DNA fingerprinting extract identifying features from RF signals using a Region of Interest (ROI) that remains constant for all transmissions such as preambles, midambles, pilot tones, etc. The true source of a transmission was correctly identified $\%C \approx 78\%$ in a test case that involves $N_{\text{devices}} = 15$ devices using Time Domain (TD) RF-DNA fingerprints.

CB-DNA fingerprinting uniquely identifies emissions from a radio by computing statistical features of the received signal projected into a constellation space. These features can be used to obtain device-specific information such as manufac-

turer, model, serial number, etc. In a test case involving $N_{\text{devices}} = 15$ devices, the mean correct classification rate was $\%C \approx 95\%$ using CB-DNA fingerprints.

The watermarking method establishes a side-channel that enables the exchange of a Hash Based Message Authentication Code (HMAC) that authenticates the source of a signal. The established side channel provides a reliable communication link even at low Signal to Noise Ratio (SNR) conditions. For example, the Bit Error Rate (BER) of the extracted watermark at an SNR=8 E_b/N_0 dB was 1.47×10^{-4}. The intellectual contributions of this dissertation are validated through experimentations.

Table of Contents

Page

Page

List of Figures

List of Tables

List of Acronyms

ACRO AFIT Cognitive Radio 51

AFIT Air Force Institute of Technology 51

AWGN Additive White Gaussian Noise 102

BER Bit Error Rate 102

CB-DNA Constellation-Based Distinct Native Attribute 1

COTS Commercial Off-The-Shelf 56

CRLB Cramer-Rao Lower Bound 94

CR Cognitive Radio 1

DFT Discrete Fourier Transform 34

DOS Denial of Service 100

DRA Dimensional Reduction Analysis 73

DSA Dynamic Spectrum Access 3

DSSS Direct Sequence Spread Spectrum 14

FPGA Field Programmable Gate Array 102

FVR False Verification Rate 22

PHY Physical Layer 100

PLL Phase-Locked Loop 89

PRNG Pseudo Random Number Generator 86

PSD Power Spectral Density 34

PSK Phase Shift Keying 89

PUEA Primary User Emulation Attack 1

PU Primary User 1

QAM Quadrature Amplitude Modulation 83

QPSK Quadrature Phase Shift Keying 102

RF-DNA Radio Frequency Distinct Native Attribute 1

RF Radio Frequency 1

ROC Receiver Operating Characteristic 21

ROI Region of Interest 1

SDR Software-Defined Radio 101

SD Spectral Domain 32

SNR Signal to Noise Ratio 102

SOI Signal of Interest 37

SU Secondary User 100

TD Time Domain 32

TVR True Verification Rate 21

USB Universal Serial Bus 14

USRP Universal Software Radio Peripheral 46

WRAN Wireless Regional Area Networks 4

PHYSICAL LAYER DEFENSES AGAINST PRIMARY USER EMULATION ATTACKS

I. Introduction

The rapid growth of wireless devices has created a strain on the available spectrum. This strain is further aggravated by the fixed allocation of spectrum resources dictated by current regulations. Spectrum surveys conducted within several cities in the United States revealed that licensed portions of the spectrum are sparsely utilized leaving large spectrum gaps unutilized [8, 9]. It is evident that a new licensing scheme to access the spectrum will be required in the near future.

Cognitive Radio (CR) is a new idea proposed by researchers at the beginning of the century to alleviate the spectrum scarcity. CR creates two classes of users: Primary User (PU) and Secondary User (SU). PUs are licensed users of the spectrum and they have priority above everybody else. SUs are unlicensed users who have equal access to the spectrum, whenever the PUs are not transmitting in their allocated space. Since SUs are unlicensed, they must access the spectrum in a way that does not cause interference with the PU. Additionally, CR aims to implement intelligent radio communication systems that are aware of their environment, and adjust their transmitter and receiver parameters to maximize spectrum efficiency while maintaining the ability of obtaining highly reliable communication system.

A Dynamic Spectrum Access (DSA) system that has two classes of users (PU and SU) who can be exploited by a malicious user who wants exclusive access of the spectrum by emulating the PU. A Primary User Emulation Attack (PUEA) is conducted by mimicking the PU signal's characteristics, causing SUs to identify the

attacker as a licensed user of the spectrum [10]. Researchers have identified that a PUEA can be used to generate a Denial of Service (DOS) – disrupting the operation of a cognitive radio system by significantly reducing the spectrum available to SUs.

Researchers have identified three main defenses against a PUEA: Naive detection, Localization based, and Physical Layer (PHY) coding. Naive detection methods detect a PUEA by estimating the mean and variance of the PU's signal and use these measurements to validate the source of transmission [11]. Localization-based defenses against PUEAs estimate the location of the source of the signal, and compare it to known PU locations for authentication [10]. PHY coding defenses estimate the location of the source of emissions by allowing a reference signal interfere with the PU's emissions and analyzing the results from the point-of-view at multiple receivers [12]. While these techniques are effective to some degree, security schemes based on geolocation are increasingly difficult to implement as they require obtaining measurements from several different sensors that are widely spaced around the PU location.

This dissertation presented three methods to detect a PUEA that are implemented at the PHY. The first method created Radio Frequency Distinct Native Attribute (RF-DNA) fingerprints and used them to authenticate the PU. The second method projected the received communication symbols into a constellation space and used these projections to create Constellation-Based Distinct Native Attribute (CB-DNA) fingerprints. Finally, the last method used watermarks to establish a communication channel that enables the exchange of Hash Based Message Authentication Code (HMAC) that authenticates the PU.

II. Detection of Primary User Emulation Attack Using Radio Frequency Distinct Native Attribute Fingerprinting Techniques

Abstract

Cognitive Radio (CR) is a promising technology that works by detecting unused parts of the spectrum and automatically reconfiguring Modulator/Demodulator (MODEM) parameters to operate in the available communication channels while minimizing interference. CR enables efficient use of the Radio Frequency (RF) spectrum by generating waveforms that can coexist with existing users in licensed spectrum bands. Spectrum sensing is one of the most important components of CR systems, because it provides awareness of the operating environment, as well as detecting the presence of primary (licensed) spectrum users. Current CR research efforts are focused on the development of new mechanisms to detect Primary Users (PUs) or improve existing ones. However, previous researchers have identified that a Primary User Emulation Attack (PUEA) can disrupt the operation of a CR system by significantly reducing the spectrum available to unlicensed users. This research proposed a transmitter verification scheme to validate PUs using RF fingerprinting. RF fingerprinting uniquely identifies a commercial radio by extracting features from the collected emissions. These features can be used to obtain device-specific information such as manufacturer, model, serial number, etc.

2.1 Introduction

Dynamic Spectrum Access (DSA) is a new paradigm that permits reutilization of unused portions of the spectrum, when the Primary User (PU) (licensed user) is not occupying its allocation of the spectrum. The Institute of Electrical and Electronics

Engineers (IEEE) is currently developing a new standard for DSA users. The Wireless Regional Area Networks (WRAN) standard provides means for DSA usage of the TV portion of the spectrum. This standard specifies the frequency allocation for the United States as: 54-60, 76-88, 174-216, 470-608 and 614-698 MHz, for a total of 282MHz spanning 47 TV channels [13].

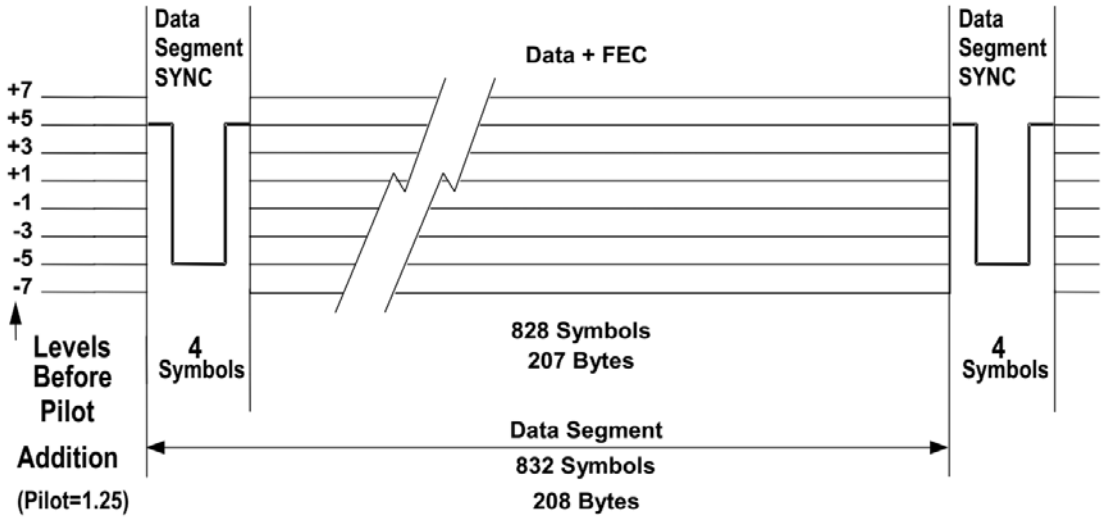

Figure 1. ATSC Digital Television Standard: RF/Transmission System Characteristics [1]

Traditional cognitive radio research centers around the parts of the spectrum set aside for TV stations, as a primary target for secondary user utilization. Digital TV signals transmit a synchronization pattern that can be exploited by using Radio Frequency Distinct Native Attribute (RF-DNA) to identify the emitter. The synchronization portion for digital TV signals is illustrated in Figure 1. This research assumed that the signal of interest contained a synchronization field that remained constant for all collections.

Software-Defined Radios (SDRs) are highly configurable and have the capability to generate arbitrary signals. It is possible for a SDR, such as the Universal Software Radio Peripheral (USRP) X310, to generate signals that closely resemble a digital TV station's transmissions. Such an attack can be easily accomplished by storing

samples of a digital TV signal and replaying them later. This research proposed a mechanism to generate RF-DNA fingerprints that can be used to classify and verify signals that contain a fixed synchronization field.

Prior researchers have determined mechanisms to detect a Primary User Emulation Attack (PUEA) based on estimating the transmitter location [10, 11, 14, 15, 16, 17, 18, 19] and comparing it to known PU emitter locations. Emitter geolocation solutions require measurements from several sensors, which are widely spaced around the emitter. This research described a novel method to verify the identity of the PU using Radio Frequency (RF) fingerprinting without the aid of a sensor network. The ability to verify the identity of the PU, without cooperation from other nodes, is one key advantage of this research.

The PU verification scheme relied on examining waveforms at the Physical Layer (PHY), which will uniquely identify devices based on inherent differences in their transmissions. This verification scheme required prior signal collection of PU's transmissions. RF fingerprints were generated using the synchronization parameters (preambles, postambles, midambles, pilot tones, etc) of the protocol used by the PU. PUEA need to mimic the protocol used by the PU in order to fool secondary users. The forged transmissions needed to include the synchronization parameters of the protocol used by the PU– enabling the verification of the signal source using RF fingerprinting.

Every device that emits RF signals has unique characteristics that are very difficult to duplicate. Thus, these characteristics may be used to uniquely identify transmitters. These characteristics are observed as transient behavior with respect to the instantaneous amplitude, phase, and frequency of the radiated signal. This behavior can be caused due to a variety of reasons, such as precision of frequency synthesis systems, modulator subsystems, and RF amplifiers. Unique transient signals can be observed even among transmitters of the same type and model. This differentiation is

due to manufacturing tolerances and component aging used in the device [20]. These transmitter anomalies can be used to create RF fingerprints.

2.2 Background

This section provides the technical background supporting the methodology described in section 2.3. The topics covered in the section include: generation of Time Domain (TD) Radio Frequency (RF) fingerprints, generation of spectral domain RF fingerprints, and classification of systems using Multiple Discriminant Analysis / Maximum Likelihood (MDA/ML).

Time Domain RF Fingerprinting.

RF fingerprints were generated by passively collecting signals generated by MODEMs, as they transmit communication symbols. The collected signal were represented in the TD as the complex vector $\mathbf{x}[n] = s_I(n) + js_Q(n)$ for $n = \{0, 1, 2, ..., N-1\}$, where n specified the time when the sample was measured, and the variable N specified the total number of samples stored in the vector. The instantaneous amplitude, phase, and frequency of $\mathbf{x}$ can be computed as follows [3]:

$$a(n) = \sqrt{(s_I(n) + js_Q(n))^2}, \qquad n = \{0, 1, 2, ..., N-1\}, \tag{1}$$

$$\phi(n) = \tan^{-1}\left[\frac{s_Q(n)}{s_I(n)}\right], \qquad s_I(n) \neq 0, \qquad n = \{0, 1, 2, ..., N-1\}, \tag{2}$$

$$f(n) = \frac{1}{2\pi}\left[\frac{d\phi(n)}{dn}\right] \qquad n = \{0, 1, 2, ..., N-1\}. \tag{3}$$

The quality of RF fingerprints generated using instantaneous amplitude, phase, and frequency can be improved by normalizing and centering the collected signal of

interest. Centering and normalization of the signal can be obtained by

$$a_c(n) = \frac{a(n) - \mu_a}{\max(a_c(n))}, \tag{4}$$

$$\phi_c(n) = \frac{\phi(n) - \mu_\phi}{\max(\phi_c(n))}, \tag{5}$$

$$f_c(n) = \frac{f(n) - \mu_f}{\max(f_c(n))}, \tag{6}$$

where μ_a, μ_ϕ, μ_f, were the respective amplitude, phase, and frequency means [3].

RF fingerprints were obtained by dividing the sequences $a_c(n)$, $\phi_c(n)$, $f_c(n)$ into R equal-length sequences. The distinct fingerprints were generated by computing the standard deviation (σ), variance (σ^2), skewness (γ), and kurtosis (κ) of these sequences to create new vectors as follows:

$$\mathbf{F}_r^a = [\sigma_a, \sigma_a^2, \gamma_a, \kappa_a], \tag{7}$$

$$\mathbf{F}_r^\phi = [\sigma_\phi, \sigma_\phi^2, \gamma_\phi, \kappa_\phi], \tag{8}$$

$$\mathbf{F}_r^f = [\sigma_f, \sigma_f^2, \gamma_f, \kappa_f]. \tag{9}$$

The composite fingerprint was generated by concatenating the individual $\mathbf{F}^\sigma$ sequences, where σ denotes a specific amplitude, phase, or frequency sequence by

$$\mathbf{F}^\sigma = \begin{bmatrix} \mathbf{F}_1^\sigma & \vdots & \mathbf{F}_2^\sigma & \cdots & \mathbf{F}_R^\sigma \end{bmatrix}. \tag{10}$$

The composite amplitude, phase, and frequency fingerprints may be combined in order to generate a complete TD fingerprint as follows:

$$\mathbf{F}_{TD} = \begin{bmatrix} \mathbf{F}^{a} & \vdots & \mathbf{F}^{\phi} & \vdots & \mathbf{F}^{f} \end{bmatrix} \tag{11}$$

A visual depiction of the generated RF fingerprints is shown in Figure 2. The Figure shows the RF fingerprints for eight different devices. The values for the variance, skewness, and kurtosis of the signal generated by the devices are shown in the horizontal bands. The colors represent the average value for each statistical measurement scaled to span 0 to 1 [2].

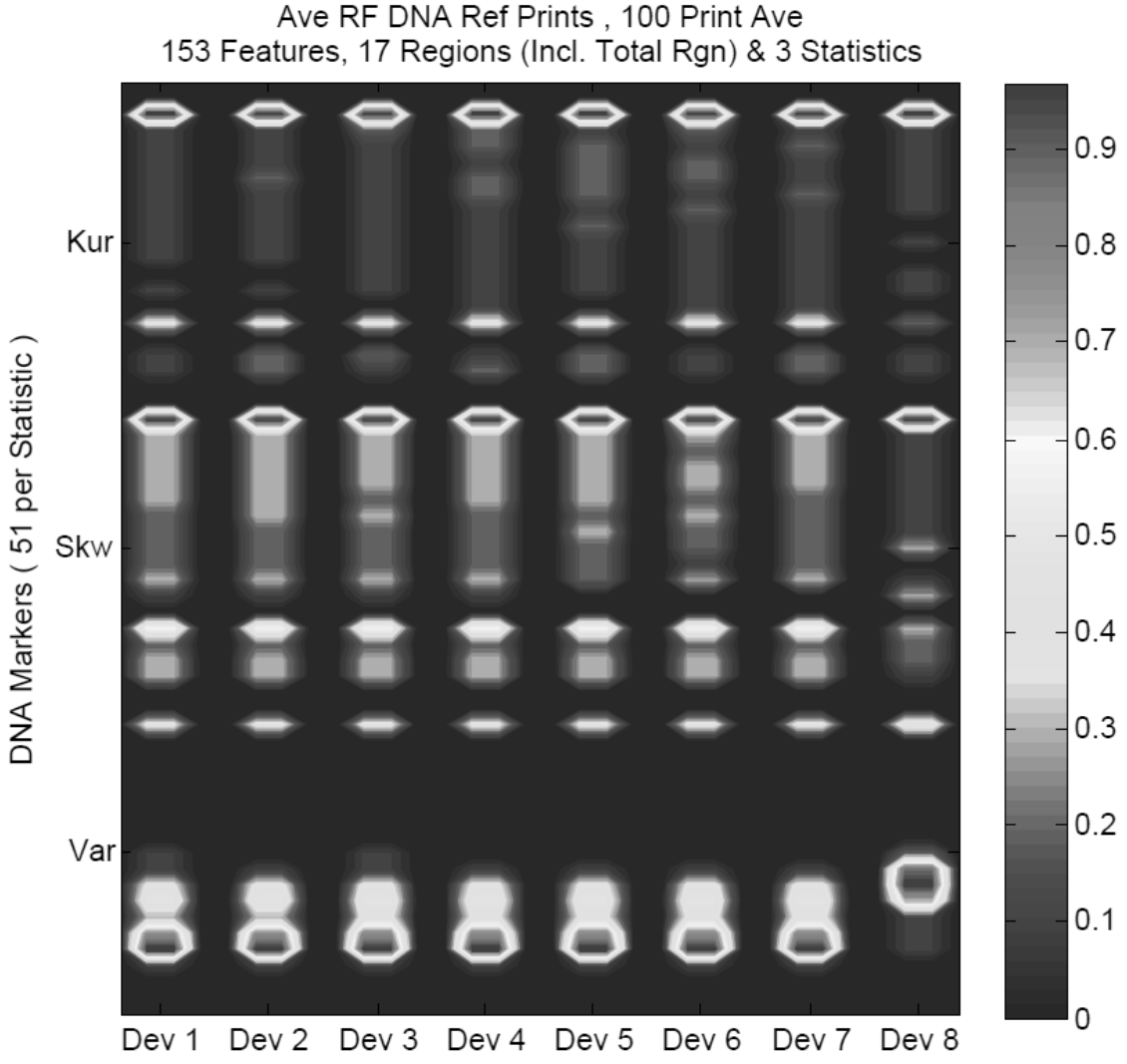

Figure 2. RF Fingerprint Visualization for 8 Devices [2]

Spectral Domain RF Fingerprinting.

Spectral Domain (SD) RF fingerprints were generated using the Power Spectral Density (PSD) of the TD signal represented in vector $\mathbf{x}$. The SD representation of $\mathbf{x}$ was computed using the Discrete Fourier Transform (DFT). The mathematical model to compute the DFT is as follows:

$$\mathbf{X}(k) = \frac{1}{N} \sum_{n=0}^{N-1} \mathbf{x}(n) e^{\frac{-j2\pi kn}{N}} \quad \text{for } k = \{0, 1, 2, ..., N-1\} \tag{12}$$

In this mathematical model, $\mathbf{X}(k)$ is a complex number representing the frequency component of a signal at band k, while $\mathbf{x}(n)$ represents the signal as it is being sampled in the time domain [21]. The PSD of the signal is normalized with respect to power in order to mitigate collection effects that may affect signal classification [3]. The average power of the signal is computed by:

$$P_{\mathbf{X}} = \frac{1}{N} \sum_{n=0}^{N-1} \mathbf{X}(n)\mathbf{X}(n)^*, \tag{13}$$

and the normalized-power PSD sequence is obtained by:

$$\overline{\mathbf{X}}(k) = \frac{1}{P_{\mathbf{X}}} \left|\mathbf{X}(k)\right|^2. \tag{14}$$

Once the normalized PSD signal was obtained, the SD fingerprints were generated by dividing the sequence into R equal length sequences. The distinct fingerprints were generated by computing the standard deviation (σ), variance (σ^2), skewness (γ), and kurtosis (κ) of these sequences to create new vectors as follows:

$$\mathbf{F}_r = [\sigma, \sigma^2, \gamma, \kappa]. \tag{15}$$

The composite fingerprint was generated by concatenating the individual **F** sequences by:

$$\mathbf{F} = \left[\mathbf{F}_1 \quad \vdots \quad \mathbf{F}_2 \quad \cdots \quad \mathbf{F}_R\right]. \tag{16}$$

The resultant full-dimensional fingerprint vector **F** from 16 contained a total of N_f = (# of Features)×(# of Statistical Metrics)×(# of Regions) elements. This vector is illustrated in Figure 3.

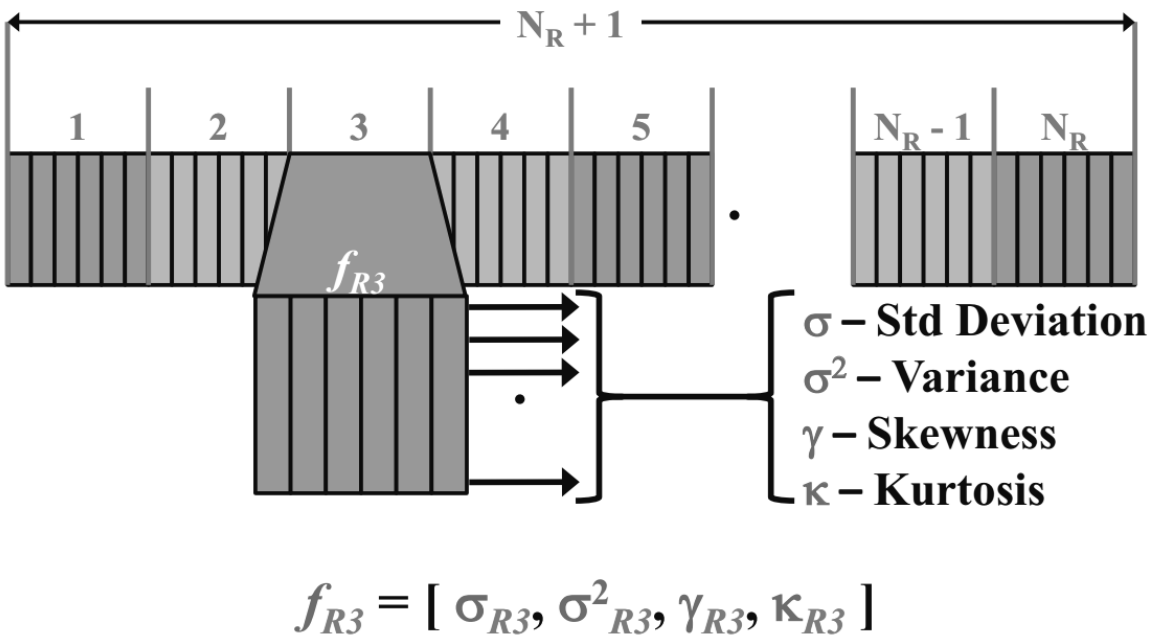

Figure 3. RF-DNA Statistical Fingerprint Generation for Centered and Normalized Feature Sequences and $N_R + 1$ Total Subregions [3]

Multiple Discriminant Analysis/Maximum Likelihood.

The purpose of RF fingerprints is to extract features from a signal, so that it can be classified. Classification of RF fingerprints requires additional processing, because they can generate a multivariate statistical model with hundreds of independent variables. Obtaining a Maximum Likelihood Estimate (MLE) of the source of a RF emanation can be computationally intensive due to the high dimensionality of the statistics. This problem was simplified by using an Multiple Discriminant

Analysis (MDA) algorithm.

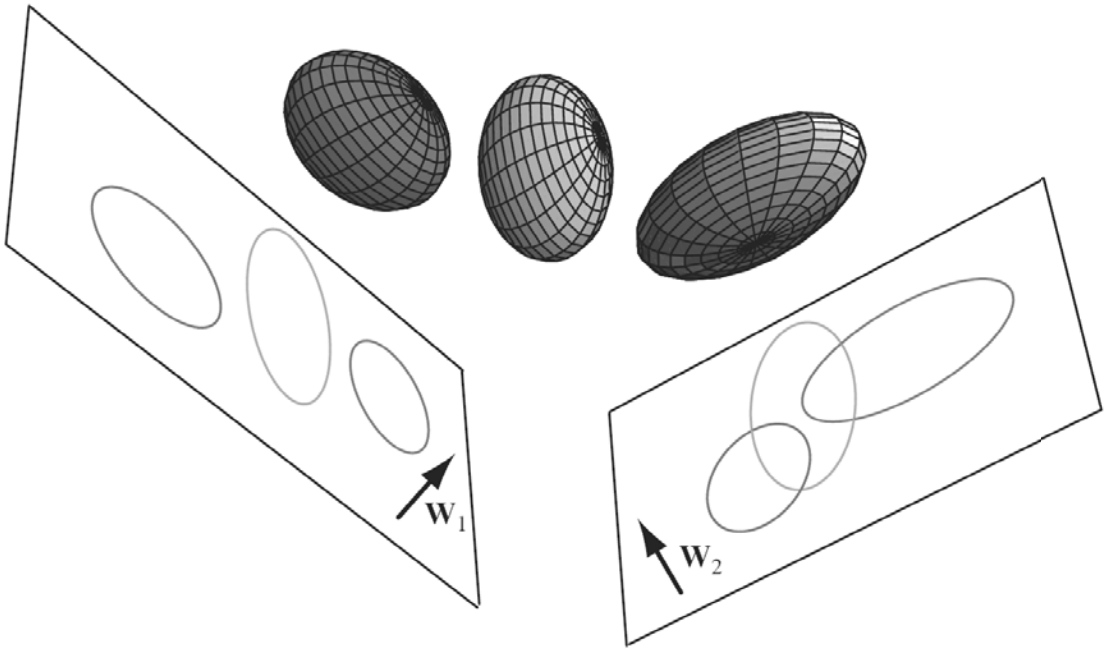

Figure 4. MDA Projection of 3D Space into 2D Space [2]

MDA is a multivariate statistical technique to apply linear discriminant analysis [22]. The objective of MDA is to classify objects into two or more mutually-exclusive classes by reducing the dimensionality of a set of independent variables. The dimensionality reduction is accomplished by identifying the smallest linear combination of variables with normal errors that best discriminate between classes [23]. For example, the 3D model shown in Figure 4 was projected onto 2D models in order to reduce the dimensionality of the problem. The 2D projections were defined by the norm vectors **W**1 and **W**2 respectively. Classification and discrimination along the **W**2 projections were significantly more difficult because the projections overlap. However, the **W**1 subspace facilitated classification and discrimination, because the projections do not overlap. The MDA protocol aimed to determine projections such as those provided by the **W**1 vector.

The MDA algorithm started by defining two scatter matrices, the inter-class ma-

trix (S_b) and the intra-class matrix (S_w) of the dataset x. The MDA projection maximized inter-class distances while minimizing intra-class spread. These matrices are defined by [3]:

$$\mathbf{S}_b = \sum_{i=1}^{N_c} P_i \sum_i, \tag{17}$$

$$\mathbf{S}_B = \sum_{i=1}^{N_c} P_i(\mu_i - \mu)(\mu_i - \mu)^T, \tag{18}$$

Where N_c is the number of classes, P_i is the prior probability of class c_i, and $\sum_i$ is the covariance matrix. Using the two scatter matrices the projection matrix $\mathbf{W}$ was formed using the eigenvectors of $\mathbf{S}_w^{-1}\mathbf{S}_b$ [3]. The multivariate statistics can be projected into a $(N_c - 1)$ dimensional subspace by [3]:

$$\mathbf{F}_i^{\mathbf{W}} = \mathbf{W}^T\mathbf{F}, \tag{19}$$

where $\mathbf{F}$ is the matrix representing the fingerprint.

2.3 Methodology

Wireless communication systems are susceptible to a myriad of attacks, because the transmission medium is hard to constrain to specific locations – making it accessible to unauthorized users. This research aimed to characterize a security mechanism that operated at the Physical Layer (PHY) in order to prevent Primary User Emulation Attack (PUEA). The proposed solution generated a unique PHY Radio Frequency Distinct Native Attribute (RF-DNA) fingerprint that can be used to authenticate the Primary User (PU). This section describes the methodology used to obtain the experimental results described in section 2.4.

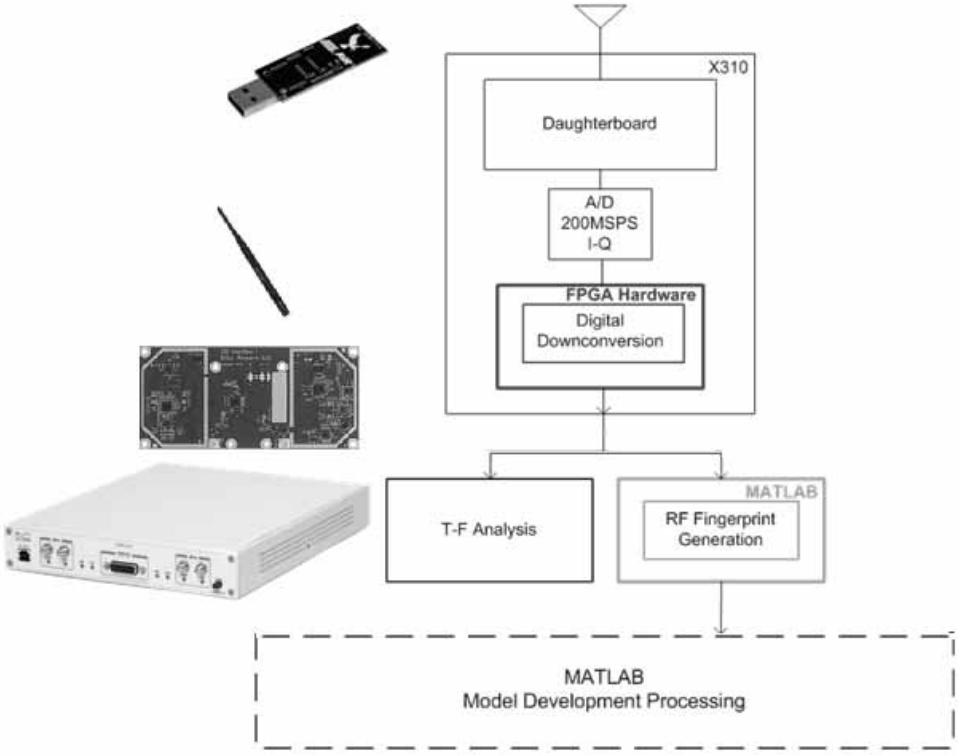

Figure 5. X310 SDR Methodology for Assessing RF-DNA Fingerprinting Using MATLAB® [4]

SDR Receiver Configuration.

The receiver/transmitter used in this research was a National Instruments (NI) Universal Software Radio Peripheral (USRP) X310 Software-Defined Radio (SDR). This research departed from the norm by using a relatively inexpensive Radio Frequency (RF) transmitter/receiver. Research of RF-DNA fingerprinting is normally conducted using highly precise and accurate collection receivers that cost over $150,000. The X310 SDR is available Commercial Off-The-Shelf (COTS), with a retail price of approximately $7,000. In addition to its price tag, the RF transmitter/receiver was chosen for this research because it had a very capable Field Programmable Gate Array (FPGA) that can be used for signal processing.

ZigBee Signal.

The RF-emitting devices used in this research included AVR RZUSBsticks and X310 SDRs. RZUSBstick is a device designed by Atmel Corporation for the development, debugging, and demonstration of Institute of Electrical and Electronics Engineers (IEEE) 802.15.4, 6LoWPAN, and ZigBee [24]. The RZUSBstick uses the Universal Serial Bus (USB) for configuration, transmission, and reception of ZigBee data.

The RZUSBstick devices were configured to transmit a beacon request at a rate of $\mathrm{BR}_{\mathrm{rate}} = 10$ BR/s (Beacon Request per second). The devices were configured to transmit using ZigBee channel 26, which has a center frequency of $F_{\mathrm{carr}} = 2.48$GHz. Only one ZigBee device was radiating RF signal at a time, and were positioned $\mathrm{Tx}_{\mathrm{Distance}} = 8.0$ cm from the receiver antenna at the time of collection.

Each beacon request transmitted had $\mathrm{N}_{\mathrm{symbols}} = 32$ Offset Quadrature Phase Shift Keying (O-QPSK) symbols. The signal had two fields: the preamble and the payload. The preamble consisted of $\mathrm{N}_{\mathrm{symbols}} = 8$ O-QPSK symbols and the payload consisted of $\mathrm{N}_{\mathrm{symbols}} = 24$ O-QPSK symbols. The symbols were transmitted at a rate of $\mathrm{Chip}_{\mathrm{rate}} = 2$ MChips/s (250 kbps). Each symbol was mapped to one of 16 pseudo-random, 32-chip sequences in order to create a Direct Sequence Spread Spectrum (DSSS) signal. The characteristics of the ZigBee beacon request signal is illustrated in Figure 6.

ZigBee specifications require that the $\mathrm{N}_{\mathrm{symbols}} = 8$ O-QPSK symbols that form the preamble are mapped to the `[1100101]` bit sequence. Therefore, the first eight symbols of every ZigBee burst are identical. Fingerprints were generated based on the preamble, since all preambles are identical at the bit-level regardless of the device transmitting. The preamble was divided into $\mathrm{N}_{\mathrm{regions}} = 8$ regions, one region per symbol, in order to compute the statistical characteristics required for a RF-DNA

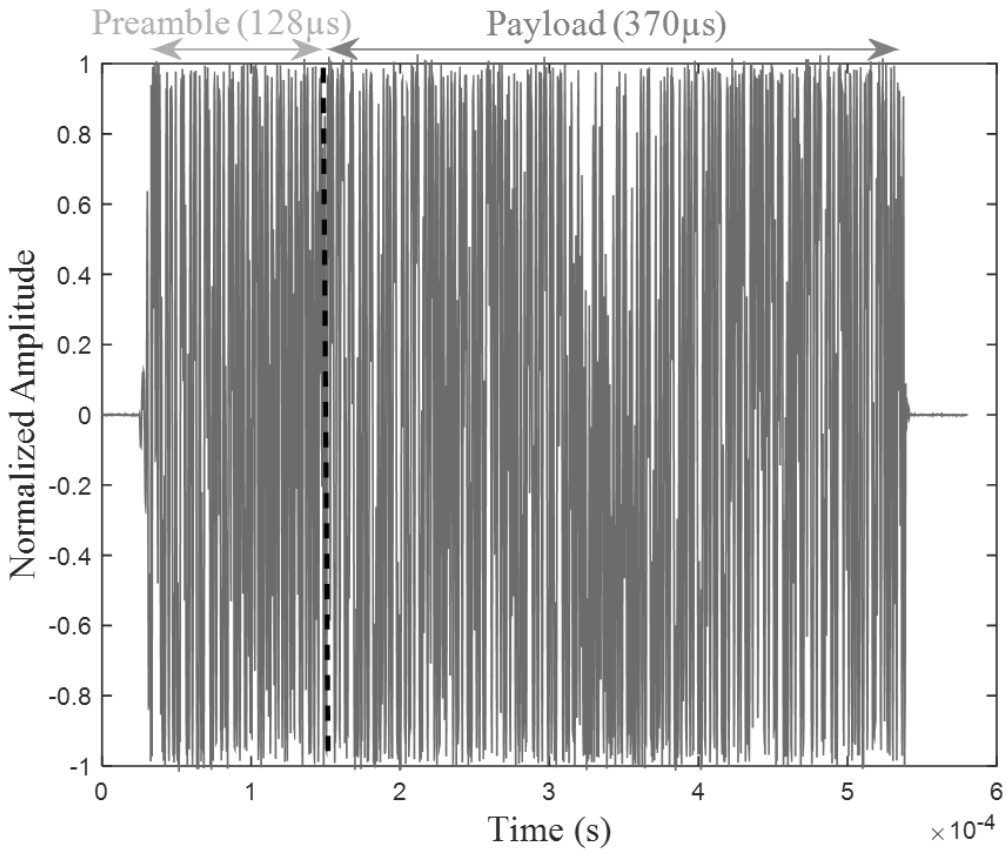

Figure 6. Time Domain Response of Experimentally-Collected ZigBee Burst

fingerprint. Figure 7 illustrates the eight regions used to generate the fingerprints.

Experimental Signal Collection.

The devices under test were inside a Ramsey STE6000 RF Shielded Test Enclosure. This test enclosure was designed for use with Industrial Scientific and Medical (ISM) band signals including Bluetooth, WiFi, and ZigBee. The STE6000 provided isolation greater than 90dB at the 2.4Ghz ISM band. Additionally, the interior had an RF absorbent foam liner that attenuated signal reflections within the test enclosure by more than 24dB. The STE6000 was equipped with Ethernet and USB connections, in order to control the devices operating inside test enclosure while it was sealed. Figure 8 shows the test enclosure used for this research.

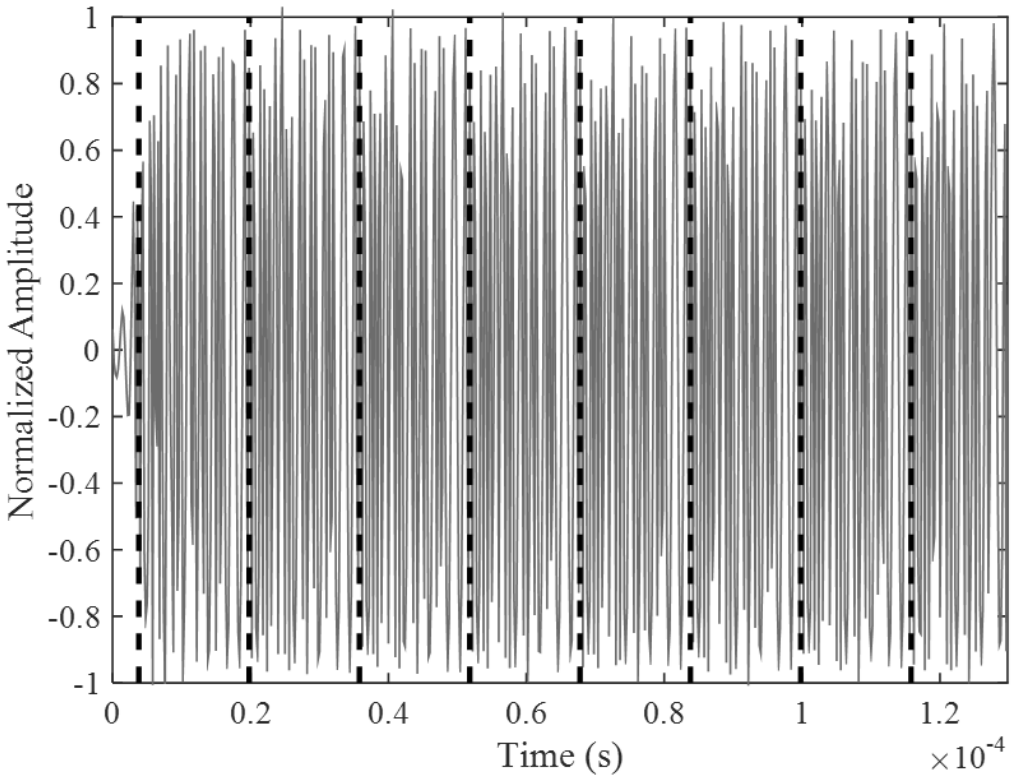

Figure 7. ZigBee Preamble Time Domain Response

The X310 SDR has transmit and receive capabilities covering from DC to 6.0 GHZ depending on daughterboard installed. For this research, the SBX-40 daughterboard was installed in the collection receiver, which provided a receive frequency range of 400-4400 MHz with a maximum instantaneous bandwidth of 40MHz. The receiver was configured to collect signals with a center frequency of $F_{carr} = 2.48$ GHz and a sampling rate of $F_{samp} = 5$MS/s. The collection receiver configuration remained fixed throughout all trials.

2.4 Results and Analysis

The simulation scenario consisted of $N_{devices} = 4$ devices. $N_{fprints} = 1000$ Time Domain (TD) Radio Frequency Distinct Native Attribute (RF-DNA) fingerprints

Figure 8. Ramsey STE6000 Shielded Test Enclosure

were generated for each device. The first three devices were RZUSBStics, and the fourth device was an X310-Software-Defined Radio (SDR) emulating the RZUSBStick device-3. The emulation was accomplished by capturing and playing back the signal radiated by RZUSBStick-3 using the X310-SDR. The signal was collected at a sample rate of $F_{samp} = 25$MS/s and replayed at the same rate of $F_{samp} = 25$MS/s.

The signal collected from the RZUSBStick had an SNR = 55 dB. The transmitter gain emulating RZUSBStick-3 was adjusted to obtain a Signal to Noise Ratio (SNR) of 55dB, in order to match the signal power collected for the other devices. The SNR was computed by taking the ratio of two measurements: the average power of the signal plus noise, and the average power collected without any signal present (noise).

The Multiple Discriminant Analysis / Maximum Likelihood (MDA/ML) algorithm was used to project the RF fingerprints onto a 2D subspace. $N_{fprints} = 500$ fingerprints per device were used to develop the 2D model. Additive White Gaussian Noise (AWGN) was used to create two noise realizations per fingerprint for a total

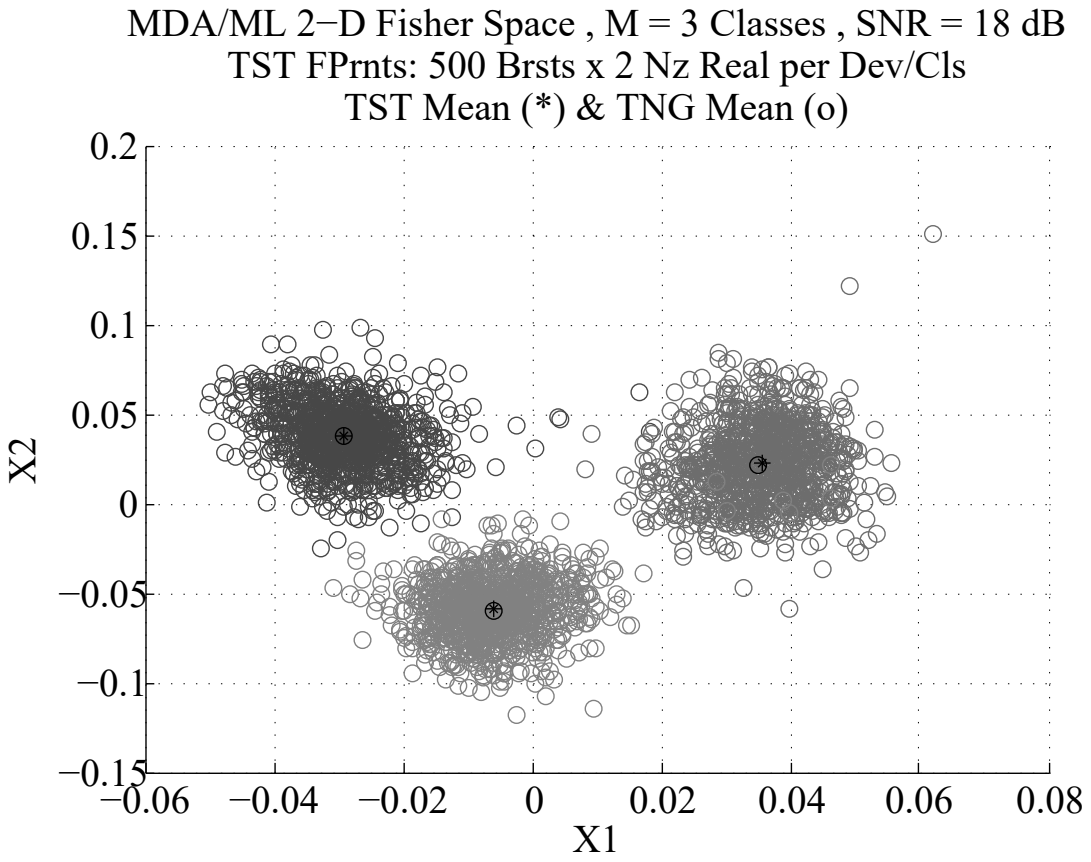

Figure 9. MDA/ML Projection for Three RZUSBStick Devices

of 1,000 fingerprints per device. The collected signal quality was degraded from a SNR=55dB down to SNR=18dB, in order to simulate transmitter normal operating conditions. The signal degradation was accomplished by adding a scaled Pseudo Random Number Generator (PRNG) to the signal of interest.

Each fingerprint had $N_{regions} = 9$ and $N_{feats} = 81$ different features. The features were generated by computing the variance (σ^2), skewness (γ), and kurtosis (κ) for the instantaneous amplitude $a_c(n)$, instantaneous frequency $f_c(n)$, and instantaneous phase $\phi_c(n)$ of each region. The 2D projections of the fingerprints are illustrated in Figure 9. The illustration shows that the vast majority of the fingerprints were clustered according to their respective devices (classes), but there are some outliers that may results in misclassification.

RF Fingerprints Classification.

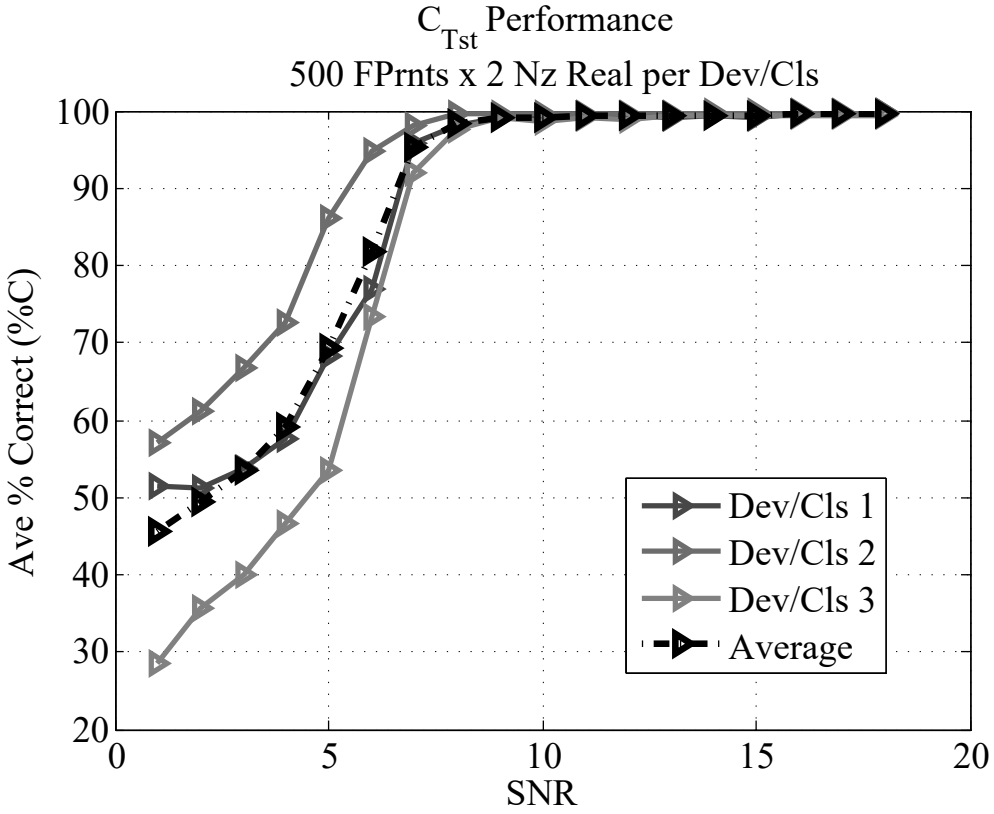

Figure 10. Fingerprint Classification Performance for Three RZUSBStick Devices

$N_{fprints} = 500$ fingerprints per device were used to compute the fingerprint classification performance. The fingerprints used to determine classification performance were different than the fingerprints used to generate the MDA/ML model. The signal quality was gradually degraded from a SNR=18dB down to a SNR=0dB using 1dB decrements. Signal degradation was accomplished by adding a scaled PRNG to the signal of interest. The scaled PRNG was used to represent the presence of AWGN in the communication channel. The PRNG used had an approximate mean of zero and an approximate standard deviation of one. Therefore, the scaling factor required to achieve an average noise power P_{avg} can be obtained by $\sqrt{P_{avg}}$.

Each step of the different curves shown in Figure 10 was computed using Nz=2

noise realizations per fingerprint, for a total of $N_{fprints} = 1000$ fingerprints per device. The resultant fingerprints were classified one at a time. The projection of each fingerprint was compared to the mean of the fingerprints projection of the four devices in the MDA/ML model. The unknown fingerprint was projected into a Fisher space. The Euclidian distance was computed for the unknown fingerprint, as well as the mean of every known device. The classification algorithm associated the unknown fingerprint with the device with the shortest Euclidian distance.

Classification algorithm performance varied depending on the SNR of the signal, as illustrated in Figure 10. The system correctly classified nearly %C=100% of the fingerprints, when the SNR was greater than 20dB and misclassified %C=50% of the fingerprints on average when the SNR equaled 0dB. The projections for Device 1 have the most compact cluster with maximum separation in respect to other devices. This separation allowed the system to correctly classify Device 1 fingerprints %C=70% of the time on average with a SNR of 3dB. However, the system correctly classified about %C=35% of the fingerprints for Device 3, when the SNR equaled 3dB, due to the widespread projections for this device.

RF Fingerprint Verification.

The MDA/ML model may be used to verify the identity of a transmitter. Verification was accomplished by measuring how similar an unknown fingerprint obtained from a signal claiming to be device-x to a model developed with actual fingerprints of device-x. For example, comparing the similarity between the fingerprint of a collected WiFi signal from a device claiming to have a Media Access Control (MAC) address (`01:23:45:67:89:ab`), to fingerprints previously collected from the device with MAC address (`01:23:45:67:89:ab`).

This test included the $N_{devices} = 3$ devices that were used for the MDA/ML clas-

sification model and introduced a new device not part of the MDA/ML classification model. Devices 1, 2, 3 were the original RZBUSBStick devices used to generate the MDA/ML model, and Device 4 was a RZBUSBStick rogue device. Table 1 shows the relationship among devices.

Table 1. RZBUSBStick Devices Plus a Rogue Device

Device	Color Code
RZUSBStick-1	Blue
RZUSBStick-2	Green
RZUSBStick-3	Red
RZUSBStick (Rogue)	Dev 4

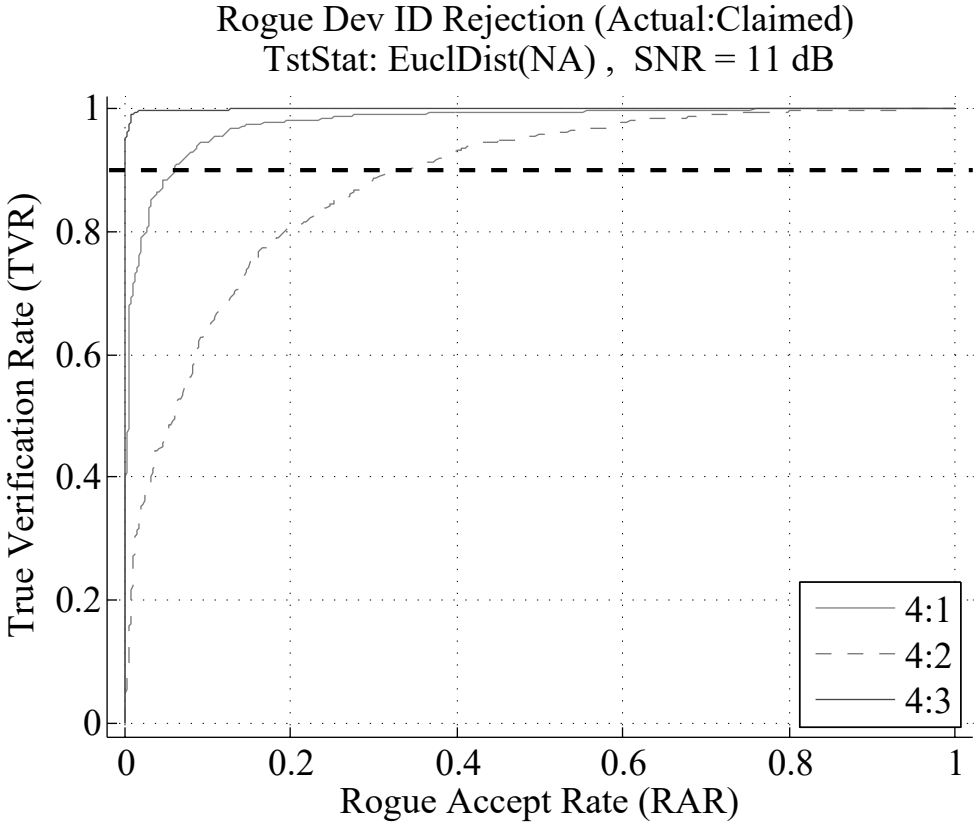

Figure 11. Fingerprints Verification Performance For Rogue RZUSBStick Device

The performance of verification system is typically characterized using Receiver Operating Characteristic (ROC) curves. The vertical axis represents the True Verifi-

cation Rate (TVR) defined as [25]:

$$\text{True Verification Rate} = \frac{\text{Positives correctly classified}}{\text{Total positives}}. \tag{20}$$

The horizontal axis represents the False Verification Rate (FVR) defined as [25]:

$$\text{False Verification Rate} = \frac{\text{Negatives incorrectly classified}}{\text{Total negatives}}. \tag{21}$$

Figure 11 illustrates system performance verifying device identity. The measure of similarity used for this system was Euclidian distance. The system classified fingerprints with a SNR$\geq$ 11dB for Device 1 nearly perfectlu (TVR $>$ 99% and FVR $<$ 3%). Device 2 achieved a TVR$\geq$ 90% with a FVR$\leq$ 6%, while Device 3 yielded a TVR$\geq$ 90% with a FVR$\leq$ 36%. The difference of verification performance among devices was explained by the distribution for the different devices, the similarity between the transmission of the rogue device and the known devices, as well as the different covariances of the distributions.

RF-DNA Fingerprint Verification for X310-SDR Replay Attack.

This test included the $N_{devices} = 3$ RZUSBStick devices that were used for the MDA/ML classification model and introduced a new device, which has not been profiled before. Devices 1, 2, and 3 were the original RZBUSBStick devices used to generate the MDA-ML model, and Device 6 was a X310-SDR replaying the signal generated by RZUSBStick Device 3. Table 2 shows the relationship among devices.

Figure 12 illustrates the performance of discriminating an SDRs identity when it is replaying the signal from a device known to the classification model. The measure of similarity used for this system was Euclidian distance. When the system was operating with an SNR greater than 11dB, device classification was perfect, i.e. the

Table 2. RZBUSBStick Devices plus X310-SDR Replay Attack

Device	Color Code
RZUSBStick-1	Blue
RZUSBStick-2	Green
RZUSBStick-3	Red
X310-SDR (Replay Dev-3)	Dev 6

SDR-replay was rejected (FVR=0%), while the actual device transmitting was classified correctly (TVR=100%). The signal was degraded using AWGN to 5dB, and

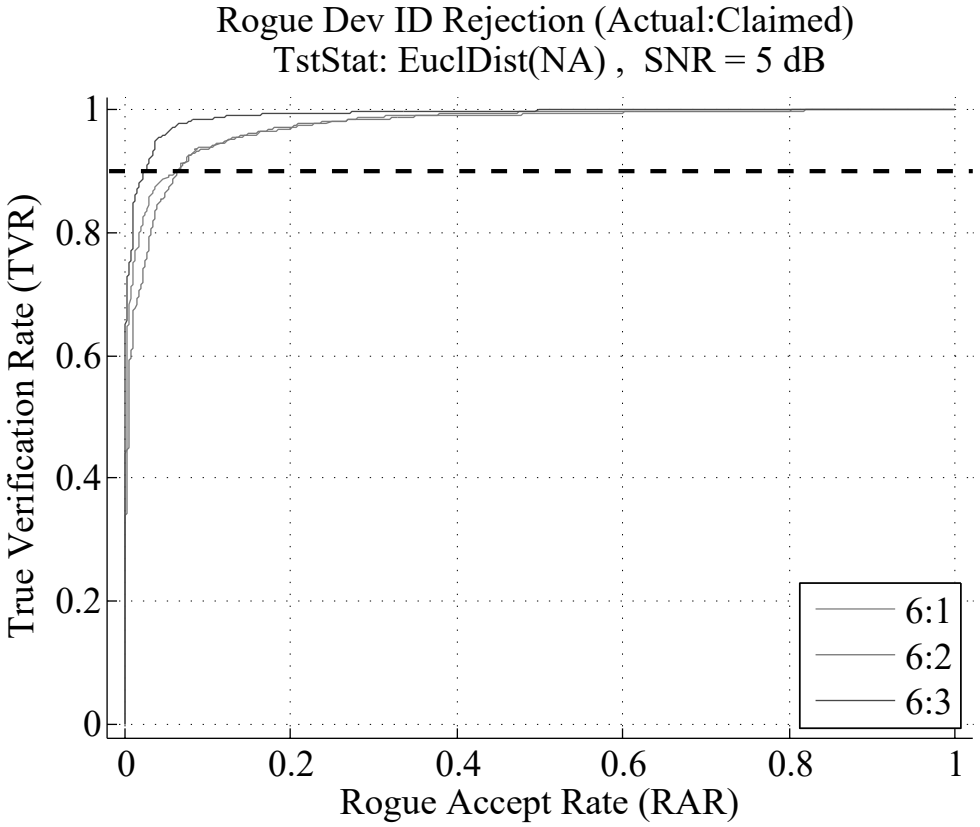

Figure 12. Fingerprints Verification Performance For X310-SDR Replay Attack

the system performance was recorded, as shown in Figure 12. The system classified fingerprints with a SNR$\geq$ 5dB for Device 1 nearly perfectly with TVR $>$ 94% and a corresponding FVR $<$ 3%. Device 2 and Device 3 also achieved high performance, as well with a TVR$\geq$ 90% and a corresponding FVR$\leq$ 6%. The difference of verifica-

tion performance among devices was explained by the distribution closeness for the different devices, the similarity between the transmission of the rogue device and the known devices, as well as the different covariances of the distributions.

RF-DNA Fingerprint Model Development for X310-SDR Devices.

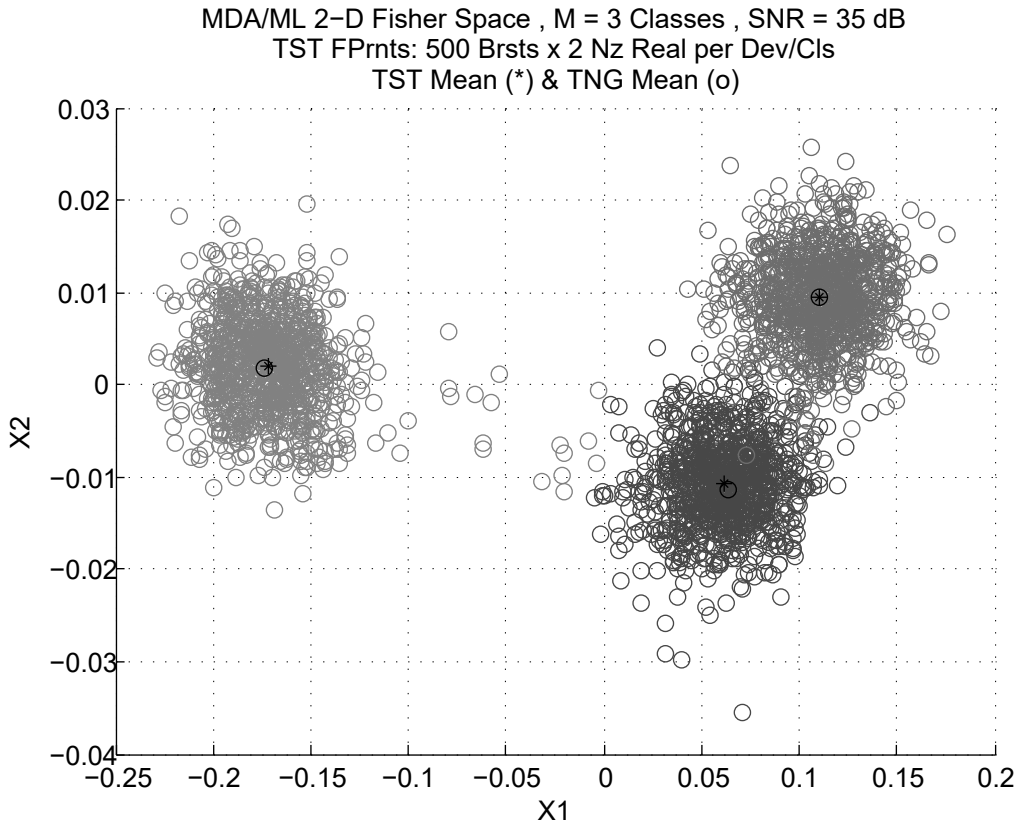

Figure 13. RF-DNA Fingerprints MDA/ML Projection of Three X310-SDR

The simulation scenario consisted of three X310-SDR. The objective of this test was to characterize the performance of the classification/verification algorithm for SDRs with like-configuration. The first two devices were X310 SDRs with an SBX daughterboard. The SBX daughterboard provided a transmit/receive frequency range of 400-4400 MHz with a maximum instantaneous bandwidth of 40 MHz. The third device was an X310 SDR configured with a CBX daughter board and a Global Positioning System Disciplined Oscillator (GPSDO). The GPSDO provided a high-

accuracy reference clock signal that minimized frequency and phase artifacts of the transmitter/receiver. The signal was collected by a X310 SDR equipped with an SBX daugtherboard at a sample rate of F_{samp} = 5MS/s. The collection receiver and its configuration remained fixed throughout all trials.

The signal generated by the X310-SDR was captured over-the-air with an antenna separation of $N_{distance}$ = 8cm. The software controlling the SDR was configured to play samples stored in a binary file. The software configuration remained fixed throughout for all SDR transmissions. The transmitter gain was adjusted via software to obtain a SNR of 55dB. This SNR was computed by taking the ratio of two measurements: the average power of the signal plus noise and the average power collected without any signal present (noise).

RF Fingerprints Classification for X310-SDR.

$N_{fprints}$ = 500 fingerprints per device were used to compute the fingerprint classification performance. The fingerprints used to determine classification performance were different than the fingerprints used to generate the MDA/ML model. Signal quality was gradually degraded from a SNR=35dB down to a SNR= 0dB using 2dB decrements. This signal degradation was accomplished by adding a scaled PRNG to the signal of interest. The scaled PRNG was used to represent the presence of AWGN in the communication channel.

Each step of the different curves shown in Figure 14 was computed using Nz=2 noise realizations per fingerprint for a total of $N_{fprints}$ = 1000 fingerprints per device. The resultant fingerprints were classified one at a time. Each fingerprint projection was compared to the mean of four devices' fingerprint projections in the MDA/ML model. The unknown fingerprint was projected into a Fisher space. The Euclidian distance was computed for the unknown fingerprint and the mean of every known

device. The classification algorithm associated the unknown fingerprint with the device with the shortest Euclidian distance.

The performance of the classification algorithm varies depending on the SNR of the signal as illustrated in Figure 14. The system correctly classify nearly %C=100% of the fingerprints when the SNR is greater than 35dB, and misclassify %C=50% of the fingerprints on average when the SNR equals 3dB. The projections for $N_{\text{devices}} = 3$ have the most compact cluster with maximum separation with respect to other devices. This separation allows the system to correctly classify device one fingerprints %C=70% of the time on average with a SNR of 3dB. However, the system correctly classifies about %C=30% of the fingerprints for device 1 when the SNR equals 3dB due to the widespread of the projections for this device.

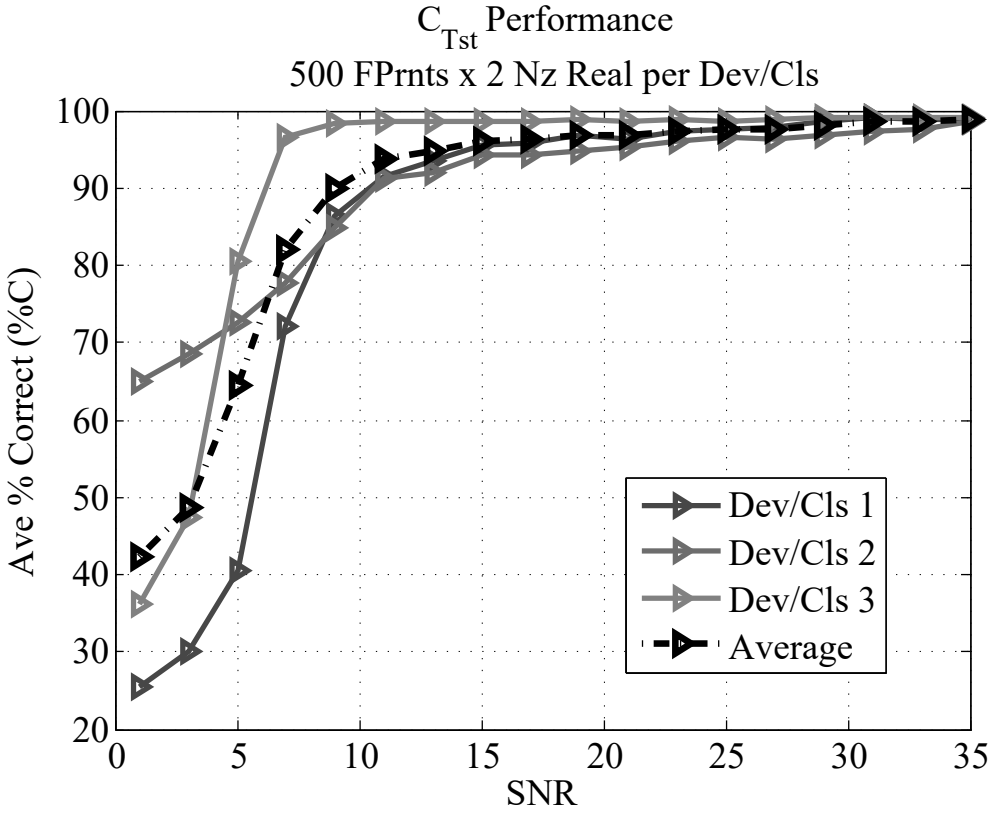

Figure 14. Fingerprint Classification Performance for Three X310-SDRs

The MDA/ML algorithm was used to project the RF fingerprints onto a 2D sub-

space as shown in Figure 13. $N_{\text{fprints}} = 500$ fingerprints per device were used to develop the 2D model. AWGN was used to create Nz=2 realizations per fingerprint, for a total of $N_{\text{fprints}} = 1000$ fingerprints per device. The quality of the signal was degraded from a SNR=55dB down to a SNR=35 in order to simulate transmitter normal operating conditions. The signal degradation was accomplished by adding a scaled PRNG to the signal of interest.

RF Fingerprints Verification for X310-SDR.

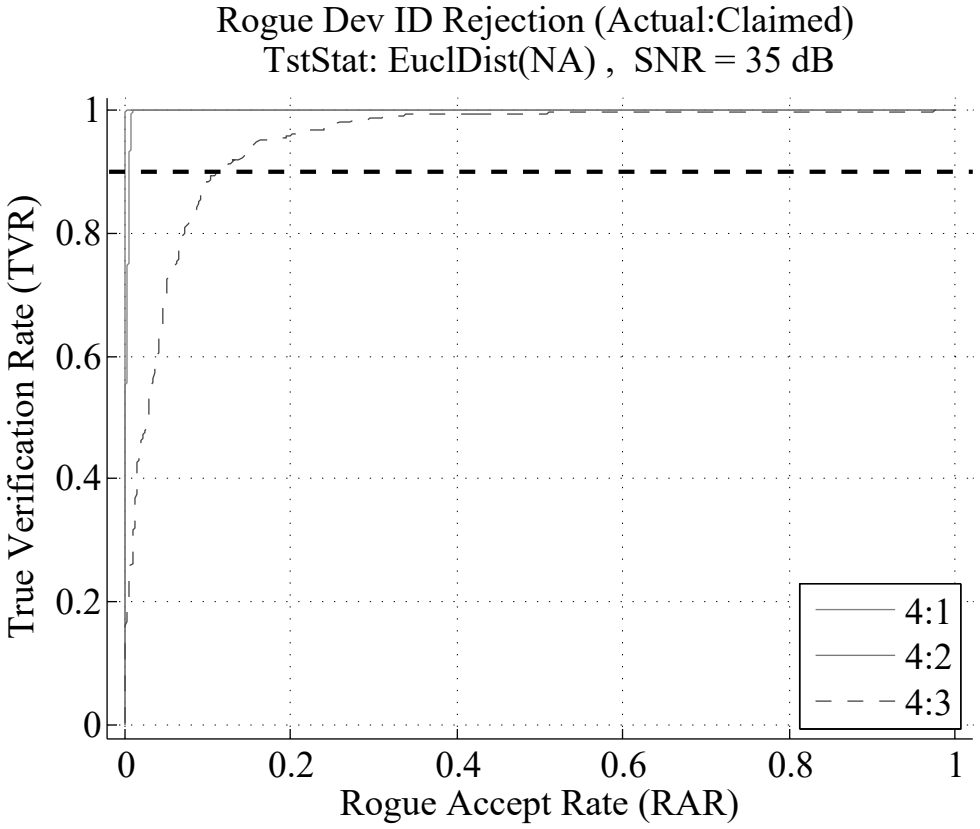

Figure 15. Fingerprints Verification Performance for Rogue X310-SDR

This test includes the $N_{\text{devices}} = 3$ X310-SDR devices that were used for the MDA/ML classification model and introduces a new device that has not been profiled before. Device 1 , 2, 3 are the original X310 devices used to generate the MDA-ML model. The rogue device (device-4) is a X310-SDR with a CBX daughterboard

and a GPSDO. All devices are using identical software configuration, and they are transmitting identical digital samples. Table 3 shows the relationship among devices.

Table 3. Device Configuration for RF Fingerprint Verification

Device	Color Code
X310-SBX-1	Blue
X310-SBX-2	Green
X310-CBX-GPSDO-1	Red
X310-CBX-GPSDO-2	Dev 4

The configuration of devices is so similar that it is very difficult to classify and verify the identity of the devices at low SNR. A very high SNR=35dB was used in order to obtain acceptable results. The system classifies fingerprints with a SNR$\geq$ 35dB for device-1 and device-2 nearly perfect (TVR $>$ 98% and FVR $<$ 2%). Device-3 yields a TVR$\geq$ 90% with a corresponding FVR$\leq$ 10%. The difference of verification performance among devices is explained by the configuration of the devices. The phase and frequency features for device-3 and the rogue device are very simmilar because both devices are configured with a GPSDO. Figure 15 illustrates the performance of the system verifying the identity of devices. The measure of similarity used for this system was Euclidian distance.

2.5 Conclusions and Future Research Recommendations

Current Radio Frequency (RF) communication systems are limited because they need to operate in spectrally dense environments. Cognitive Radio (CR) systems are a new area of research that focus on maximizing the performance of communication systems in spectrally dense environments by dynamically adjusting transmitter and receiver parameters to operate in under-utilized areas of the spectrum.

One of the goals of CR systems is to avoid interference with the primary (licensed) users of the spectrum. This goal can be accomplished by avoiding transmissions in

the areas of the spectrum currently utilized by the primary user. Spectrum sensing is needed in CR systems to provide information about the surrounding radio spectrum and to be able to detect the presence of the primary user. Current CR research efforts are focused on the development of new mechanisms to detect primary user (PU) or improve existing ones. However, previous researchers have identified that a Primary User (PU) emulation attack can disrupt the operation of a cognitive radio system by significantly reducing the spectrum available to secondary (unlicensed) users.

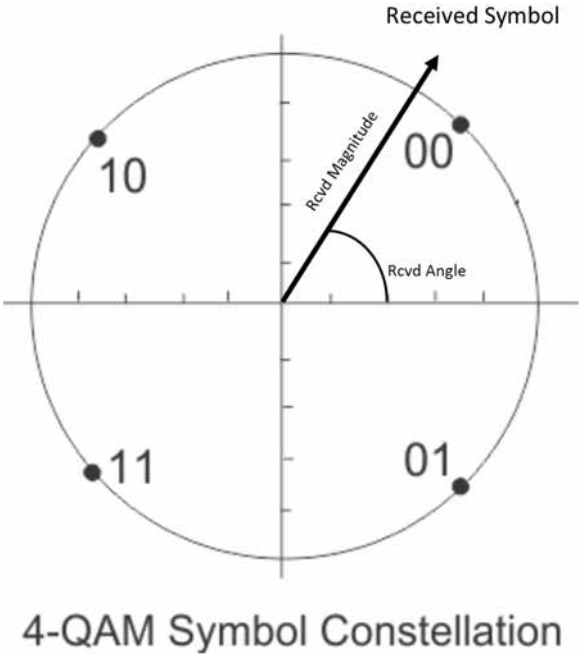

Figure 16. IQ Channel Deviation for 4QAM Constellation Projection

This paper describes an algorithm that detects a primary user emulation attack using Radio Frequency Distinct Native Attributes (RF-DNA) fingerprinting techniques. Several tests were conducted to characterize the performance of the algorithm.

Test results demonstrated that the proposed solution can detect a Software-Defined Radio (SDR) replaying the signal of a primary user. Even under a relatively low Signal to Noise Ratio (SNR) 5dB, the true verification rate of the primary user exceeds 90%, while the false verification rate of the replay was less than 6%. These

experiments consider the most challenging scenario case by classifying devices from the same manufacturer and model number. Results are expected to improve in cases where the devices are from different manufacturers.

Future research includes the generation of fingerprints based on the unavoidable In-Phase/Quadrature-Phase (I/Q) channel deviations generated by the transmitter while they emit communication symbols. The I/Q channel deviations from the ideal symbol are illustrated in Figure 16. Features that have the potential to discriminate devices can be obtained by computing the variance (σ^2), skewness (γ) and kurtosis (κ) for the symbol magnitude $a_c(n)$, and the phase angle $\phi_c(n)$ between the in-phase and quadrature phase axes.

This research highlighted that it is possible to provide reliable discrimination of devices through Radio Frequency Distinct Native Attribute (RF-DNA) fingerprinting techniques using relatively inexpensive ($\sim$\$7,000) equipment such as the National Instruments (NI) X310 Software Defined Radio (SDR). The ability to verify the true source of an RF emission can be used to prevent a Primary User Emulation Attack (PUEA).

III. Detection of Primary User Emulation Attack Using Constellation-Based Distinct Native Attribute Techniques

3.1 Introduction

Cognitive-Radio refers to a new development of intelligent radio communication systems that are aware of their environment, and adjust their transmitter and receiver parameters in order to maximize spectrum efficiency while maintaining the ability of a highly reliable communication system. Understanding current and future spectrum usage is one of the most difficult problems in the design and implementation of Cognitive Radios (CRs). Detection and classification of signals is a critical design problem in cognitive radios in order to detect the presence of Primary User (PU) (licensed) of the spectrum. Current CR spectrum sensing research efforts tend to focus on developing new mechanisms to detect PU presence or improving existing ones [26]. However, previous researchers have identified that a Primary User Emulation Attack (PUEA) can disrupt the operation of a cognitive radio system by significantly reducing the spectrum available to secondary (unlicensed) users [10, 11, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36].

Traditional security techniques for preventing PUEAs are based on identifying the location of the transmission source in order to compare it to the known PU location [10, 28]. While these techniques are effective to some degree, security schemes that are geolocation based are increasingly difficult to implement as they require obtaining measurements from several different sensors that are widely spaced around the PU location. Additionally, geolocation based algorithms do not work well when the PU is a mobile node. Recent research demonstrates that the analysis of signals at the Physical Layer (PHY) layer can be used to thwart PUEAs [31, 37, 38, 19, 15, 39, 40, 12]. This paper describes an innovative algorithm that detects PUEAs using

Constellation-Based Distinct Native Attribute (CB-DNA) fingerprinting techniques.

The proposed PU verification system relies on examining waveforms at the PHY layer to uniquely identify devices based on inherent differences in their transmissions. This verification scheme requires prior signal collection of PU's transmissions. Every device that emits Radio Frequency (RF) signals has unique characteristics that are very hard to duplicate, making these features useful to uniquely identify transmitters. These characteristics are observed as transient behavior with respect to the instantaneous amplitude, phase, and frequency of the radiated signal. This behavior can be caused due to a variety of reasons such as precision of frequency synthesis systems, modulator subsystems, and RF amplifiers. Unique transient signals can be observed even among transmitters of the same type and model due to manufacturing tolerances and aging of used components [20]. These transmitter anomalies can be used to create Radio Frequency Distinct Native Attribute (RF-DNA) fingerprints.

3.2 Background

This section provides the technical background supporting the methodology described in section 3.3. The topics covered in the section include: generation of Time Domain (TD) Radio Frequency Distinct Native Attribute (RF-DNA) fingerprints, generation of Spectral Domain (SD) RF-DNA fingerprints, generation of Constellation-Based Distinct Native Attribute (CB-DNA) fingerprints, and classification of signals using Multiple Discriminant Analysis / Maximum Likelihood (MDA/ML).

Time Domain RF-DNA Fingerprinting.

RF-DNA fingerprints can be generated by passively collecting signals generated by Modulator/Demodulators (MODEMs) as they transmit communication symbols. The collected signal can be represented in the TD as the complex vector $\mathbf{x}[n] =$

$s_I(n) + js_Q(n)$ for $n = \{0, 1, 2, ..., N-1\}$, where n specifies the time when the sample was measured and the variable N specifies the total number of samples stored in the vector. The instantaneous amplitude, phase, and frequency of $\mathbf{x}$ can be computed as follows [41]:

$$\mathbf{a}(n) = \sqrt{(s_I(n) + js_Q(n))^2}, \qquad n = \{0, 1, 2, ..., N-1\}, \tag{22}$$

$$\boldsymbol{\phi}(n) = \tan^{-1}\left[\frac{s_Q(n)}{s_I(n)}\right], \quad s_I(n) \neq 0, \quad n = \{0, 1, 2, ..., N-1\}, \tag{23}$$

$$\mathbf{f}(n) = \frac{1}{2\pi}\left[\frac{d\phi(n)}{dn}\right] \qquad n = \{0, 1, 2, ..., N-1\}. \tag{24}$$

The quality of RF-DNA fingerprints generated using instantaneous amplitude, phase, and frequency can be improved by normalizing the range and subtracting their respective means in order to remove any existing bias. Bias removal and signal normalization can be obtained by:

$$\mathbf{a}_c(n) = \frac{\mathbf{a}(n) - \mu_a}{\max(\mathbf{a}_c(n))}, \tag{25}$$

$$\boldsymbol{\phi}_c(n) = \frac{\boldsymbol{\phi}(n) - \mu_\phi}{\max(\boldsymbol{\phi}_c(n))}, \tag{26}$$

$$\mathbf{f}_c(n) = \frac{\mathbf{f}(n) - \mu_f}{\max(\mathbf{f}_c(n))}, \tag{27}$$

where μ_a, μ_ϕ, μ_f, are the respective amplitude, phase, and frequency means [41].

RF-DNA fingerprints are obtained by dividing the sequences $a_c(n)$, $\phi_c(n)$, $f_c(n)$, into R equal length sequences. The distinct fingerprints are generated by computing

the standard deviation (σ), variance (σ^2), skewness (γ), and kurtosis (κ) of these sequences to create new vectors as follows:

$$\mathbf{F}_r^a = [\sigma_a, \sigma_a^2, \gamma_a, \kappa_a], \tag{28}$$

$$\mathbf{F}_r^\phi = [\sigma_\phi, \sigma_\phi^2, \gamma_\phi, \kappa_\phi], \tag{29}$$

$$\mathbf{F}_r^f = [\sigma_f, \sigma_f^2, \gamma_f, \kappa_f]. \tag{30}$$

The composite fingerprint is generated by concatenating the individual $\mathbf{F}^\sigma$ sequences, where σ denotes a specific amplitude, phase or frequency sequence by

$$\mathbf{F}^\sigma = \left[\mathbf{F}_1^\sigma \quad \vdots \quad \mathbf{F}_2^\sigma \quad \cdots \quad \mathbf{F}_R^\sigma\right]. \tag{31}$$

The composite amplitude, phase, and frequency fingerprints can be combined in order to generate a complete TD fingerprint as follows:

$$\mathbf{F}_{TD} = \left[\mathbf{F}^a \quad \vdots \quad \mathbf{F}^\phi \quad \vdots \quad \mathbf{F}^f\right] \tag{32}$$

A visualization depiction of the generated Radio Frequency (RF) fingerprints is shown in Figure 17. The figure shows the RF fingerprints for 4 different devices. The values for the variance, skewness, and kurtosis of the signal generated by the devices is shown in the horizontal bands. The colors represent the average value for each statistical measurement scaled to span 0 to 1.

Spectral Domain RF-DNA Fingerprinting.

SD RF-DNA fingerprints are generated using the Power Spectral Density (PSD) of the TD signal represented in vector $\mathbf{x}$. The SD representation of $\mathbf{x}$ can be computed using the Discrete Fourier Transform (DFT). The mathematical model to compute

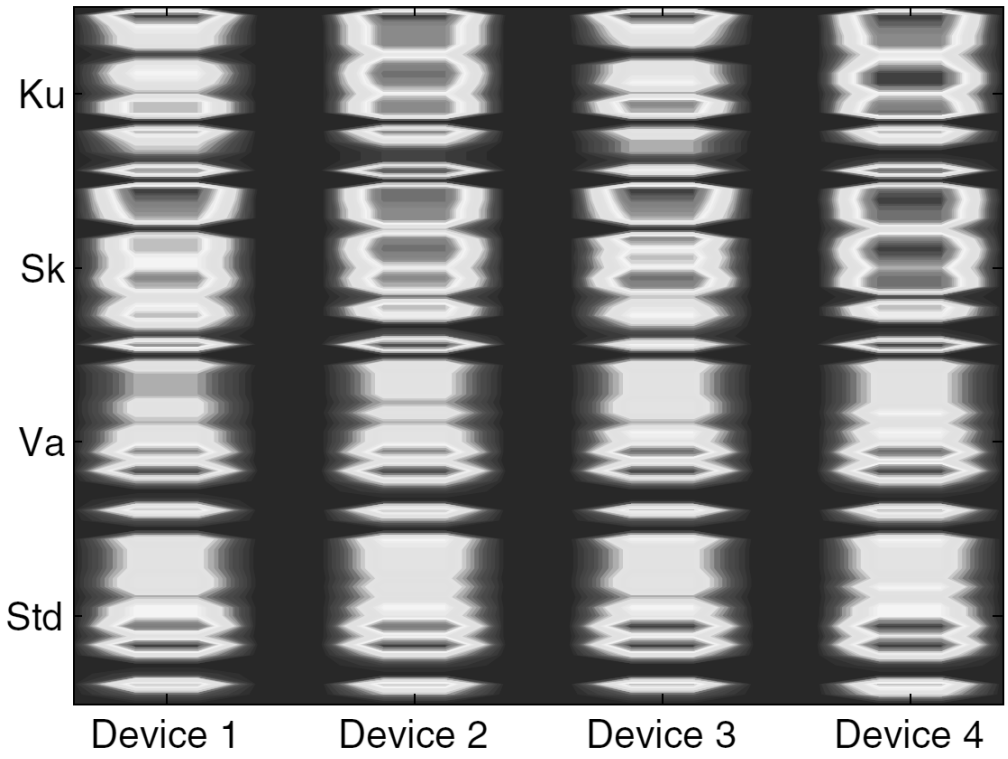

Figure 17. Visualization for RF-DNA Fingerprints for 4 Devices [5]

the DFT is as follows:

$$\mathbf{X}(k) = \frac{1}{N}\sum_{n=0}^{N-1}\mathbf{x}(n)e^{\frac{-j2\pi kn}{N}} \quad \text{for } k = \{0, 1, 2, ..., N-1\} \tag{33}$$

In this mathematical model $\mathbf{X}(k)$ is a complex number that represents the frequency component of a signal at band k, while $\mathbf{x}(n)$ represents the signal as it is being sampled in the time domain [21]. The PSD of the signal is normalized with respect to power in order to mitigate collection effects that may affect signal classification [3]. The average power of the signal is computed by:

$$P_{\mathbf{X}} = \frac{1}{N}\sum_{n=0}^{N-1}\mathbf{X}(n)\mathbf{X}(n)^*, \tag{34}$$

and the normalized-power PSD sequence is obtained by:

$$\overline{\mathbf{X}}(k) = \frac{1}{P_{\mathbf{X}}} \left| \mathbf{X}(k) \right|^2 . \tag{35}$$

Once the normalized PSD signal is obtained, the SD fingerprints are generated by dividing the sequence into R equal length sequences. The distinct fingerprints are generated by computing the standard deviation (σ), variance (σ^2), skewness (γ), and kurtosis (κ) of these sequences to create new vectors as follows:

$$\mathbf{F}_r = [\sigma, \sigma^2, \gamma, \kappa]. \tag{36}$$

The composite fingerprint is generated by concatenating the individual $\mathbf{F}$ sequences by:

$$\mathbf{F} = \left[\mathbf{F}_1 \quad \vdots \quad \mathbf{F}_2 \quad \cdots \quad \mathbf{F}_R \right]. \tag{37}$$

The resultant full-dimensional fingerprint vector $\mathbf{F}$ from (37) contains a total of N_f = (# of Features)×(# of Statistical Metrics)×(# of Regions) elements. This vector is illustrated in Figure 18.

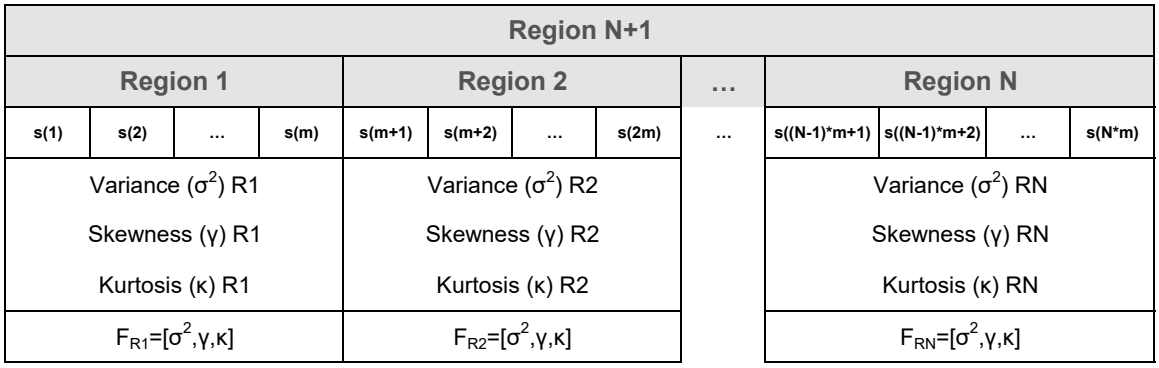

Figure 18. RF-DNA Statistical Fingerprint Generation for Centered and Normalized Feature Sequences and $N+1$ Total Subregions

Constellation-Based RF Fingerprinting.

RF-DNA fingerprints are generated using synchronization parameters (preambles, postambles, midambles, pilot tones, etc) of the protocol used by the Primary User (PU). Primary User Emulation Attacks (PUEAs) need to mimic the protocol used by the PU in order to fool secondary users. The forged transmissions need to include the synchronization parameters of the protocol used by the PU, enabling the verification of the source of the signal using RF-DNA fingerprinting. RF-DNA generates features based on the portions that remain constant in the Signal of Interest (SOI). In contrast, CB-DNA uses the entire SOI by generating features from the projections of communication symbols.

It is possible to extract unique features from a transmitter that is operating in a steady state condition using Constellation-Based Distinct Native Attribute (CB-DNA) [6, 42, 43] . A constellation projection is computed using a linear transformation, which projects each received symbol as a single point in the I/Q plane. A given modulation scheme will have an ideal location for each symbol in the alphabet, which will maximize the performance of the communication link [44]. The projection of received symbols collected over-the-air will have unintended and unavoidable deviations compared to ideal symbol locations due to variability in the receiver and transmitter's hardware. These imperfections are introduced by: component tolerances, oscillators' phase noise, spurious tones from mixers and power amplifiers, manufacturing processes, etc [42].

Transmitter's modulated signals plus imperfections can be modeled as follows:

$$\begin{aligned} Z(t) =& I(t)\cos(2\pi f_c t + \phi/2) \\ &+ Q(t)\sin(2\pi f_c t - \phi/2), \end{aligned} \tag{38}$$

where $Z(t)$ represents the TD transmitted signal, $I(t)$ the in-phase component of the signal, $Q(t)$ the quadrature phase component of the signal, f_c the intermediate carrier frequency, and ϕ the quadrature error induced by the transmitter's components [42]. The individual $I(t)$ and $Q(t)$ bands can be modeled as follows:

$$I(t) = G_{I/Q} \sum_{k=-\infty}^{\infty} I_k(t - kT_s - \tau) + O_I(t), \tag{39}$$

$$Q(t) = \sum_{k=-\infty}^{\infty} Q_k(t - kT_s - \tau - \tau_D) + O_Q(t), \tag{40}$$

where $G_{I/Q}$ is the I/Q gain imbalance, I_k and Q_k represent the modulated symbols in their respective I and Q bands, τ_D is the time delay between the I and Q channels, $O_I(t)$ and $Q_I(t)$ represents the I/Q offsets, and T_s is the symbol period [42]. The imperfections $G_{I/Q}$, τ_D, $O_I(t)$, and $O_Q(t)$ are generated by the transmitter's hardware components and are unique for each transmitter. The projection of $I(t)$ and $Q(t)$ in the constellation space will deviate from the ideal symbol locations due to imperfections described by (39) and (40).

Prior research shows that constellation projection deviations reflect a bias that is conditional to the previous symbol transmitted, and the next symbol to be transmitted. If the projections are grouped based on prior estimated symbol, current estimated symbol, and next estimated symbol, then these bias show as clusters in the I/Q plane. Figure 19 illustrates this phenomenon by color coding constellation points according to the symbol values preceding and succeeding the symbol being estimated, i.e., [0 X 0], [0 X 1], [1 X 0], and [1 X 1], where X denotes the symbol being estimated. The clusters formed by applying conditional constellation are caused by the

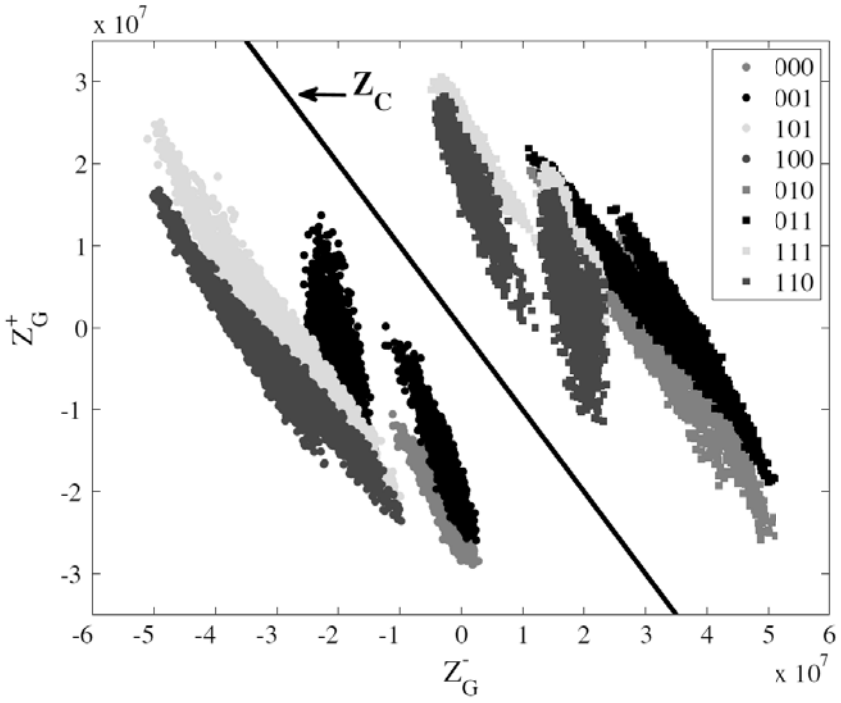

Figure 19. Binary Constellation for Unintentional Ethernet Cable Emissions Symbol Estimation Showing Non-Gaussian Multimodal Symbol Sub-Clusters and Linear Bit Estimation Boundary (Z_C). [6]

transmitter's hardware components and can be used to uniquely identify the source of the RF emanations [6].

Multiple Discriminant Analysis/Maximum Likelihood.

The purpose of RF fingerprints is to extract features from signals so that they can be classified. Classification of RF fingerprints requires additional processing because they can generate a multivariate statistical model with hundreds of independent variables. Obtaining a Maximum Likelihood Estimate (MLE) of the source of a RF emanation can be computationally intensive due to the high dimensionality of the statistics. This problem can be simplified using MDA algorithm.

MDA is a multivariate statistical technique to apply linear discriminant analysis [22]. The objective of MDA is to classify objects into two or more mutually exclusive classes by reducing the dimensionality of a set of independent variables. The dimen-

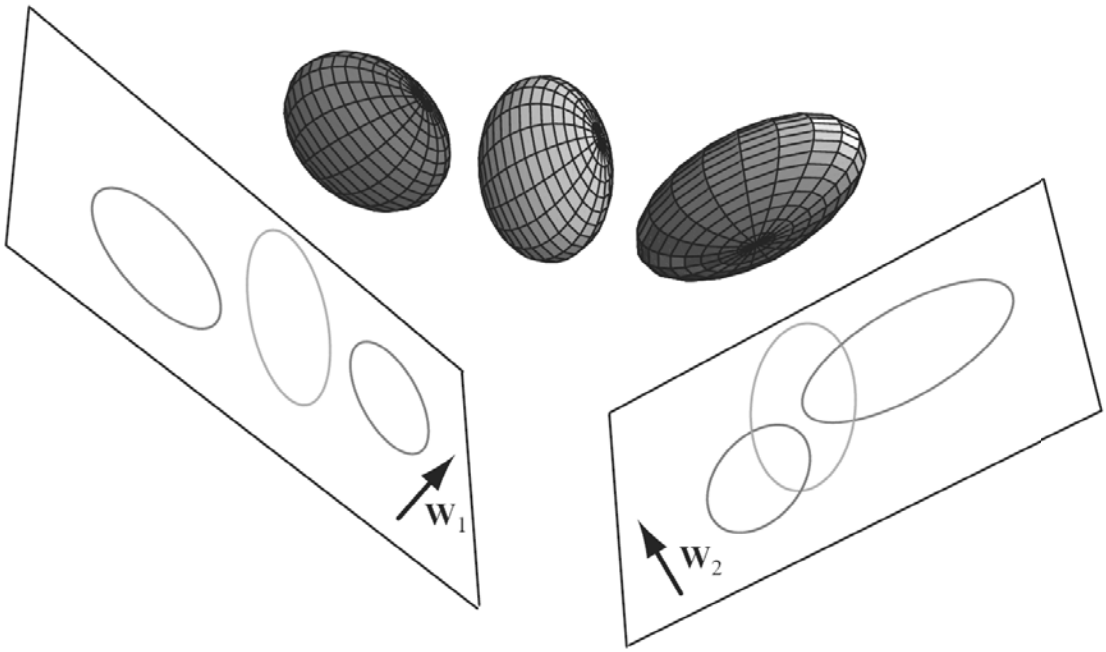

Figure 20. MDA Projection of 3D Space into 2D Space [7]

sionality reduction is accomplished by identifying the smallest linear combination of variables with normal errors that best discriminate between classes [23]. For example, the 3D model shown in Figure 20 is projected onto 2D models in order to reduce the dimensionality of the problem. The 2D projections are defined by the norm vectors **W**1 and **W**2 respectively. It is significantly more difficult to classify and discriminate the **W**2 projections because the projections overlap. However, the **W**1 subspace facilitates classification and discrimination because the projections do not overlap. The MDA algorithm aims to find projections such as those provided by the **W**1 vector.

The MDA algorithm starts by defining two scatter matrices, the inter-class matrix (S_b) and the intra-class matrix (S_w) of the dataset x. The MDA projection maximizes inter-class distances while minimizing intra-class spread. These matrices are defined by [7]:

$$\mathbf{S}_b = \sum_{i=1}^{N_c} P_i \sum i, \tag{41}$$

$$\mathbf{S}_B = \sum_{i=1}^{N_c} P_i(\mu_i - \mu)(\mu_i - \mu)^T, \tag{42}$$

Where N_c is the number of classes, P_i is the prior probability of class c_i and $\sum_i$ is the covariance matrix. Using the two scatter matrices, the projection matrix $\mathbf{W}$ is formed using the eigenvectors of $\mathbf{S}_w^{-1}\mathbf{S}_b$. The multivariate statistics can be projected into a $(N_c - 1)$ dimensional subspace by [7]:

$$\mathbf{F}_i^{\mathbf{W}} = \mathbf{W}^T\mathbf{F}, \tag{43}$$

where $\mathbf{F}$ is the matrix representing the fingerprint.

3.3 Methodology

This section outlines the methodology used to determine the applicability of the Constellation-Based Distinct Native Attribute (CB-DNA) concept to detect the presence of a Primary User Emulation Attack (PUEA). Additionally, this section outlines the goals and hypotheses of this research, elaborates on the problem, and describes the measures of merit on which the results of the algorithm will be judged. An outline of the experiments to be performed as well as the hardware and software configuration is given. The expected results are given and the expected performance factors are stated.

Research Objectives.

Wireless communication systems are susceptible to a myriad of attacks because the transmission medium is hard to constrain to specific locations, making it accessible to unauthorized users. This research aims to characterize a security mechanism that operates at the Physical Layer (PHY) layer in order to detect a PUEA. The proposed

solution generates unique CB-DNA fingerprints that can be used to authenticate the Primary User (PU).

The objective of this research is to develop an algorithm that capitalizes on the unavoidable spurious signals emitted by transmitters, as they try to radiate communication symbols, by generating CB-DNA fingerprints that uniquely identify the transmitter. The algorithm requires prior collections of the PU Radio Frequency (RF) emanations in order to generate a Multiple Discriminant Analysis / Maximum Likelihood (MDA/ML) classification model that will be used to discriminate unidentified signals.

Research Hypotheses.

There are three hypotheses that will be considered throughout this research:

- CB-DNA fingerprints can be used to uniquely identify the source of a transmission.
- CB-DNA fingerprints can provide a better RF source discrimination performance than Radio Frequency Distinct Native Attribute (RF-DNA) fingerprints.
- The average correct classification rate of devices will exceed %C=90% for like-model devices, passband device discrimination, and baseband device discrimination.

Measure of Merit.

The measure of merit of this algorithm is its ability to persistently perform cross-device discrimination, and more specifically like-model discrimination. Like-model discrimination presents a greater classification challenge because the devices use identical components, assembly line procedures, quality assurance standards, etc. The

measure of merit will be quantified as $\%C$: the average percentage of correct classification.

Quadrature Phase Shift Keying Transmitter Design.

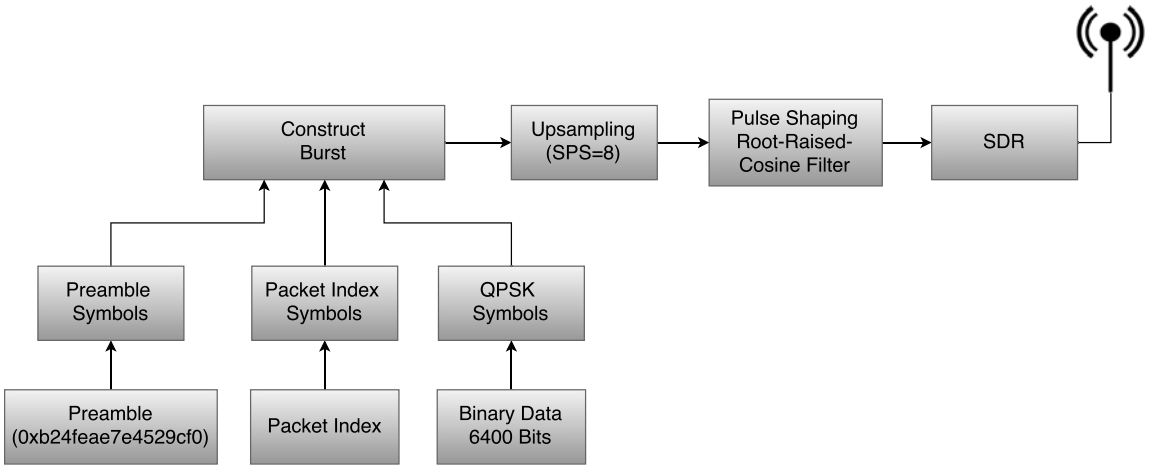

Figure 21. Block Diagram for Burst-Mode QPSK Transmitter Implementation

A QPSK modulated signal was developed to serve as a proof of concept as there are currently no standardized CR systems. The signal is constructed from a data packet that consists of three fields: $\mathrm{P}_{\mathrm{length}} = 64$ bits training sequence, $\mathrm{Pid}_{\mathrm{length}} = 16$ bits packet index, and $\mathrm{Pload}_{\mathrm{length}} = 6400$ bits payload.

The training sequence serves as a preamble, and it is used to aid the receiver during the synchronization process. This $P_{\mathrm{length}} = 64$ bits sequence has very good periodic autocorrelation properties [45], which enables the receiver to detect burst presence, estimate symbol boundaries, and estimate phase offset between the transmitter and receiver. The autocorrelation function of this binary sequence is shown in Figure 22. The $Pid_{len} = 16$ bits packet index field is used to identify the specific packet transmitted, in order to conduct Bit Error Rate (BER) computations. Finally, the $Pload_{len} = 6400$ bits payload is used to represent the data to be transmitted and is populated with a sequence obtained from a Pseudo Random Number Generator

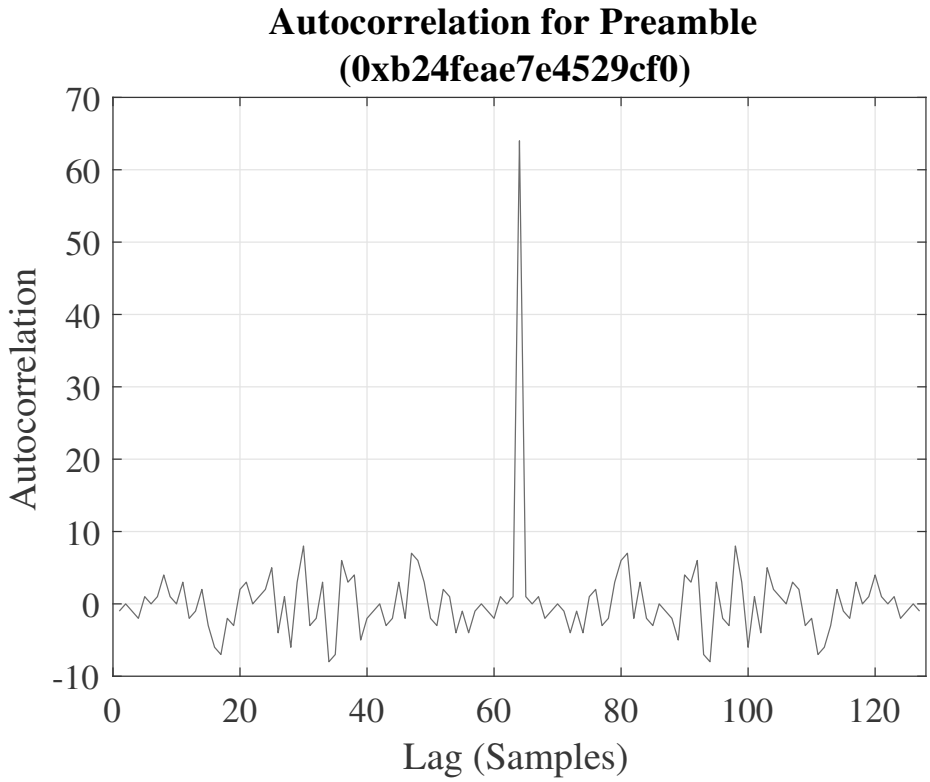

Figure 22. Autocorrelation Function for the Preamble Sequence

(PRNG).

The QPSK transmitter implemented for this research takes the preamble, packet index and data payload as inputs, and converts them into QPSK symbols. Following this conversion, these communication symbols are upsampled by a factor of $sps = 8$ by inserting seven zeros in between each symbol. Finally, a pulse-shaping root-raised-cosine Nyquist filter is applied to the signal in order to minimize Intersymbol Interference (ISI) and interpolate the samples in between symbols. The implementation of this QPSK transmitter is illustrated in Figure 21. The resultant signal generated by the transmitter has a bandwidth of $\mathrm{Tx}_{\mathrm{bandwidth}} = 1$ MHz as shown in Figure 23.

The impulse response of the pulse shaping filter is shown in Figure 24. This filter implementation minimizes ISI because the only non-zero component is the symbol

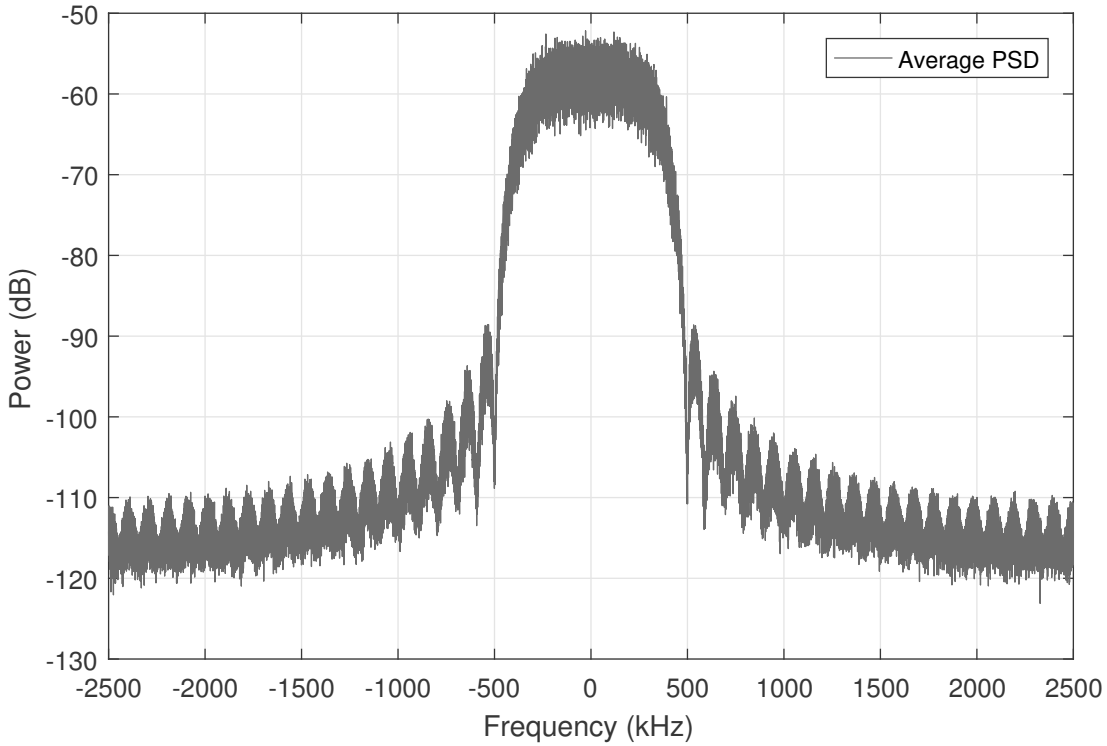

Figure 23. PSD of Baseband QPSK Signal Computed Using Welch's Overlapped Segment Averaging Estimator, Sample Rate $F_{\mathbf{samp}}$=5 MS/s

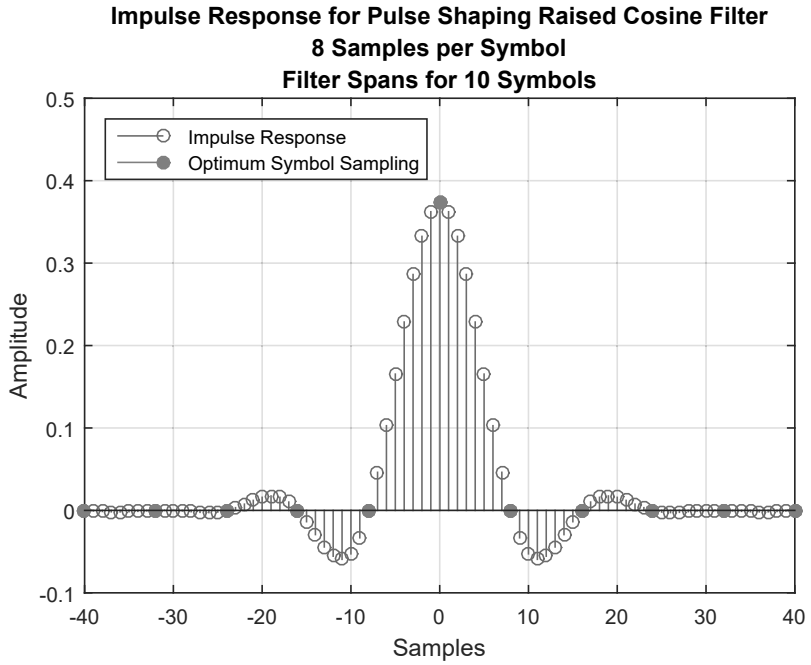

Figure 24. Root Raised Cosine Filter Impulse Response, sps=8 Samples per Symbol, Filter Spans for $FSpan$=10 Symbols Showing Optimum Symbol Sampling

that is currently being sampled.

This QPSK transmitter was implemented in MATLAB® and the resulting discrete waveform samples were stored in a binary file. The file contained $NBursts = 1000$ individual bursts with a random number of zeros in between each burst. The zeros in between bursts are used to disable the transmitter, so that the system operates in burst-mode. GNU-Radio was configured to open the binary file and send those samples to National Instruments (NI) Universal Software Radio Peripheral (USRP) X310 and the BladeRF respectively.

Software-Defined Radio Receiver Configuration.

The collection receiver used in this research was a NI USRP X310 Software-Defined Radio (SDR). This research departs from the norm by using a relatively inexpensive RF transmitter and receiver. Research of RF-DNA fingerprinting is normally conducted using very precise and accurate collection receivers equipped with high quality expensive analog components especially designed for sensitive measurements in order to minimize receiver coloration effects [31, 46, 47]. The X310 SDR is available Commercial Off-The-Shelf (COTS) with a retail price of approximately $7,000. In addition to its price tag, this RF transmitter/receiver was chosen for this research because it has a very capable Field Programmable Gate Array (FPGA) that can be used for signal processing.

GNU Radio was used as the controlling software for all RF transmissions and signal collections. The transmissions were preprocessed using MATLAB® and stored in a file. GNU Radio was configured to read the preprocessed file and play the samples through the SDR platform. Signal collection was accomplished by configuring GNU Radio to store the collected samples in a file. Signal postprocessing was accomplished using MATLAB®.

Quadrature Phase Shift Keying Receiver Design.

The main objective of this research is to assess the performance of a device discrimination algorithm based on CB-DNA fingerprints. A burst-mode QPSK receiver was implemented to project the received symbols in constellation space. The constellation points obtained from this receiver were used to generate CB-DNA based fingerprints. Figure 25 illustrates the burst-mode QPSK receiver implemented in this project.

The choice of implementation for the burst detector, carrier frequency recovery, and phase recovery components can significantly affect the resulting constellation projection. The respective implementations for these components are detailed in this document.

Burst Detector.

Burst detection is normally implemented using an energy detection algorithm. Using this scheme, the beginning of a burst is detected by computing when the input signal power exceeds a specified threshold. However, this research cross-correlates the received signal with the known preamble sequence to detect the presence of a burst. Using this technique it is possible to estimate symbol boundary, since the peak of the cross-correlation aligns with the beginning of the preamble. This technique only works

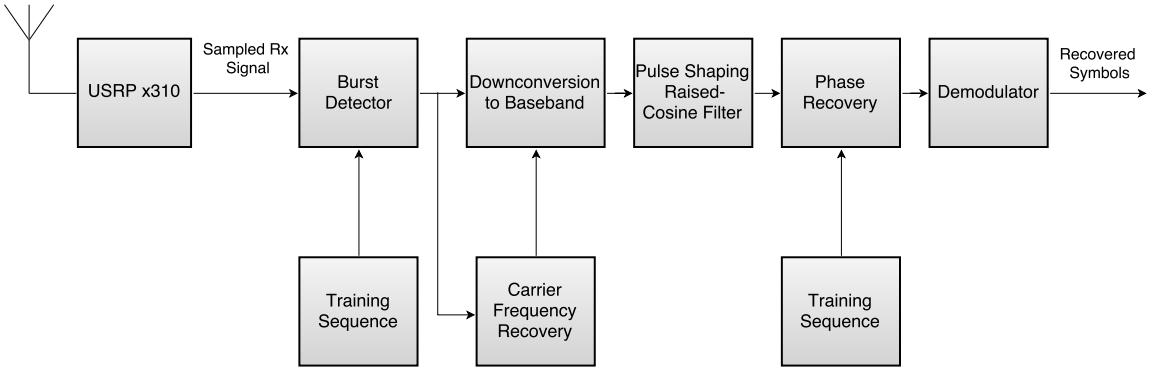

Figure 25. Block Diagram for Burst-Mode QPSK Receiver Implementation

when the preamble has very good correlation properties, and the center frequency offset between the transmitter and receiver is relatively small.

Intermediate Carrier Frequency Recovery.

Communication systems implemented using Phase Shift Keying (PSK) modulation have zero average energy transmitted at the carrier frequency [44].

A QPSK signal sampled at the output of the receiver's matched filter can be modeled as the complex vector:

$$\boldsymbol{R}(n) = S\mathbf{a}(n)\exp(j2\pi f_c t) + \boldsymbol{\omega}(n), \quad \text{for } n = 1, ..., N \tag{44}$$

where S is a real scalar, $\mathbf{a}(n)$ is the transmitted QPSK symbols of unit magnitude, f_c is the carrier frequency, and $\boldsymbol{\omega}(n)$ represents the noise in the communication channel [48].

The carrier frequency of a M-PSK signal can be estimated by raising the sampled M-PSK signal to the M power in order to remove the modulation. Raising the signal to the M power creates a significant tone at M times the carrier frequency, revealing the suppressed carrier [49]. In the specific case of QPSK the tone at four times the carrier frequency is evident in the following expression:

$$\begin{aligned} \boldsymbol{R}^4(n) =& S^4\mathbf{a}^4(n)\exp(j8\pi f_c t) + \\ & 4S^3\mathbf{a}^3(n)\exp(j6\pi f_c t)\boldsymbol{\omega}(n) + \\ & 6S^2\mathbf{a}^2(n)\exp(j4\pi f_c t)\boldsymbol{\omega}^2(n) + \\ & 4S\mathbf{a}(n)\exp(j2\pi f_c t)\boldsymbol{\omega}^3(n) + \boldsymbol{\omega}^4(n). \end{aligned} \tag{45}$$

This research estimated the intermediate carrier frequency in a burst-by-burst basis by computing $\hat{F}\text{Carr} = (\arg\max_n(|\mathcal{F}\{R^4(n)\}|)/4$. This technique produces

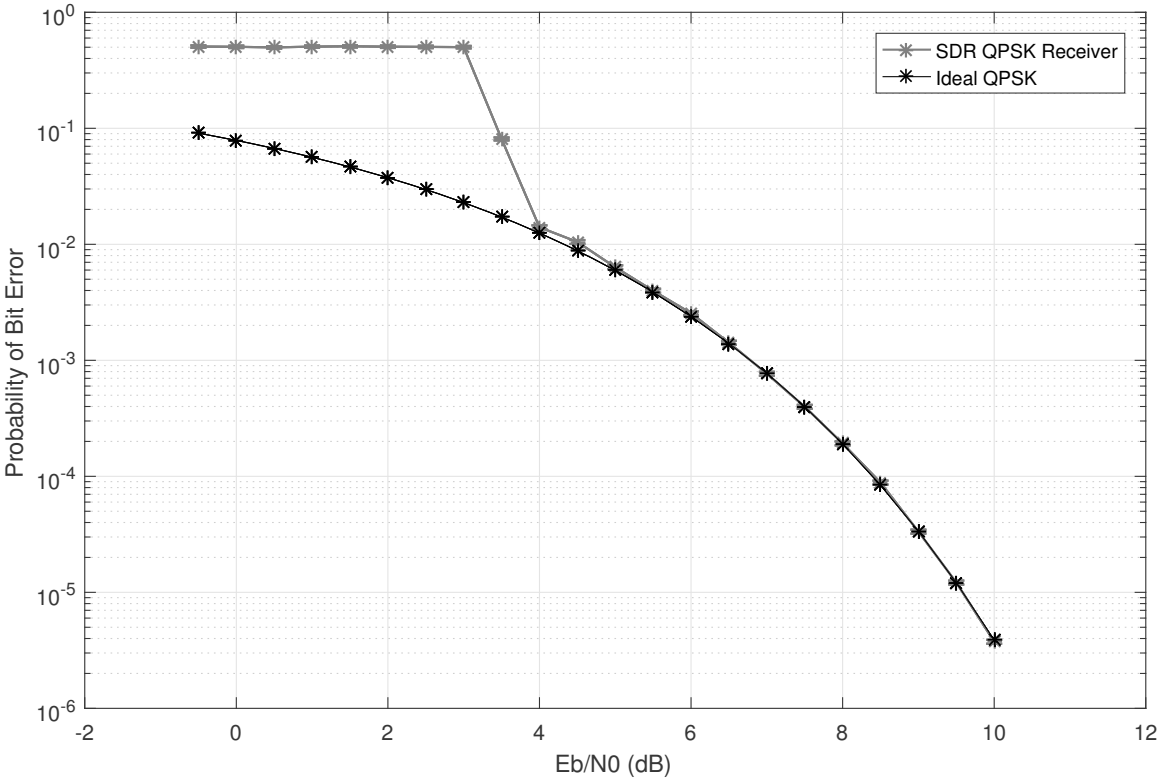

Figure 26. Probability of Bit Error vs $E_b/N0$ for SDR QPSK Receiver

reliable intermediate frequency estimates when the Signal to Noise Ratio (SNR) $E_b/N0$=4 dB. It is not possible to synchronize the receiver when the SNR$E_b/N0 \leq$ 4 dB because the intermediate frequency estimates obtained are unreliable as illustrated in Figure 26. These limitations in the computation of intermediate frequency estimates is consistent with the Cramer-Rao Lower Bound (CRLB) for QPSK signals [50, 51].

Each data point in Figure 26 was computed with at least $N_{\text{bitErrors}}$=2500 bit errors. This large number of trials reduced the mean error bars to within the vertical extent of the plotted data markers. Therefore, trial mean error bars are intentionally omitted to enhance visual clarity.

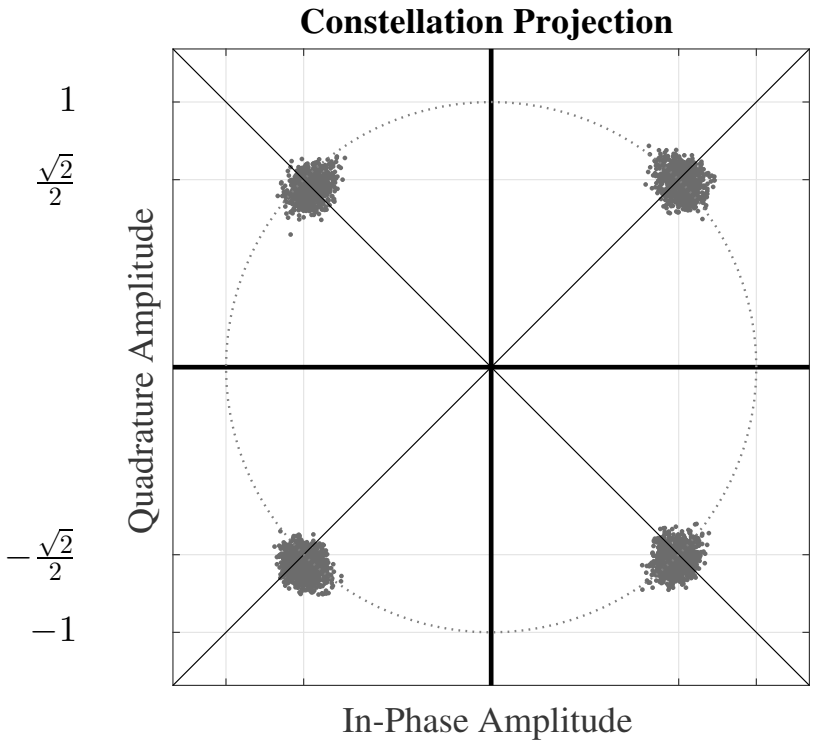

Figure 27. Derotated and Normalized Constellation Projection for One Received Burst with $Eb/N0$=20dB

Phase Recovery.

Typical implementations of QPSK receivers use a Phase-Locked Loop (PLL) to reconstruct the suppressed carrier. PLL algorithms use feedback to detect and compensate for phase errors [52]. The auto-compensation feature inherent in PLL algorithms could potentially hide some of the features used to uniquely identify a transmitter. Therefore, this research implements a phase detection algorithm that rotates the received constellation points from 0 radians to $\pi/2$ radians in $N = 100$ increments, and finds the phase angle that projects symbols closer to ideal locations. The pseudo-code for this algorithm is presented in Algorithm 1.

There are four different phase angle ambiguities after derotating the constellation. This research resolves these ambiguities by comparing the four possible phase angles with the known preamble. Finally, the constellation projection is normalized by

Algorithm 1 Phase Angle Estimator

Require: Received Constellation Projections(rxConstProj)
  rotationVariances $\leftarrow \infty$
  for $N = 1$ to 100 **do**
    $\theta \leftarrow \frac{N\pi}{2\times 100}$
    rotatedCProj $\leftarrow$ rxConstProj $\cdot e^{j\theta}$
    temp $\leftarrow$ |real(rotatedCProj)| + j |imaginary(rotatedCProj)|
    rotationVariances(N)$\leftarrow$ variance(temp)
  end for
  $N \leftarrow \arg\min_N$(rotationVariances)
  return rxConstProj $\cdot e^{\frac{jN\pi}{2\times 100}}$

scaling each constellation point as follows:

$$constPoint = \frac{constPoint}{mean\left(|rxConstProj|\right)}. \tag{46}$$

The derotated and normalized constellation projections for one burst is illustrated in Figure 27.

Experimental Signal Collection.

The experiments were conducted in the AFIT Cognitive Radio (ACRO) Laboratory located at the Air Force Institute of Technology (AFIT). The devices under test were inside a Ramsey STE6000 RF Shielded Test Enclosure. This test enclosure was designed for use with Industrial Scientific and Medical (ISM) band signals including Bluetooth, WiFi, and ZigBee. The STE6000 provides isolation greater than 90dB at the 2.4Ghz ISM band. Additionally, the interior has an RF absorbent foam liner that attenuates signal reflections within the test enclosure by more than 24dB. The STE6000 was equipped with Ethernet and USB connections in order to control the devices operating inside test enclosure while it was sealed.

The X310 SDR has transmit and receive capabilities covering from DC to 6.0 GHz depending on daughterboard installed. For this research, the CBX daughterboard

revision 3 serial number F59192 was installed in the collection receiver, providing a receive frequency range of 1200-6000 MHz with a maximum instantaneous bandwidth of 40MHz. The collection receiver was configured to collect signals with a center frequency of $f_c = 2.48$ GHz, and a sampling rate of FSamp = 5MS/s. The collection receiver configuration remained fixed throughout all trials.

The performance of the MDA/ML discrimination algorithm is a function of the collected signal's $E_b/N0$, with higher $E_b/N0$ achieving better performance. Four independent Additive White Gaussian Noise (AWGN) realizations were generated to assess the performance of the MDA/ML discrimination algorithm at varying $E_b/N0$. The AWGN realizations were power scaled to represent $E_b/N0 \in [0, 3, 6, ..., 27]$. The AWGN realizations used to generate RF-DNA fingerprints were like-filtered to match the QPSK receiver passband. These AWGN noise realizations facilitate analysis of RF-DNA and CB-DNA fingerprint generation and device classification under various degraded SNR conditions. The block diagram that depicts the process to generate RF-DNA and CB-DNA fingerprints at varying $E_b/N0$s is illustrated in Figure 28.

CB-DNA Features Extraction and Fingerprints Generation.

The constellation projections were grouped based on the previous estimated symbol, current estimated symbol, and the next estimated symbol. Figure 29 illustrates this phenomenon by placing each constellation point in one of the following four groups: $[S_j, S_x, S_k]$, $[90, S_x, 90]$, $[180, S_x, 180]$, $[S_x, S_x, S_x]$, where S_x denotes current estimated symbol, and the other variables indicate a different communication symbols or angular relationship in degrees.

There are 64 possible permutations of prior, current and next estimated symbols in QPSK (i.e., [($S1$,$S1$,$S1$),($S1$,$S1$,$S2$), ..., ($S4$,$S4$,$S4$)]). CB-DNA fingerprints were generated by placing each received symbol in one of the 64 different groups. The

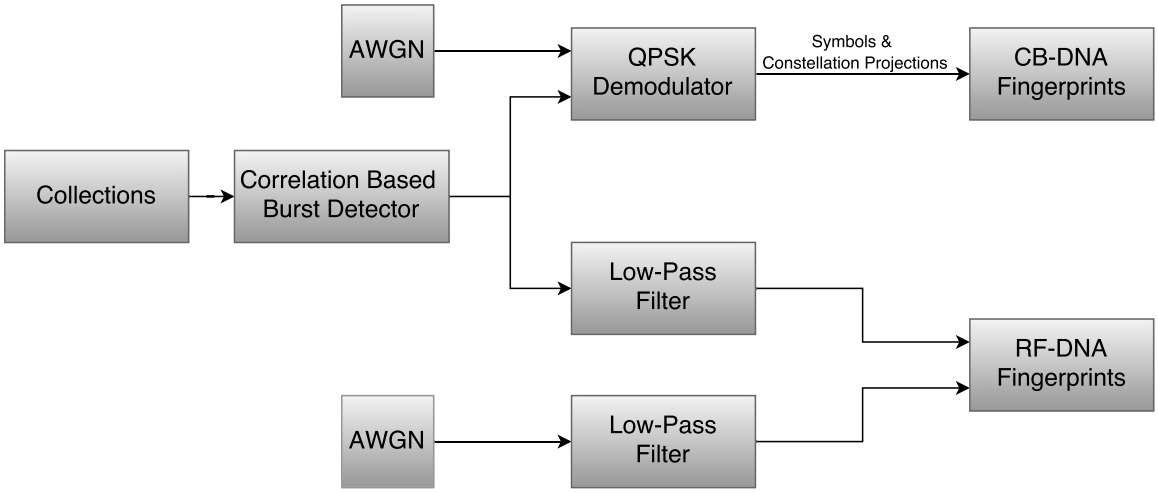

Figure 28. Block Diagram for CB-DNA and RF-DNA Fingerprint Generation Procedure

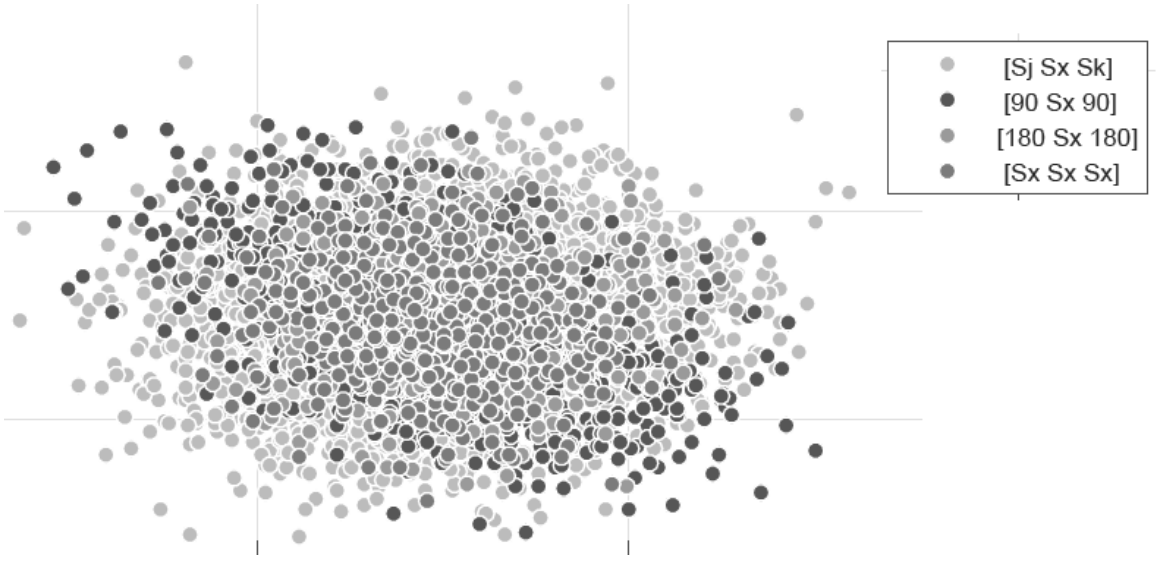

Figure 29. Conditional QPSK Projection. S_x denotes current estimated symbol, and the other variables indicate a different communication symbol or angular relationship in degrees.

identifying features were extracted by computing the following features for each of the conditional projections:

- Variance of the projected phase angle (radians)
- Variance of the projected magnitude
- Skewness of the projected phase angle (radians)
- Skewness of the projected magnitude
- Kurtosis of the projected phase angle (radians)
- Kurtosis of the projected magnitude
- Main diagonal of the covariance(real(const),imag(const))

The variance σ, skewness γ, and kurtosis κ where computing as follows:

$$\sigma^2 = \frac{1}{N_{\mathbf{x}}} \sum_{n=1} N_{\mathbf{x}} \left(\bar{\mathbf{x}}_c(N) - \mu \right)^2 , \tag{47}$$

$$\gamma = \frac{1}{N_{\mathbf{x}} \sigma^3} \sum_{n=1} N_{\mathbf{x}} \left(\bar{\mathbf{x}}_c(N) - \mu \right)^3 , \tag{48}$$

$$\kappa = \frac{1}{N_{\mathbf{x}} \sigma^4} \sum_{n=1} N_{\mathbf{x}} \left(\bar{\mathbf{x}}_c(n) - \mu \right)^4 . \tag{49}$$

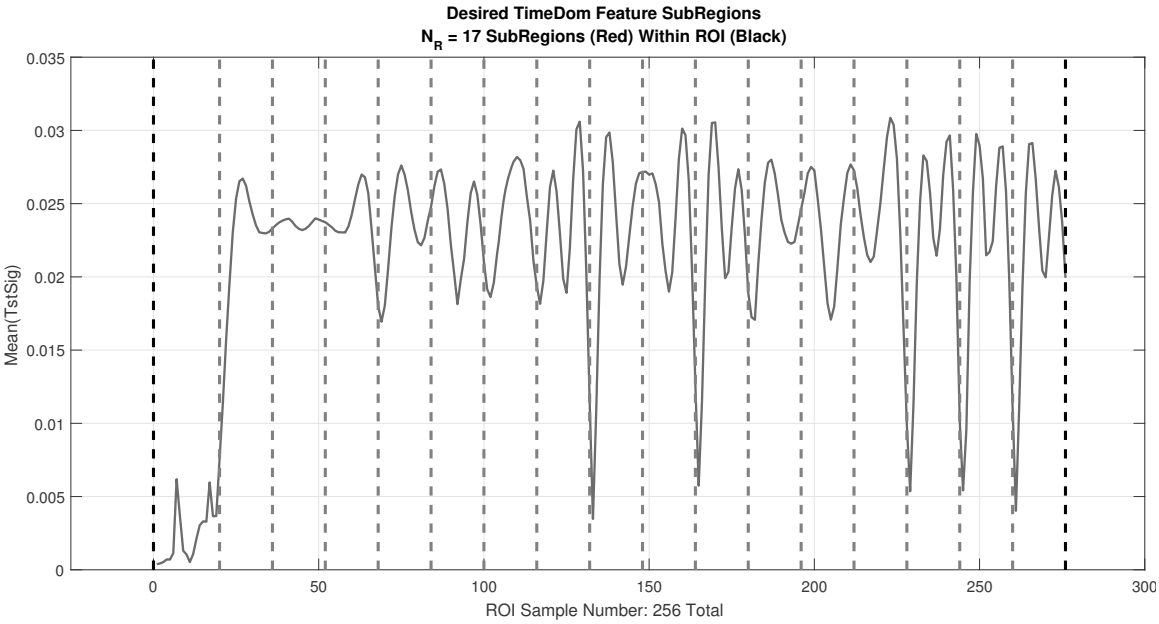

Figure 30. Mean of 1000 Bursts Preamble Response Depicting the $N_R = 17$ Sub-Regions Used for RF-DNA Fingerprint Generation. Each Sub-Region Contains 2 QPSK Symbols.

RF-DNA Features Extraction and Fingerprints Generation.

RF-DNA fingerprints are generated by extracting identifying features from portions of the signal that remain constant in between bursts such as: preambles, postambles, midambles, pilot tones, etc. This research utilizes the preamble portion of the signal as the Region of Interest (ROI). The ROI was divided into 17 subregions as shown in Figure 30. The first subregion F_{R1} shows the transmitter response as it switches from standby mode to transmit mode. Each subregion F_{R2} to F_{R17} contains the transmitter response as it emits two QPSK communication symbols.

The normalized and centered instantaneous amplitude $\mathbf{a}_c$, the normalized and centered instantaneous phase $\boldsymbol{\phi}_c$, and the normalized and centered frequency $\mathbf{f}_c$ was computed for each subregion. The vector $\mathbf{a}_c$ was computed using (22) and (25), the vector $\boldsymbol{\phi}_c$ using (23) and (26), and the vector $\mathbf{f}_c$ using (24) and (27).

The RF-DNA features were extracted by computing the standard deviation σ^2, the skewness γ, and kurtosis κ for each subregion. The values for σ^2 were computed

using (47), γ were computed using (48), and κ computed using (49).

3.4 Results

This section presents and analyzes the results of the Multiple Discriminant Analysis / Maximum Likelihood (MDA/ML) discrimination algorithm using Radio Frequency Distinct Native Attribute (RF-DNA) and Constellation-Based Distinct Native Attribute (CB-DNA) fingerprints. Until recently it was very hard to design a test that isolates the effects of baseband components on device discrimination from the effects of passband components. Nowadays we have Commercial Off-The-Shelf (COTS) Software-Defined Radio (SDR) platforms that have separable baseband and passband components. This research designed six test cases that address the worst-case scenarios for Primary User Emulation Attacks (PUEAs). The objectives of the six test cases are as follows:

- Discrimination performance based on passband components
- Discrimination performance based on baseband modulators
- Discrimination performance of like-model devices
- Discrimination performance of large number of like-model devices with mixed configurations
- Discrimination performance based on passband components across multiple baseband boards
- Discrimination performance based on baseband boards across multiple passband components

Classification experiments were conducted using $N_{\text{bursts}} = 1000$ independent bursts; $N_{\text{trainbst}} = 500$ bursts were used for MDA/ML training, and $N_{\text{tstbst}} = 500$ bursts were

used for testing. For each burst $N_{\text{Nz}} = 4$ Monte Carlo noise realizations were created at each $E_b/N0$. Each test described in this section has a total of N_{tests}=(500 bursts)$\times$($N_{\text{Nz}} = 4$)=2000 independent tests per each $E_b/N0$.

Passband Classification Performance.

CB-DNA and TD RF-DNA classification performance was assessed using one NI X310 SDR with seven different configurations. The NI X310 SDR configuration was modified by swapping the daughterboard seven times. The objective of these tests was to demonstrate the algorithm's ability to differentiate features generated by the passband components (daughterboard) while ignoring features generated by the baseband modulator (X310 mainboard). Individual configuration and average MDA/ML %C correct classification performance at $Eb/N0 \in [0, 27.0]$ dB using TD RF-DNA is shown in Figure 31, and the performance using CB-DNA fingerprints is shown in Figure 32.

For TD RF-DNA fingerprints, five of the seven individual X310 configurations achieve %C=90% or better correct classification at $Eb/N0 \geq 21$ dB. Individual classification of the remaining two X310 configurations fail to achieve %C=90% using TD RF-DNA fingerprints. The average classification performance using TD RF-DNA fingerprints exceeded %C=90% for $Eb/N0 \geq 24$ dB.

CB-DNA fingerprints achieve %C=90% or better for three configurations at $Eb/N0 \geq 21$ dB, four configurations at $Eb/N0 \geq 24$ dB, and six configurations at $Eb/N0 = 27$ dB. Individual classification of the remaining X310 configuration fails to achieve %C=90% using CB-DNA fingerprints. The average classification performance using CB-DNA fingerprints exceeded %C=90% for $Eb/N0 \geq 24$ dB.

The mean classification rate for both TD RF-DNA and CB-DNA fingerprints at $E_b/N0 = 24$ dB is $\%C \approx 91\%$ as shown in Table 4. Individual classification perfor-

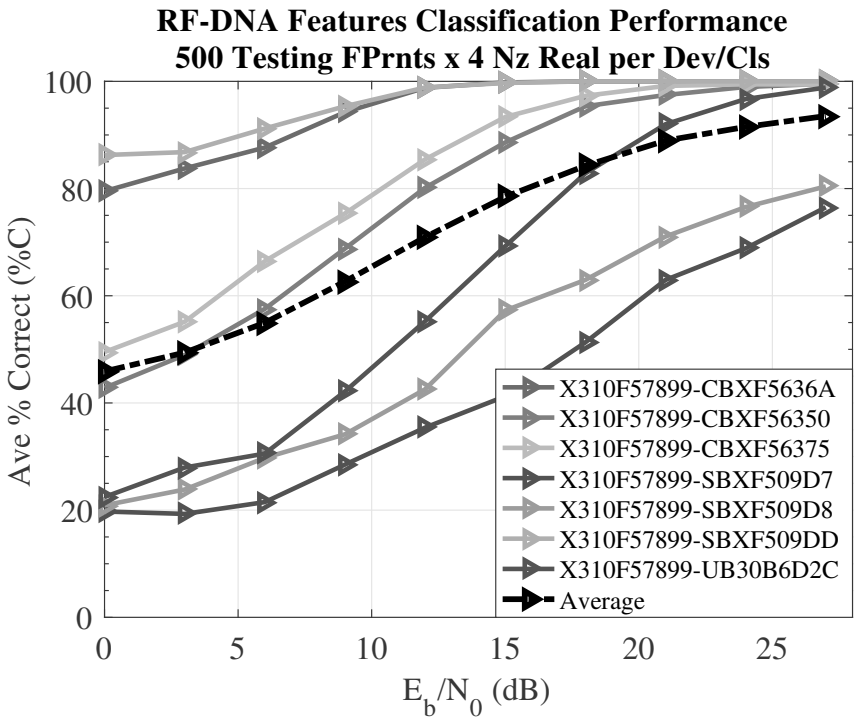

Figure 31. Passband MDA/ML Classification Performance Using TD RF-DNA Fingerprints from Seven Daughterboards and One NI X310 SDR

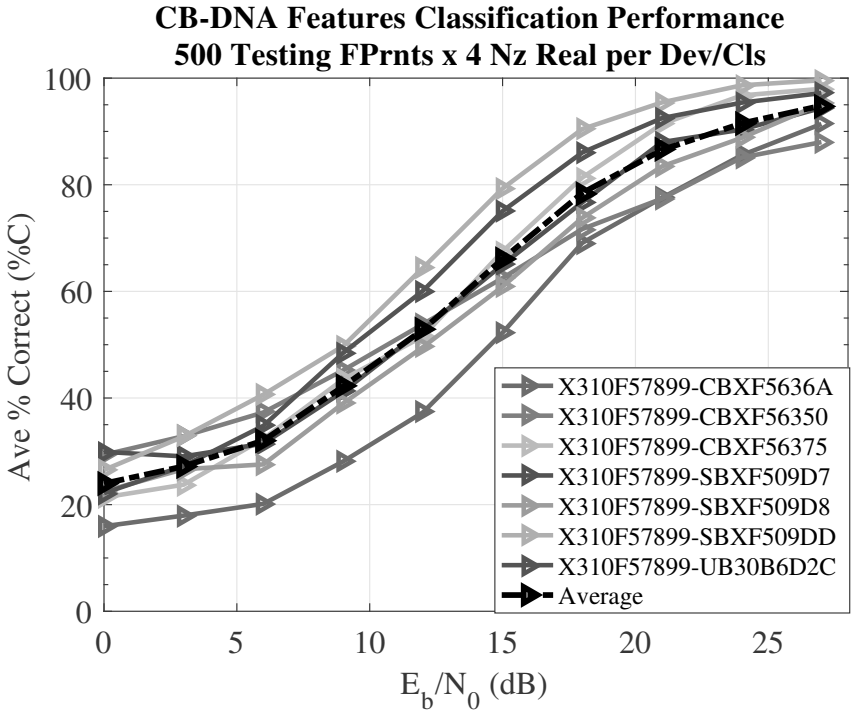

Figure 32. Passband MDA/ML Classification Performance Using CB-DNA Fingerprints, from Seven Daughterboards and One NI X310 SDR

Table 4. Confusion Matrix for $N_d = 7$ Devices Passband Classification Performance using RF-DNA/CB-DNA Fingerprints at $E_b/N0 = 24$ dB

Input Class \ Called Class	X310F57899 CBXF5636A	X310F57899 CBXF56350	X310F57899 CBXF56375	X310F57899 SBXF509D7	X310F57899 SBXF509D8	X310F57899 SBXF509DD	X310F57899 UB30B6D2C
X310F57899 CBXF5636A	**100.0%** / 85.2%	0.0% / 12.4%	0.0% / 2.3%	0.0% / 0.1%	0.0% / 0.0%	0.0% / 0.1%	0.0% / 0.1%
X310F57899 CBXF56350	0.0% / 12.1%	**98.5%** / 85.3%	0.0% / 2.6%	0.0% / 0.0%	0.0% / 0.0%	0.0% / 0.0%	1.5% / 0.0%
X310F57899 CBXF56375	0.0% / 1.1%	0.0% / 1.8%	**99.3%** / 96.7%	0.0% / 0.0%	0.8% / 0.0%	0.0% / 0.5%	0.0% / 0.0%
X310F57899 SBXF509D7	0.0% / 0.0%	0.0% / 0.1%	0.0% / 0.0%	68.2% / **92.6%**	28.2% / 5.2%	0.0% / 1.7%	3.6% / 0.1%
X310F57899 SBXF509D8	0.0% / 0.0%	0.0% / 0.0%	0.1% / 0.0%	22.6% / 8.0%	77.3% / **86.4%**	0.0% / 0.9%	0.1% / 4.8%
X310F57899 SBXF509DD	0.0% / 0.0%	0.0% / 0.0%	0.0% / 0.1%	0.0% / 0.9%	0.0% / 0.7%	**100.0%** / 98.4%	0.0% / 0.1%
X310F57899 UB30B6D2C	0.0% / 0.0%	0.6% / 0.1%	0.0% / 0.1%	3.1% / 0.6%	0.1% / 3.6%	0.0% / 0.6%	**96.3%** / 95.1%

mance for TD RF-DNA is %C$\geq$ 68%, while the individual classification performance for CB-DNA is %C$\geq$ 85%. The confusion matrix shows that the majority of misclassifications are for daughterboards from the same family (i.e., SBX is mostly confused with another SBX, CBX is mostly confused with another CBX and so forth).

Baseband Classification Performance.

CB-DNA and TD RF-DNA classification performance were assessed using four NI X310 SDRs and one daughterboard, which corresponds to four different configurations. These configurations were assembled by putting the same daughterboard into each of the four NI X310 SDRs. The objective of these tests was to demonstrate the algorithm's ability to differentiate features by the baseband modulators (X310 mainboard), while ignoring features generated by the passband component (daughterboard). Individual configuration and average MDA/ML %C performance at $Eb/N0 \in [0, 27.0]$ dB using TD RF-DNA is shown in Figure 33, and the performance using CB-DNA fingerprints is shown in Figure 34.

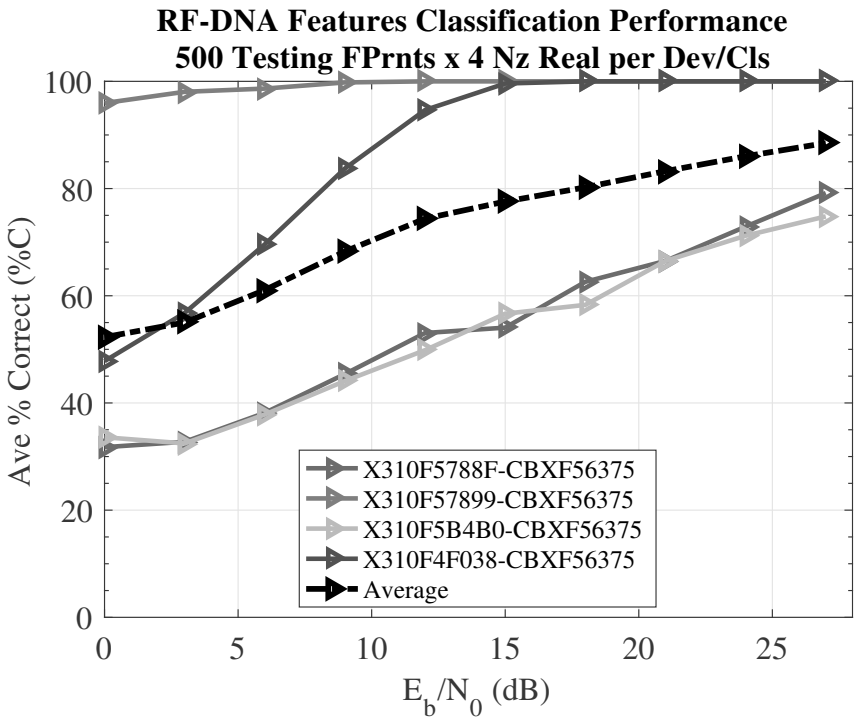

Figure 33. Baseband MDA/ML Classification Performance Using TD RF-DNA Fingerprints from One Daughterboard and Four NI X310 SDR

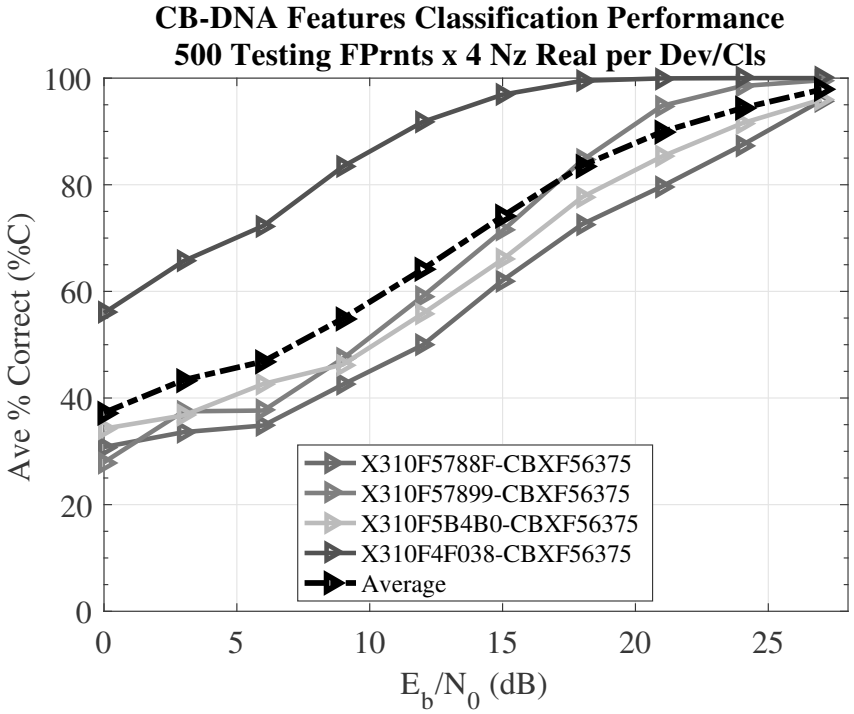

Figure 34. Baseband MDA/ML Classification Performance Using CB-DNA Fingerprints, from One Daughterboard and Four NI X310 SDR

Table 5. Confusion Matrix for $N_d = 4$ Devices Baseband Classification Performance using RF-DNA/CB-DNA Fingerprints at $E_b/N0 = 24$ dB

Called Class

Input Class	**RF-DNA/CB-DNA**	**X310F5788F CBXF56375**	**X310F57899 CBXF56375**	**X310F5B4B0 CBXF56375**	**X310F4F038 CBXF56375**
	X310F5788F CBXF56375	71.1% / **89.4%**	0.0% / 2.5%	29.0% / 8.2%	0.0% / 0.1%
	X310F57899 CBXF56375	0.0% / 1.6%	**100.0%** / 98.3%	0.0% / 0.2%	0.0% / 0.0%
	X310F5B4B0 CBXF56375	28.0% / 9.0%	0.0% / 0.4%	72.1% / **90.7%**	0.0% / 0.0%
	X310F4F038 CBXF56375	0.0% / 0.0%	0.0% / 0.0%	0.0% / 0.0%	**100.0%** / **100.0%**

For TD RF-DNA fingerprints, two of the four individual X310 configurations achieve %C=90% or better correct classification at $Eb/N0 \geq 12$ dB. Individual classification of the remaining two X310 configurations fail to achieve %C=90% using TD RF-DNA fingerprints. The average classification performance using TD RF-DNA fingerprints fails to achieve %C=90%.

CB-DNA fingerprints achieve %C=90% or better for two configurations at $Eb/N0 \geq 18$ dB and for three configurations at $Eb/N0 \geq 24$ dB. Individual classification of the remaining X310 configuration achieves %C=90% at $Eb/N0 \geq 27$ dB. The average classification performance using CB-DNA fingerprints exceeded %C=90% for $Eb/N0 \geq 21$ dB.

The mean classification rate for TD RF-DNA fingerprints at $E_b/N0 = 24$ dB is $\%C \approx 86\%$, and CB-DNA fingerprints is $\%C \approx 95\%$ as shown in Table 5. Individual classification performance for TD RF-DNA is %C$\geq$ 71%, while the individual classification performance for CB-DNA is %C$\geq$ 89%. The confusion matrix shows that the majority of misclassifications are for devices X310 serial number F5788F and X310 serial number F5B4B0. The other two devices have nearly perfect classification

performance.

Like-Model Classification Performance.

CB-DNA and TD RF-DNA classification performance were assessed using eight BladeRF SDRs. The BladeRF SDR configurations are unlike the X310 configurations, because they do not have interchangeable daughterboards, therefore each BladeRF SDR is a separate configuration. The objective of these tests was to demonstrate the algorithm's ability to differentiate features of like-model SDR by exclusively using BladeRF SDRs. Individual configuration and average MDA/ML %C performance at $Eb/N0 \in [0, 27.0]$ dB using TD RF-DNA is shown in Figure 35, and the performance using CB-DNA fingerprints is shown in Figure 36.

For TD RF-DNA fingerprints, five of the eight individual BladeRF SDRs achieve %C=90% or better correct classification at $Eb/N0 \geq 21$ dB. Individual classification of the remaining three X310 configurations fail to achieve %C=90% using TD RF-DNA fingerprints. The average classification performance using TD RF-DNA fingerprints exceeded %C=90% for $Eb/N0 \geq 24$ dB.

CB-DNA fingerprints achieve %C=90% or better for two configurations at $Eb/N0 \geq 6$ dB, five configurations at $Eb/N0 \geq 12$ dB, and eight configurations at $Eb/N0 = 18$ dB. The average classification performance using CB-DNA fingerprints exceeded %C=90% for $Eb/N0 \geq 15$ dB.

The mean classification rate for TD RF-DNA fingerprints at $E_b/N0 = 24$ dB is $\%C \approx 86\%$, and for CB-DNA fingerprints is $\%C \approx 99\%$ as shown in Table 6. The confusion matrix shows that TD RF-DNA misclassifies the Blade-RFs with serial numbers 2592, 31C4, and E078, which have an average classification rate of $\%C \approx 70\%$. Meanwhile, the lowest classification rate for CB-DNA is $\%C = 96.0\%$ for the Blade-RF with serial number CDF8.

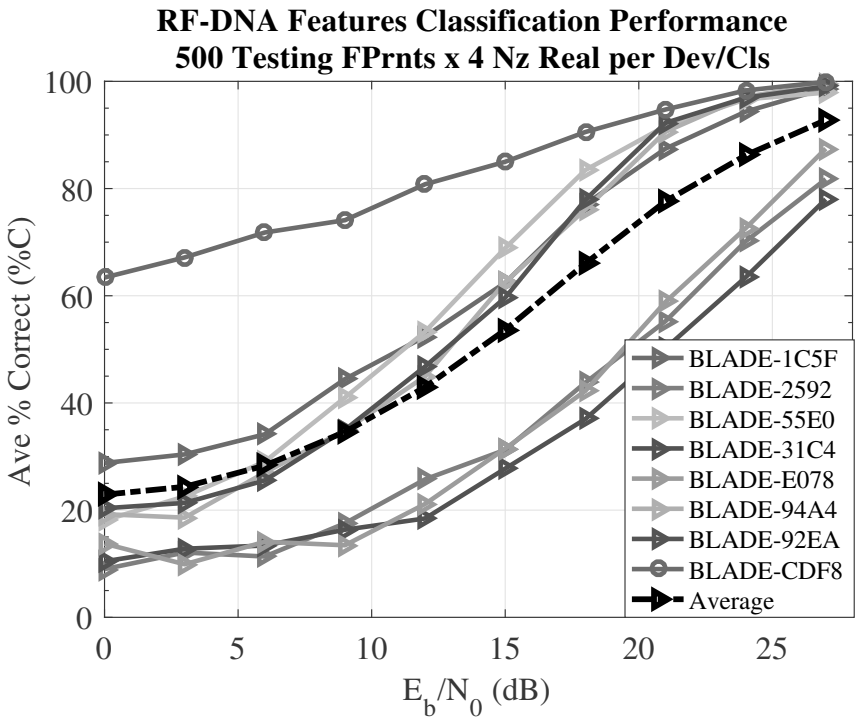

Figure 35. Like-Model MDA/ML Classification Performance Using TD RF-DNA Fingerprints from Eight BladeRFs and One NI X310 SDR with Seven Daughterboards

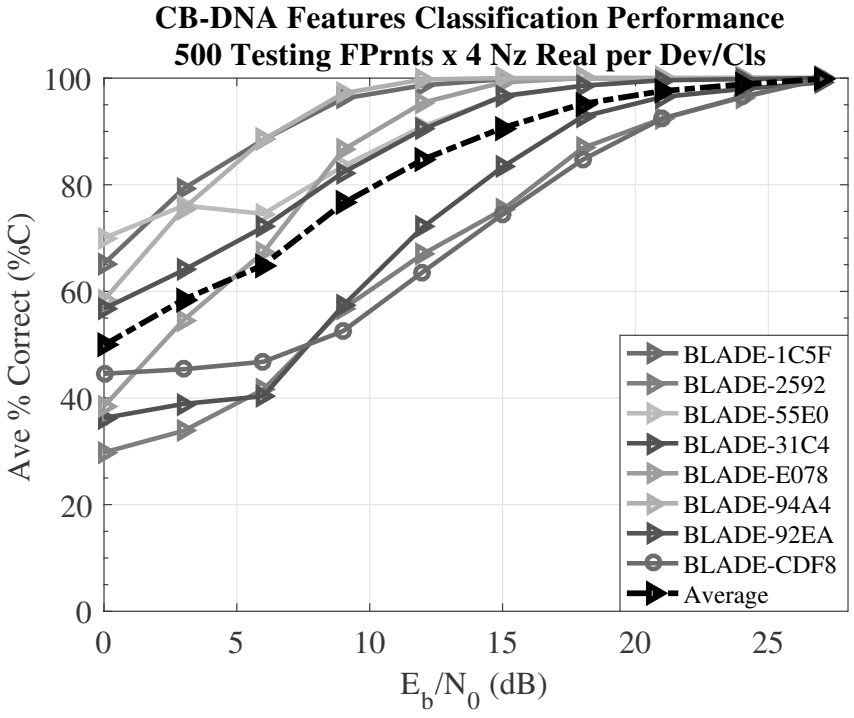

Figure 36. Like-Model MDA/ML Classification Performance Using CB-DNA Fingerprints, from Eight BladeRFs and One NI X310 SDR with Seven Daughterboards

Table 6. Confusion Matrix for $N_d = 8$ Like-Model Device Classification Performance using RF-DNA/CB-DNA Fingerprints at $E_b/N0 = 24$ dB

Input Class \ Called Class	Blade 1C5F	Blade 2592	Blade 55E0	Blade 31C4	Blade E078	Blade 94A4	Blade 92EA	Blade CDF8
Blade 1C5F	93.0% / **100.0%**	0.6% / 0.0%	0.0% / 0.0%	0.5% / 0.0%	5.9% / 0.0%	0.1% / 0.0%	0.0% / 0.0%	0.0% / 0.0%
Blade 2592	0.7% / 0.0%	73.6% / **96.3%**	0.0% / 0.0%	16.3% / 0.3%	6.9% / 0.0%	1.5% / 0.0%	0.8% / 0.0%	0.1% / 3.4%
Blade 55E0	0.0% / 0.0%	0.0% / 0.0%	95.0% / **100.0%**	0.1% / 0.0%	0.0% / 0.0%	0.1% / 0.0%	4.8% / 0.1%	0.1% / 0.0%
Blade 31C4	1.5% / 0.0%	21.0% / 0.6%	0.1% / 0.0%	64.8% / **98.0%**	11.7% / 0.0%	0.5% / 0.0%	0.1% / 0.0%	0.5% / 1.4%
Blade E078	3.9% / 0.0%	9.5% / 0.0%	0.0% / 0.0%	13.5% / 0.0%	72.2% / **100.0%**	1.0% / 0.0%	0.0% / 0.0%	0.1% / 0.0%
Blade 94A4	0.8% / 0.0%	0.6% / 0.0%	0.1% / 0.0%	0.2% / 0.0%	0.4% / 0.0%	96.2% / **100.0%**	0.0% / 0.0%	1.8% / 0.0%
Blade 92EA	0.1% / 0.0%	0.2% / 0.0%	1.2% / 0.1%	0.1% / 0.0%	0.0% / 0.0%	0.1% / 0.0%	98.4% / **100.0%**	0.0% / 0.0%
Blade CDF8	0.0% / 0.0%	0.1% / 2.9%	0.2% / 0.0%	0.6% / 1.2%	0.2% / 0.0%	0.8% / 0.0%	0.0% / 0.0%	**98.3%** / 96.0%

Mixed Device Configuration Classification Performance.

CB-DNA and TD RF-DNA classification performance was assessed using one NI X310 SDR, seven daughterboards, and eight BladeRF SDRs. Seven of the fifteen configurations were assembled with one NI X310 SDR and seven daughterboards, while the other eight configurations were BladeRF SDRs. The objective of these tests was to demonstrate the algorithm's ability to differentiate a large number of like-model devices from two different manufacturers with mixed configurations. Individual configuration and average MDA/ML %C performance at $Eb/N0 \in [0, 27.0]$ dB using TD RF-DNA is shown in Figure 37, and the performance using CB-DNA fingerprints is shown in Figure 38.

For TD RF-DNA fingerprints, two of the fifteen individual configurations achieve %C=90% or better correct classification for $Eb/N0 \geq 18$ dB, five of the fifteen individual configurations achieve %C=90% or better correct classification for $Eb/N0 \geq 21$ dB, and seven of the fifteen individual configurations achieve %C=90% or better correct classification for $Eb/N0 \geq 24$ dB. Individual classification of the remaining eight

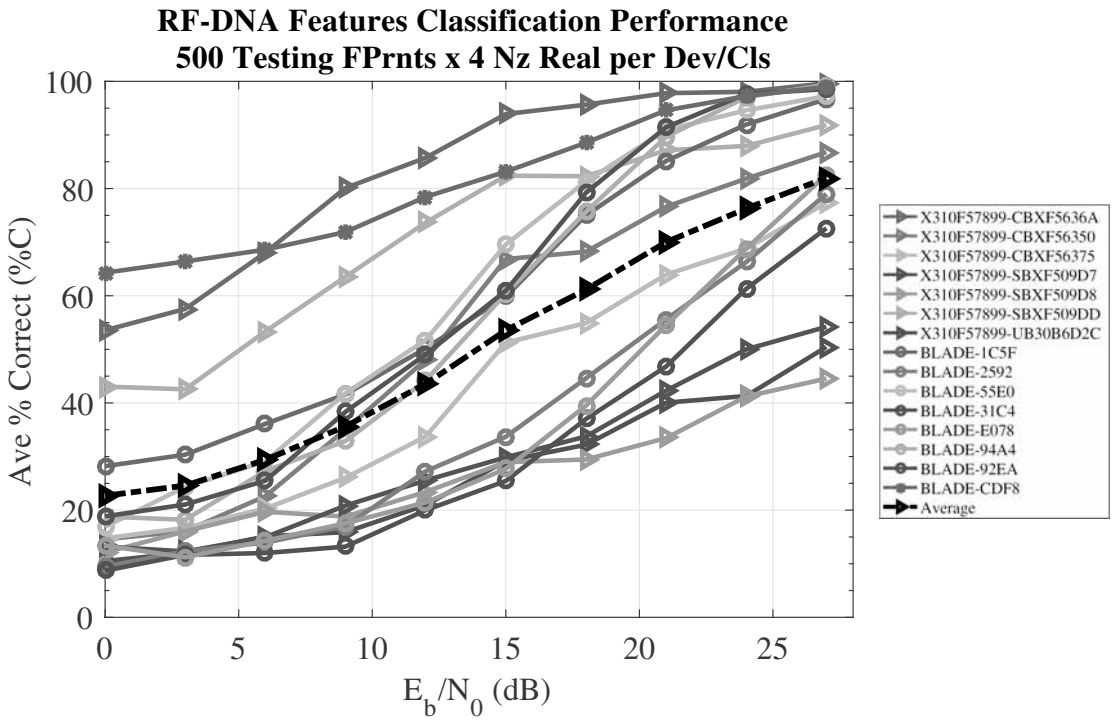

Figure 37. Mixed Device Configuration MDA/ML Classification Performance Using TD RF-DNA Fingerprints from Eight BladeRFs and Seven X310 Configurations

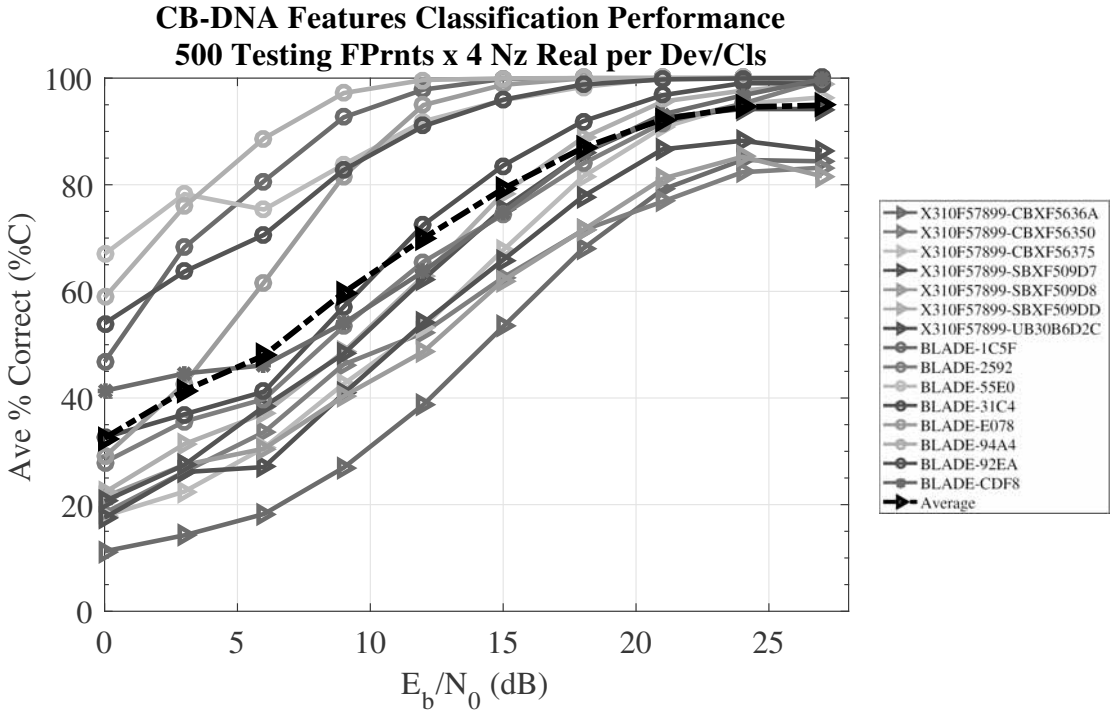

Figure 38. Mixed Device Configuration MDA/ML Classification Performance Using CB-DNA Fingerprints from Eight BladeRFs and Seven X310 Configurations

Table 7. Confusion Matrix for $N_d = 15$ Mixed Device Classification Performance using RF-DNA/CB-DNA Fingerprints at $E_b/N0 = 24$ dB

Input Class \ Called Class	X310F57899 CBXF5636A	X310F57899 CBXF56350	X310F57899 CBXF56375	X310F57899 SBXF509D7	X310F57899 SBXF509D8	X310F57899 SBXF509DD	X310F57899 UB30B6D2C	Blade 1C5F	Blade 2592	Blade 55E0	Blade 31C4	Blade E078	Blade 94A4	Blade 92EA	Blade CDF8
X310F57899 CBXF5636A	**99.1%** / 83.4%	0.9% / 13.1%	0.0% / 3.2%	0.0% / 0.1%	0.0% / 0.0%	0.0% / 0.1%	0.0% / 0.2%	0.0% / 0.0%	0.0% / 0.0%	0.0% / 0.0%	0.0% / 0.0%	0.0% / 0.0%	0.0% / 0.0%	0.0% / 0.0%	0.0% / 0.0%
X310F57899 CBXF56350	1.0% / 12.2%	83.9% / **84.2%**	0.2% / 3.6%	2.1% / 0.0%	0.2% / 0.0%	0.0% / 0.0%	12.6% / 0.1%	0.0% / 0.0%	0.0% / 0.0%	0.0% / 0.0%	0.0% / 0.0%	0.0% / 0.0%	0.0% / 0.0%	0.0% / 0.0%	0.0% / 0.0%
X310F57899 CBXF56375	0.0% / 1.0%	0.2% / 1.7%	76.9% / **96.6%**	6.6% / 0.1%	8.0% / 0.0%	6.4% / 0.7%	2.1% / 0.1%	0.0% / 0.0%	0.0% / 0.0%	0.0% / 0.0%	0.0% / 0.0%	0.0% / 0.0%	0.0% / 0.0%	0.0% / 0.0%	0.0% / 0.0%
X310F57899 SBXF509D7	0.0% / 0.0%	4.8% / 0.0%	8.2% / 0.0%	44.3% / **88.0%**	26.1% / 8.1%	0.3% / 3.3%	16.4% / 0.7%	0.0% / 0.0%	0.0% / 0.0%	0.0% / 0.0%	0.0% / 0.0%	0.0% / 0.0%	0.0% / 0.0%	0.0% / 0.0%	0.0% / 0.0%
X310F57899 SBXF509D8	0.0% / 0.0%	0.9% / 0.0%	12.1% / 0.0%	26.1% / 7.4%	43.9% / **86.6%**	0.4% / 1.5%	16.8% / 4.6%	0.0% / 0.0%	0.0% / 0.0%	0.0% / 0.0%	0.0% / 0.0%	0.0% / 0.0%	0.0% / 0.0%	0.0% / 0.0%	0.0% / 0.0%
X310F57899 SBXF509DD	0.0% / 0.0%	0.0% / 0.0%	7.4% / 0.4%	0.2% / 1.0%	0.4% / 0.7%	92.1% / **97.8%**	0.0% / 0.2%	0.0% / 0.0%	0.0% / 0.0%	0.0% / 0.0%	0.0% / 0.0%	0.0% / 0.0%	0.0% / 0.0%	0.0% / 0.0%	0.0% / 0.0%
X310F57899 UB30B6D2C	0.0% / 0.0%	15.7% / 0.0%	2.3% / 0.1%	14.7% / 0.9%	12.5% / 5.9%	0.0% / 0.3%	54.9% / **92.9%**	0.0% / 0.0%	0.0% / 0.0%	0.0% / 0.0%	0.0% / 0.0%	0.0% / 0.0%	0.0% / 0.0%	0.0% / 0.0%	0.0% / 0.0%
Blade 1C5F	0.0% / 0.0%	0.0% / 0.0%	0.0% / 0.0%	0.0% / 0.0%	0.0% / 0.0%	0.0% / 0.0%	0.0% / 0.0%	93.8% / **100.0%**	7.5% / 0.0%	0.0% / 0.0%	1.45% / 0.0%	3.6% / 0.0%	0.4% / 0.0%	0.1% / 0.0%	0.0% / 0.0%
Blade 2592	0.0% / 0.0%	0.0% / 0.0%	0.0% / 0.0%	0.0% / 0.0%	0.0% / 0.0%	0.0% / 0.0%	0.0% / 0.0%	1.2% / 0.0%	65.7% / **95.7%**	0.0% / 0.0%	20.7% / 0.4%	10.2% / 0.0%	1.6% / 0.0%	0.8% / 0.0%	0.0% / 4.0%
Blade 55E0	0.0% / 0.0%	0.0% / 0.0%	0.0% / 0.0%	0.0% / 0.0%	0.0% / 0.0%	0.0% / 0.0%	0.0% / 0.0%	0.0% / 0.0%	0.0% / 0.0%	95.4% / **100.0%**	0.1% / 0.0%	0.0% / 0.0%	0.1% / 0.0%	4.6% / 0.0%	0.0% / 0.0%
Blade 31C4	0.0% / 0.0%	0.0% / 0.0%	0.0% / 0.0%	0.0% / 0.0%	0.0% / 0.0%	0.0% / 0.0%	0.0% / 0.0%	1.1% / 0.0%	21.7% / 0.3%	0.1% / 0.0%	62.1% / **99.0%**	13.2% / 0.0%	0.9% / 0.0%	0.8% / 0.0%	0.4% / 0.8%
Blade E078	0.0% / 0.0%	0.0% / 0.0%	0.0% / 0.0%	0.0% / 0.0%	0.0% / 0.0%	0.0% / 0.0%	0.0% / 0.0%	4.7% / 0.0%	11.5% / 0.0%	0.0% / 0.0%	15.6% / 0.0%	67.7% / **100.0%**	0.5% / 0.0%	0.0% / 0.0%	0.1% / 0.0%
Blade 94A4	0.0% / 0.0%	0.0% / 0.0%	0.0% / 0.0%	0.0% / 0.0%	0.0% / 0.0%	0.0% / 0.0%	0.0% / 0.0%	0.2% / 0.0%	0.4% / 0.0%	0.1% / 0.0%	0.7% / 0.0%	0.5% / 0.0%	96.3% / **100.0%**	0.1% / 0.0%	1.9% / 0.0%
Blade 92EA	0.0% / 0.0%	0.0% / 0.0%	0.0% / 0.0%	0.0% / 0.0%	0.0% / 0.0%	0.0% / 0.0%	0.0% / 0.0%	0.1% / 0.0%	0.4% / 0.0%	2.2% / 0.2%	0.4% / 0.0%	0.0% / 0.0%	0.1% / 0.0%	97.0% / **99.9%**	0.0% / 0.0%
Blade CDF8	0.0% / 0.0%	0.0% / 0.0%	0.0% / 0.0%	0.0% / 0.0%	0.0% / 0.0%	0.0% / 0.0%	0.0% / 0.0%	0.0% / 0.0%	0.0% / 3.4%	0.1% / 0.0%	0.7% / 0.8%	0.4% / 0.0%	1.2% / 0.0%	0.0% / 0.0%	**97.7%** / 95.9%

configurations fail to achieve %C=90% using TD RF-DNA fingerprints. The average classification performance using TD RF-DNA fingerprints did not exceed %C=90%.

CB-DNA fingerprints achieve %C=90% or better for two configurations for $Eb/N0 \geq$ 9 dB, five configurations for $Eb/N0 \geq 12$ dB, seven configurations for $Eb/N0 \geq 18$ dB, and eleven configurations at $Eb/N0 = 21$ dB. Individual classification of the remaining four configurations fail to achieve %C=90% using CB-DNA fingerprints. The average classification performance using CB-DNA fingerprints exceeded %C=90% for $Eb/N0 \geq 18$ dB.

The detailed performance of TD RF-DNA and CB-DNA fingerprints at $E_b/N0=$ 24dB is shown in Table 7. The algorithm correctly classified BladeRF devices with %C≥62% and the X310 devices with %C≥43% using TD RF-DNA. The mean classification rate for BladeRF devices is %C≈ 84%, for X310 devices is %C≈ 71%, and for all devices is %C≈ 78% using TD RF-DNA. The algorithm correctly classified BladeRF devices with %C≥95% and the X310 devices with %C≥83% using CB-DNA. The mean classification rate for BladeRF devices is %C≈ 99%, for X310 devices is %C≈ 90%, and for all devices is %C≈ 95% using CB-DNA. The X310 misclassifi-

cations were from configurations using passband components from the same family (i.e., SBX is mostly confused with another SBX, CBX is mostly confused with another CBX and so forth). The classification rate of the UB30B6D2C for TD RF-DNA fingerprints was low, even though there were no other UBX daughterboards within the group of devices. The confusion matrix shows that TD RF-DNA misclassifies the Blade-RFs with serial numbers 2592, 31C4, and E078, which have an average classification rate of $\%C \approx 65\%$. Meanwhile, the lowest classification rate of Blade-RFs using CB-DNA is $\%C = 95.7\%$ for the Blade-RF with serial number 2592. These results are consistent with previous tests conducted in this research.

Passband Component Classification Across Multiple Baseband Boards.

CB-DNA and TD RF-DNA classification performance was assessed for all seven passband components (daughterboards), with each passband component being tested across four baseband components (mainboards). Fingerprints that came from the same daughterboard were combined into a single class disregarding the mainboard in which the daughterboard was installed. Seven new classes were created using this technique, one class for each daughterboard. The objective of this test was to demonstrate the algorithm's ability to differentiate passband components regardless of the baseband component in which it was installed. Individual classes as well as average MDA/ML %C performance at $Eb/N0 \in [0, 27.0]$ dB using TD RF-DNA is shown in Figure 39, and the performance using CB-DNA fingerprints is shown in Figure 40.

For TD RF-DNA fingerprints, individual classification of the seven configurations fail to achieve %C=90%. Individual classification did not show much improvement as $Eb/N0$ increased, however the performance of individual classifications converged. The average classification performance using TD RF-DNA fingerprints did not exceed

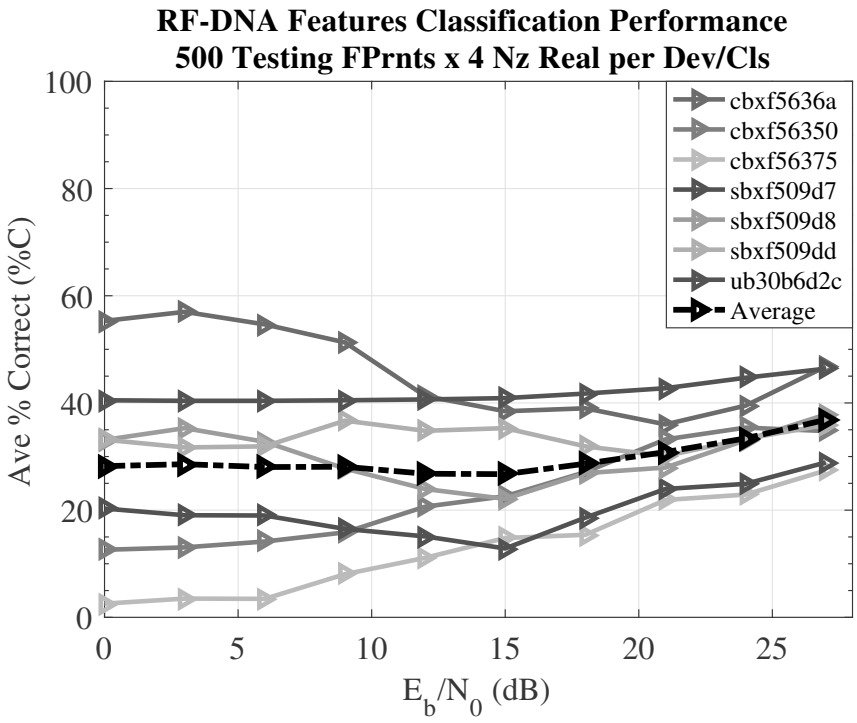

Figure 39. MDA/ML Classification Performance Using TD RF-DNA Fingerprints for Seven Daugtherboards, Each Daughterboard Tested Across Four Mainboards

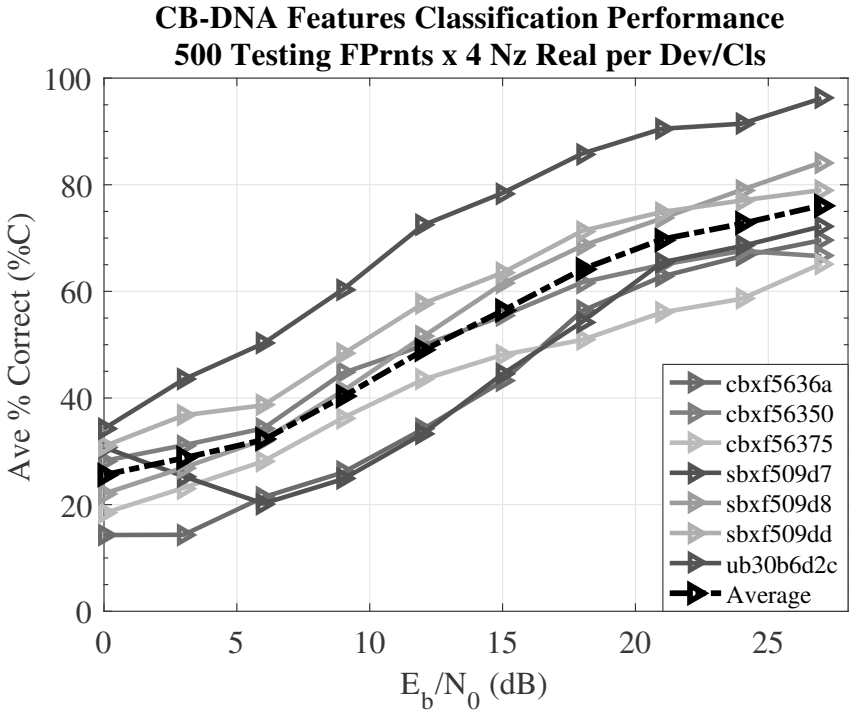

Figure 40. MDA/ML Classification Performance Using CB-DNA Fingerprints for Seven Daugtherboards, Each Daughterboard Tested Across Four Mainboards

Table 8. Confusion Matrix for MDA/ML Classification Performance Using RF-DNA/CB-DNA Fingerprints for $N_d = 7$ Daugtherboards, Each Daughterboard Tested Across Four Mainboards at $E_b/N0 = 27$ db

Called Class

Input Class	RF-DNA/CB-DNA	CBXF5636A	CBXF56350	CBXF56375	SBXF509D7	SBXF509D8	SBXF509DD	UB30B6D2C
	CBXF5636A	46.7% / **71.4%**	24.7% / 16.2%	14.1% / 12.4%	6.2% / 0.0%	3.6% / 0.0%	4.0% / 0.0%	0.8% / 0.0%
	CBXF56350	38.0% / 27.2%	33.9% / **63.6%**	15.8% / 9.0%	4.7% / 0.2%	3.9% / 0.0%	2.9% / 0.0%	0.9% / 0.0%
	CBXF56375	31.5% / 26.0%	22.8% / 15.8%	25.2% / **58.2%**	5.8% / 0.0%	4.2% / 0.0%	9.8% / 0.0%	0.9% / 0.0%
	SBXF509D7	12.9% / 0.0%	6.9% / 0.0%	7.2% / 0.0%	30.0% / **69.8%**	23.0% / 16.4%	18.9% / 13.8%	1.3% / 0.0%
	SBXF509D8	8.8% / 0.0%	7.3% / 0.0%	7.0% / 0.0%	23.3% / 14.4%	37.3% / **84.4%**	14.3% / 0.6%	2.1% / 0.6%
	SBXF509DD	9.6% / 0.0%	5.7% / 0.0%	9.5% / 0.0%	21.9% / 18.2%	16.0% / 1.2%	35.6% / **80.6%**	2.0% / 0.0%
	UB30B6D2C	9.8% / 0.0%	11.3% / 0.0%	5.2% / 0.0%	5.7% / 0.2%	19.6% / 1.6%	3.6% / 0.0%	44.9% / **98.2%**

%C=90% and only achieved %C≈37% at $E_b/N0$= 27dB.

CB-DNA fingerprints achieve %C=90% or better for one configuration for $Eb/N0 \geq$ 24 dB. Individual classification of the remaining six configurations fail to achieve %C=90% using CB-DNA fingerprints. Unlike TD RF-DNA, individual classification did show improvement as $Eb/N0$ increased, and individual classifications were clustered closer together. The average classification performance using CB-DNA fingerprints did not exceed %C=90%, but achieved %C≈77% at $E_b/N0$= 27dB. The detailed performance of TD RF-DNA and CB-DNA fingerprints at $E_b/N0$= 27dB is shown in Table 8. The algorithm correctly classified passband components from the CBX family with %C≥25%, SBX family with %C≥30%, and UBX family with %C=44.9% using TD RF-DNA. The algorithm correctly classified passband components from the CBX family with %C≥58%, SBX family with %C≥69%, and UBX family with %C=98.2% using CB-DNA. The mean classification rate for the CBX family is %C≈ 35%, SBX family is %C≈ 34%, and for all passband components is %C≈ 36% using TD RF-DNA. The mean classification rate for the CBX family is %C≈ 68%, SBX family is %C≈ 78%, and for all passband components is %C≈

75% using CB-DNA. The misclassifications were from passband components from the same family (i.e., SBX is mostly confused with another SBX and CBX is mostly confused with another CBX), although there were more misclassifications between families for TD RF-DNA.

Baseband Board Classification Across Multiple Passband Components.

CB-DNA and TD RF-DNA classification performance were assessed for all four baseband components (mainboards), with each baseband component tested across seven passband components (daughterboards). Fingerprints that came from the same mainboard were combined into a single class disregarding the daughterboard that was installed. Four new classes were created using this technique, one class for each mainboard. The objective of this test was to demonstrate the algorithm's ability to differentiate baseband components regardless of the passband component installed. Individual configuration as well as average MDA/ML %C performance at $Eb/N0 \in [0, 27.0]$ dB using TD RF-DNA is shown in Figure 41, and the performance using CB-DNA fingerprints is shown in Figure 42.

For TD RF-DNA fingerprints, individual classification of the four configurations fail to achieve %C=90%. Individual classification showed slight improvement as $Eb/N0$ increased, however the performance of individual classifications did not converge. The average classification performance using TD RF-DNA fingerprints did not exceed %C=90% and achieved %C≈55% at $E_b/N0$= 27dB.

Individual classification using CB-DNA fingerprints for all four configurations fail to achieve %C=90%. Individual classification improved as $Eb/N0$ increased and individual classifications were clustered very close together. The average classification performance using CB-DNA fingerprints did not exceed %C=90%, but achieved %C≈70% at $E_b/N0$= 27dB.

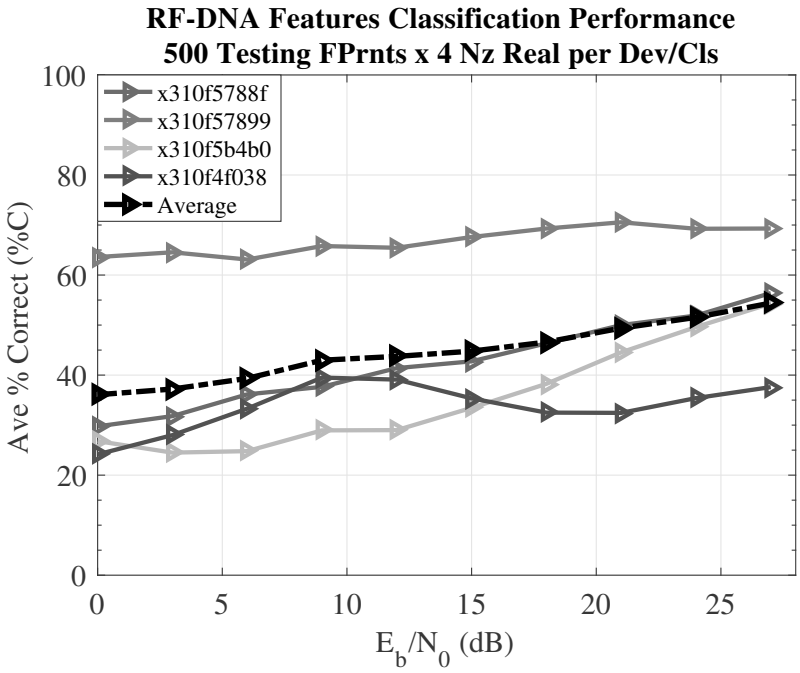

Figure 41. MDA/ML Classification Performance Using TD RF-DNA Fingerprints for N_d = 4 Mainboards, Each Mainboard Tested Across Seven Daughterboards

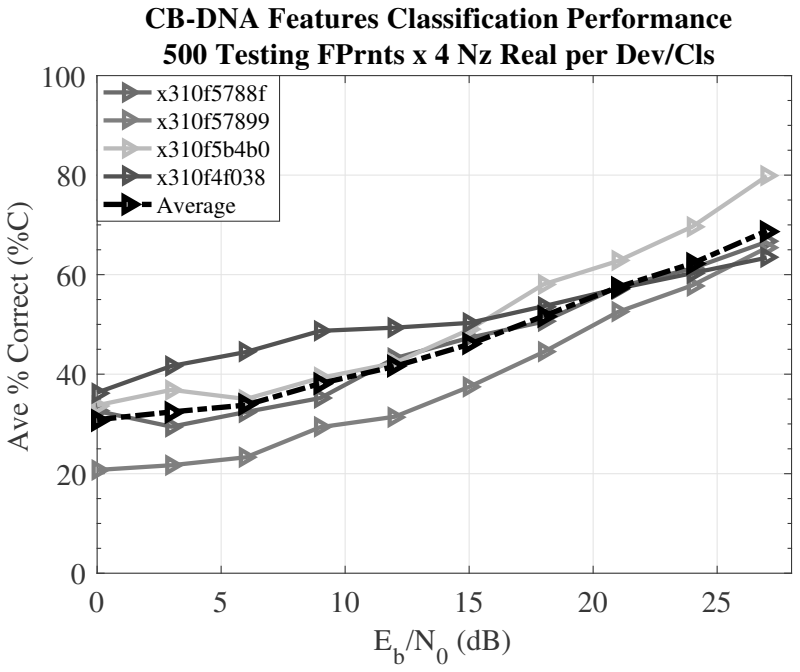

Figure 42. MDA/ML Classification Performance Using CB-DNA Fingerprints for N_d = 4 Mainboards, Each Mainboard Tested Across Seven Daughterboards

Table 9. Confusion Matrix for MDA/ML Classification Performance using RF-DNA/CB-DNA Fingerprints for N_d=4 Mainboards Tested Across Seven Daughterboards at $E_b/N0 = 27$ dB

Input Class \ Called Class	RF-DNA/CB-DNA	X310F5788F	X310F57899	X310F5B4B0	X310F4F038
	X310F5788F	54.0% / **64.6%**	11.1% / 11.0%	29.5% / 15.2%	5.5% / 9.2%
	X310F57899	14.0% / 22.8%	**68.7%** / 63.0%	13.9% / 12.8%	3.4% / 1.4%
	X310F5B4B0	27.8% / 15.8%	15.5% / 3.4%	51.7% / **77.8%**	5.1% / 3.0%
	X310F4F038	22.0% / 22.0%	18.6% / 5.4%	20.6% / 15.2%	39.0% / **57.4%**

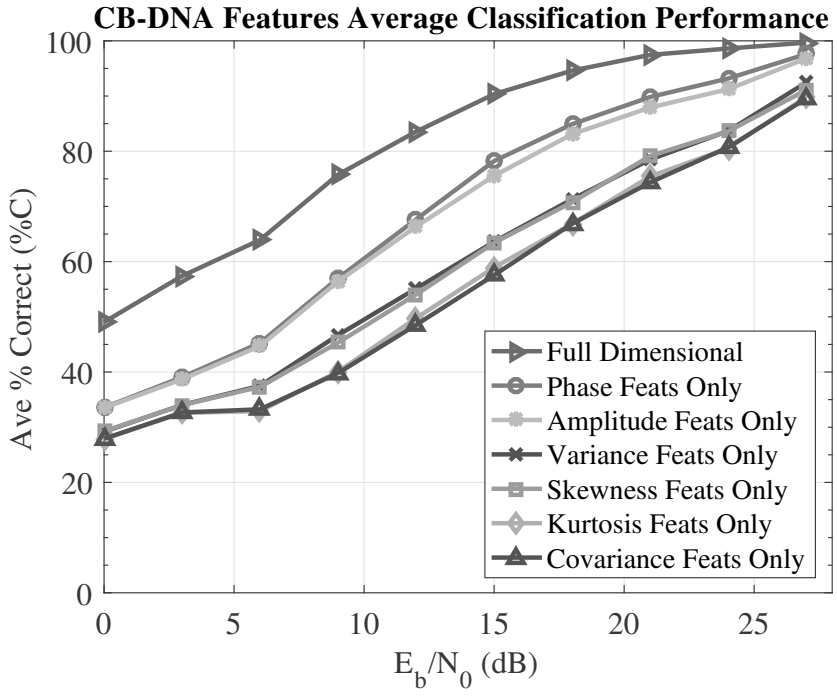

Figure 43. Comparison of Qualitative MDA/ML Classification Performance for Average %C of N_d=8 Blade-RF Like-Models Using CB-DNA Fingerprints. Qualitative Metrics Include: Covariance, Kurtosis (κ), Skewness (γ), Variance (σ^2), Magnitude, Phase Angle, and All Available Features.

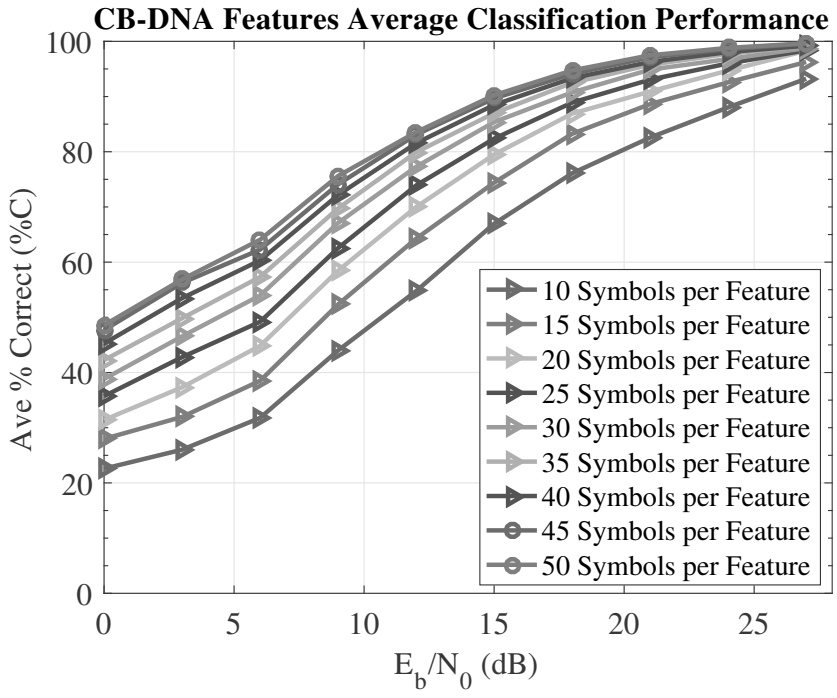

Figure 44. Average MDA/ML Classification Performance for N_d=8 Blade-RF Like-Models Using CB-DNA Fingerprints. Statistical Features Computed Using $N_{\text{symbols}} \in [10, 15, ..., 50]$.

The detailed performance of TD RF-DNA and CB-DNA fingerprints at $E_b/N0$= 27dB is shown in Table 9. The algorithm correctly classified baseband components with %C≥39% using TD RF-DNA and %C≥57% using CB-DNA. The mean classification rate for the baseband components is %C≈ 53% using TD RF-DNA. The mean classification rate for the baseband components is %C≈ 66% using CB-DNA.

Dimensional Reduction Analysis.

Full dimensional CB-DNA fingerprints have $N_{\text{feats}} = 512$ features (64 conditional Quadrature Phase Shift Keying (QPSK) projections × [2 variance + 2 skewness + 2 kurtosis + 2 covariance entries]) as described in this document. Dimensional Reduction Analysis (DRA) techniques were used to identify a proper subset of features that provide an acceptable performance, thus reducing the computational cost of the process. DRA was applied to the $N_d = 8$ Blade-RF like-model devices test case

illustrated in Figure 36. Seven new test cases were created by limiting the CB-DNA fingerprints to the following features respectively: $N_{\text{feats}} = 192$ phase angle, $N_{\text{feats}} = 192$ magnitude, $N_{\text{feats}} = 128$ covariance, $N_{\text{feats}} = 128$ variance (σ^2), $N_{\text{feats}} = 128$ skewness (γ), $N_{\text{feats}} = 128$ kurtosis (κ), and full dimensional. The detailed CB-DNA classification performance for fingerprints created with the specified proper subset of available features, as well as the full dimensional fingerprint is illustrated in Figure 43.

DRA test shows that the performance of covariance only, and kurtosis only features were nearly identical, with the lowest correct classification rate (%C). Additionally, the performance of CB-DNA fingerprints created using variance only and skewness only features were nearly identical, outperforming the previous case. Finally the performance of the classification algorithm using CB-DNA fingerprints with phase angle only and magnitude only features were nearly identical, outperforming all of the previously mentioned cases. The full dimensional fingerprints (512 features) provides a significant performance improvement over the qualitative DRA tests conducted in this research as shown in Figure 43.

All of the fingerprints generated in this research used $N_{\text{symbols}} = 50$ symbols to compute the statistics. The number of symbols used to compute the statistics (features) can affect the performance of the classification algorithm. The $N_d = 8$ Blade-RF like-model devices test was computed for 9 cases $N_{\text{symbols}} \in [10, 15, ...50]$ to illustrate how the number of symbols used to compute statistics affect the performance of the classification algorithm as shown in Figure 44. The performance of the classification algorithm improves as the number of symbols increases. However, the performance improvements for this test case asymptotically reach a limit for statistics computed using more than $N_{\text{symbols}} = 40$ symbols.

3.5 Conclusions

Traditional security techniques for preventing Primary User Emulation Attacks (PUEAs) are based on identifying the location of the source of transmission and comparing it to known Primary User (PU)'s locations. Detection of PUEAs using geolocation techniques requires a sensor network to share Radio Frequency (RF) measurements. This research presents an algorithm that identifies the true source of an emission without the aid of a sensor network by analyzing signals at the Physical Layer (PHY) layer. The proposed algorithm identifies the source of a PU emission by computing Radio Frequency Distinct Native Attribute (RF-DNA) and Constellation-Based Distinct Native Attribute (CB-DNA) fingerprints.

The effectiveness of RF-DNA and CB-DNA fingerprints to thwart a PUEA was analyzed experimentally. The performance of the algorithm was tested in four worst-case scenarios for PUEAs: like-model devices, like-model passband components, like-model baseband components, and large number of like-model devices. The tests exceeded a mean of %C=90% correct classification rate for all test cases using CB-DNA fingerprints when $E_b/N0 \geq$24 dB. Additionally, CB-DNA fingerprints outperformed RF-DNA fingerprints in all test cases.

These experiments consider the most-challenging case because all Software-Defined Radio (SDR) devices, baseband components, and passband components are brand new with the same manufacturer and model number. Classification results are expected to improve for SDR devices that are of a different brand or model number.

3.6 Appendix

Additional Results.

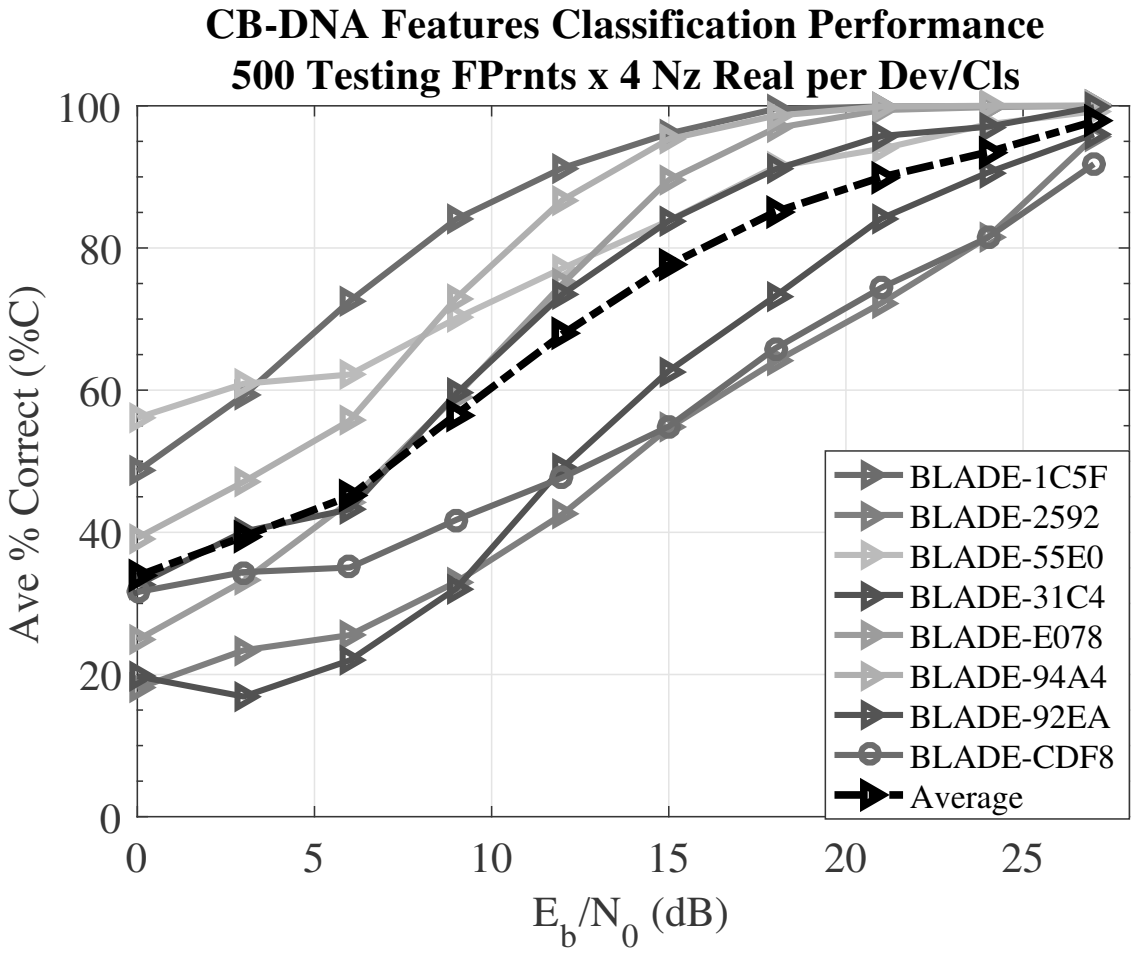

Figure 45. MDA/ML Classification Performance of CB-DNA Fingerprints Using $N_{\text{feats}} = 192$ Phase Angle Features Only: Variance (σ^2) of Phase Angle, Skewness (γ) of Phase Angle and Kurtosis (κ) of Phase Angle

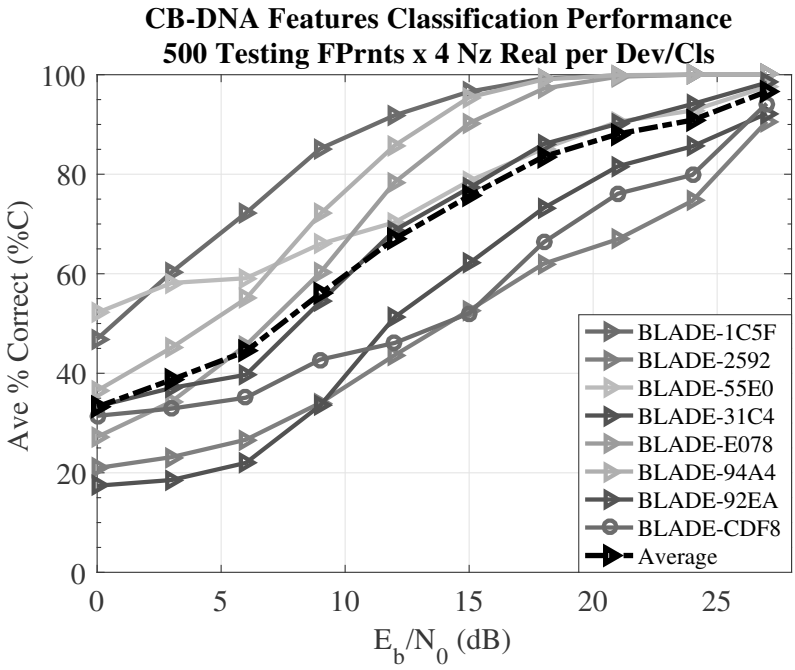

Figure 46. MDA/ML Classification Performance of CB-DNA Fingerprints Using $N_{\mathbf{feats}} = 192$ Magnitude Features Only: Variance (σ^2) of Magnitude, Skewness (γ) of Magnitude and Kurtosis (κ) of Magnitude

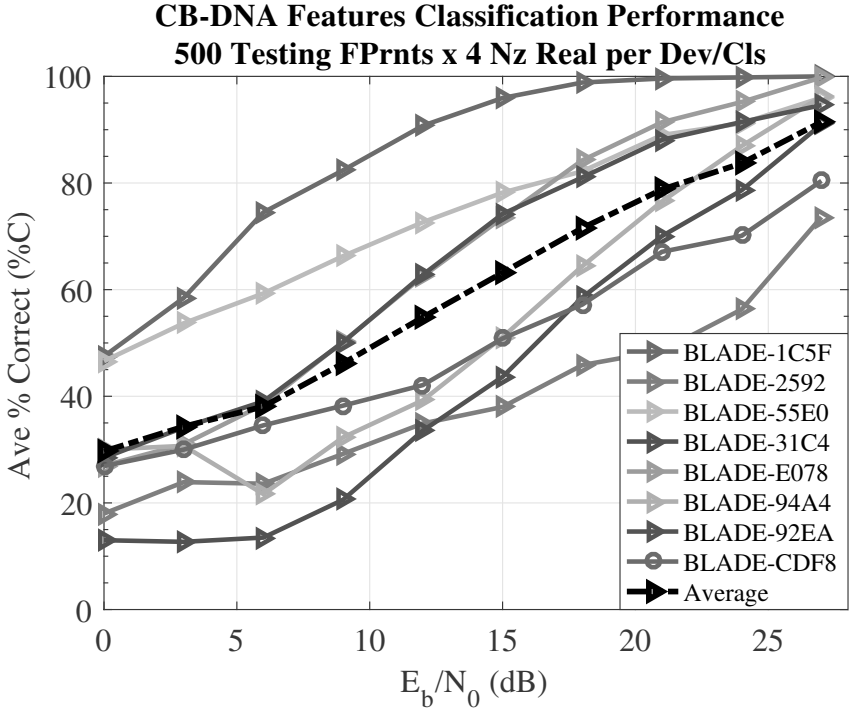

Figure 47. MDA/ML Classification Performance of CB-DNA Fingerprints Using $N_{\mathbf{feats}} = 128$ Variance Features Only: Variance (σ^2) of Phase Angle, and Amplitude

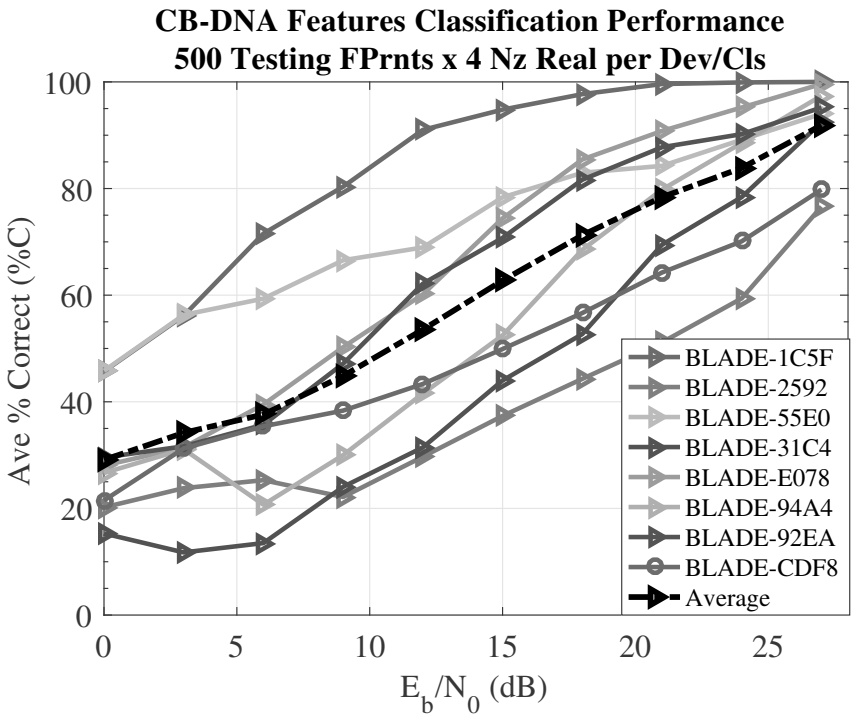

Figure 48. MDA/ML Classification Performance of CB-DNA Fingerprints Using $N_{\mathbf{feats}} = 128$ Skewness Features Only: Skewness (γ) of Phase Angle, and Magnitude

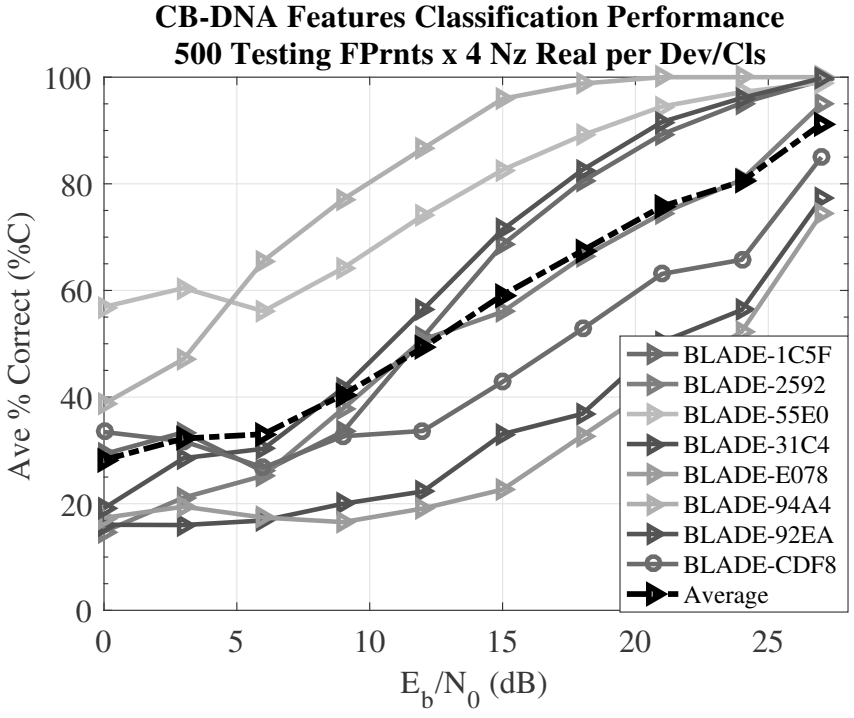

Figure 49. MDA/ML Classification Performance of CB-DNA Fingerprints Using $N_{\mathbf{feats}} = 128$ Kurtosis Features Only: Kurtosis (κ) of Phase Angle, and Magnitude

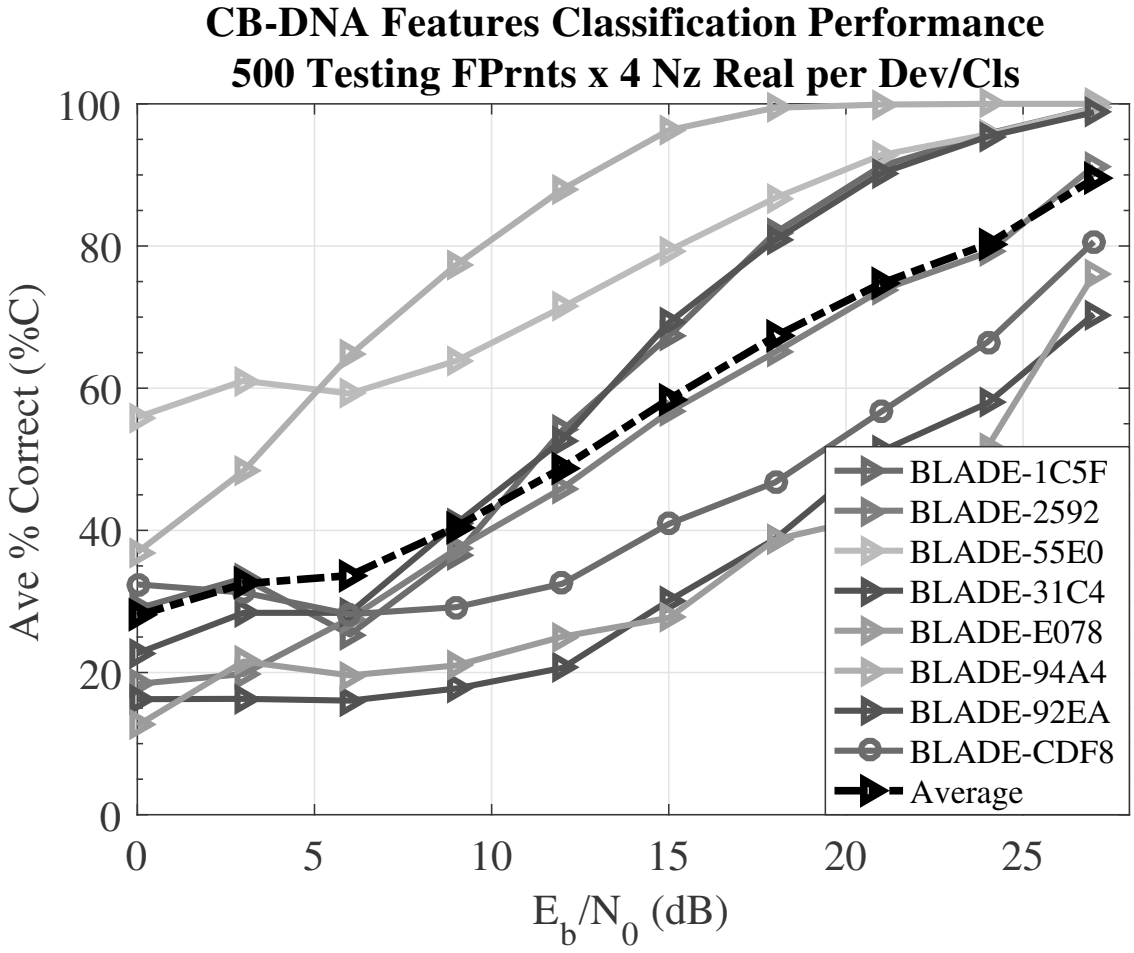

Figure 50. MDA/ML Classification Performance of CB-DNA Fingerprints Using $N_{\mathbf{feats}} = 128$ Covariance Features Only: Main Diagonal of Covariance Matrix of Real(Symbol) and Imaginary(Symbol)

IV. Robust Emitter Authentication Scheme Using Orthogonal Polyphase Based Watermarks

4.1 Introduction

The deployment of wireless networks has been growing exponentially in the last couple of decades because they provide high speed data rates and maximum mobility. The demand for wireless network access is currently saturating portions of the spectrum. Cognitive Radio (CR) is an idea proposed by researchers to alleviate spectrum scarcity by defining two types of users: Primary User (PU) and Secondary User (SU). PUs have priority above all other users, because they are licensed users of the spectrum. SUs are unlicensed users that have equal access to the spectrum whenever the PUs are not transmitting in its allocated space. Since SUs are unlicensed, they cannot interfere with the PU when utilizing their portion of the spectrum. The goal of CR is to implement intelligent and reliable radio communication systems that are aware of their environment, while adjusting their transmitter and receiver parameters to maximize spectrum efficiency.

A potential problem with the CR paradigm is a Primary User Emulation Attack (PUEA), which is when a malicious user emulates the characteristics of the PU to prevent SUs from using a portion of the spectrum. The unconstrained access to high speed data links facilitates networks exploitation by malicious users. The malicious user has two possible motives for a PUEA: gain exclusive access to a portion of the spectrum and Denial of Service (DOS).

The exploitation risks of wireless networks can be mitigated by authenticating the users participating in the network. Most authentication schemes rely on information obtained in Open Systems Interconnection (OSI) layers 2-7. This research implements an authentication scheme at the Physical Layer (PHY) to authenticate users

by embedding a watermark. Watermarking is a form of communication that embeds a concealed signal into another signal. There are multiple applications for concealed signaling, which include: copyright enforcement, steganography, and authentication. Watermarks can also be described as a method of establishing an imperceptible side-channel to exchange information [53].

The watermark signal was used to exchange information that authenticated the PU. There are multiple cryptographic solutions that may be supported in the new communication channel for message authentication. The Hash Based Message Authentication Code (HMAC) as described in [54] provides integrity of the message and authentication of transmitter with only one hash value. Another transmitter authentication method is the cryptographic link signatured implemented using a hash chain as described in [55]. The authentication codes embedded in the watermark are added in such way that does not affect receivers that are unable to extract the watermark.

4.2 Background

The objectives of this section are to provide the necessary background information to precisely define the problem and review the current state-of-the-art technologies contributing to the proposed solution. This section presents the background information using a top to bottom approach, beginning with Phase Shift Keying (PSK), orthogonal signaling, burst detection, frequency estimation, and finally narrowing down to the specific focus of this research and how to create a concealed channel by embedding information using orthogonal signaling into a PSK signal.

Phase Shift Keying Modulation.

PSK is a digital modulation scheme that encodes the information by changing the phase of a reference signal. PSK modulation is widely popular in high data-

rate Modulator/Demodulator (MODEM) implementations because this modulation scheme generates a constant power signal. Constant power signals can be implemented with non-linear power amplifiers, simplifying the receiver/transmitter design while reducing power consumption [56]. PSK signals can be represented as follows:

$$s(t) = A \exp\left(j\left(2\pi f_c t + \theta_n\right)\right) \tag{50}$$

where A represents the magnitude of the signal, f_c represents carrier frequency, t represents time, and θ_n represents the phase shift associated with a given communication symbol. Quadrature Phase Shift Keying (QPSK) is a special case of PSK modulation that can be modeled as follows [57]:

$$s_n(t) = A \exp\left(j\left(2\pi f_c t + \theta_n\right)\right) \qquad \theta_n \in \left[\frac{\pi}{4}, \frac{3\pi}{4}, \frac{5\pi}{4}, \frac{7\pi}{4}\right] \tag{51}$$

Orthogonal M-ary Signaling.

A set of N signals $\{\phi_1(t), \phi_2(t), ..., \phi_N(t)\}$ defined over a time interval $0 \leq t \leq T$ are orthonormal if:

$$\int_0^T \phi_i(t) \cdot \phi_k^*(t) dt = \begin{cases} 1, & i = k \\ 0, & i \neq k \end{cases} \tag{52}$$

Orthonormal signals can be used to transmit information by assigning a value to each $\phi_n(t)$. The optimum receiver for an orthogonal signaling system transmitted over an Additive White Gaussian Noise (AWGN) can be implemented as follows:

$$\underset{n=1,2,\ldots,N}{\arg\max} \quad \int_0^T Rx(t) \cdot \phi_n^*(t)dt, \tag{53}$$

where $Rx(t)$ represents the received signal over an AWGN, $\phi_n(t)$ represents the set of orthonormal symbols, and t represents time [58].

Signal Watermarking.

One technique to counter a Primary User Emulation Attack (PUEA) is to identify the authenticity of a user at the physical layer. Researchers at Syracuse University have developed an authentication scheme that superimposes a watermark onto the transmitted signal [59]. The watermarks are hidden in the signal by shifting the phase angle of the constellation projections, where each bit in the watermark sequence determines the direction of the phase offset. However, each phase offset is small enough to appear as noise, thereby mitigating signal degradation and hiding the watermark from malicious users. The researchers tested the implementation of this watermarking technique on two modulation schemes: QPSK and 16-ary Quadrature Amplitude Modulation (QAM). The results of the watermark Bit Error Rate (BER) for 16-ary QAM showed that the error rate decreased as the watermark length increased, and had a BER $< 10^{-5}$ when $WM_{length} = 40$ bits. Consequently, the watermark for typical authentication purposes could virtually be error free, because a $WM_{length} > 100$ bits would most likely be used.

4.3 Methodology

This section outlines the methodology used to determine the applicability of signal watermarking to authenticate the source of a Radio Frequency (RF) emission. Additionally, this section outlines the goals and hypotheses of this research, elaborates on

the problem, and describes the measures of merit on which the algorithm results will be judged. An outline of the experiments to be performed as well as the hardware and software configuration is given. The expected results are given and the expected performance factors are stated.

Research Objectives.

Physical Layer (PHY) access to wireless communication systems is hard to constrain because the transmission medium is accessible from remote locations. The unconstrained access allows malicious users to launch attacks from hidden locations. One way to mitigate these attacks is to authenticate users accessing the wireless network. This research describes a mechanism that can be used to establish the identity of RF emission. The proposed solution creates a side-channel that can be used to exchange information to authenticate the Primary User (PU).

The objective of this research is to establish a concealed communication channel to exchange information that authenticates a source of transmission in the form of watermarks. The transmitted signal degradation due to the inclusion of a watermark must be negligible.

Research Hypotheses.

There are two hypotheses that will be considered throughout this research:

- Watermarked signals should be undistinguishable from unmarked signals for users without prior knowledge.
- The addition of watermarks should have minimum impact on the communication system performance.

Measures of Merits.

The measures of merits of this algorithm are the Bit Error Rate (BER) performance of the main communication channel and the effective BER performance of the concealed signal as compared to theoretical values. Results are presented as the probability of BER in an Additive White Gaussian Noise (AWGN) channel vs Energy per Bit to Noise Power Spectral Density Ratio (E_b/N_0).

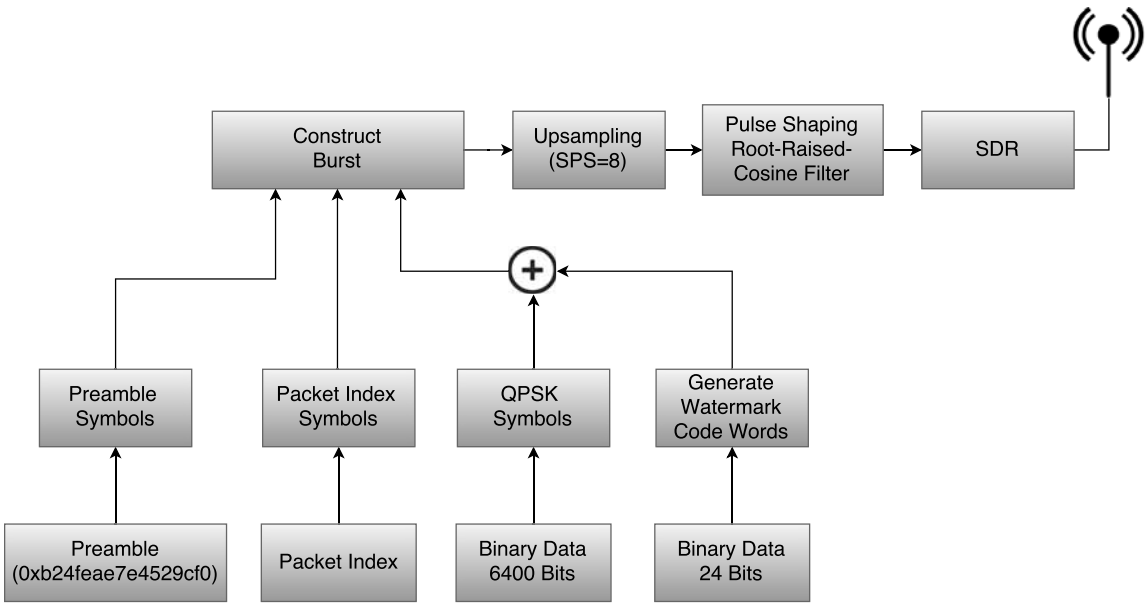

Figure 51. Block Diagram for QPSK Transmitter Implementation with Watermark Codes

QPSK Transmitter.

A QPSK modulated signal was developed to serve as a proof of concept since there are currently no standardized Cognitive Radio (CR) systems. The signal is constructed from a data packet that consists of three fields: $P_{\text{length}} = 64$ bits training sequence, $\text{Pid}_{\text{length}} = 16$ bits packet index, and $\text{Pload}_{\text{length}} = 6400$ bits payload. A watermark is constructed using $N_{\text{codes}} = 6$ code sequences that are associated with $N_{\text{bits}} = 24$ bits that were used to authenticate the transmitter. The watermark codes

were superimposed to the $\text{Pload}_{\text{length}} = 6400$ bits payload.

The training sequence serves as a preamble, and is used to aid the receiver during the synchronization process. The $P_{\text{length}} = 64$ bits sequence has very good periodic autocorrelation properties [45], which enable the receiver to detect burst presence, estimate symbol boundaries, and estimate phase angle offset between the transmitter and receiver. The $Pid_{len} = 16$ bits packet index field is used to identify the specific packet transmitted to conduct BER computations. Finally, the $Pload_{len} = 6400$ bits payload is used to represent the data to be transmitted and is populated with a sequence obtained from a Pseudo Random Number Generator (PRNG).

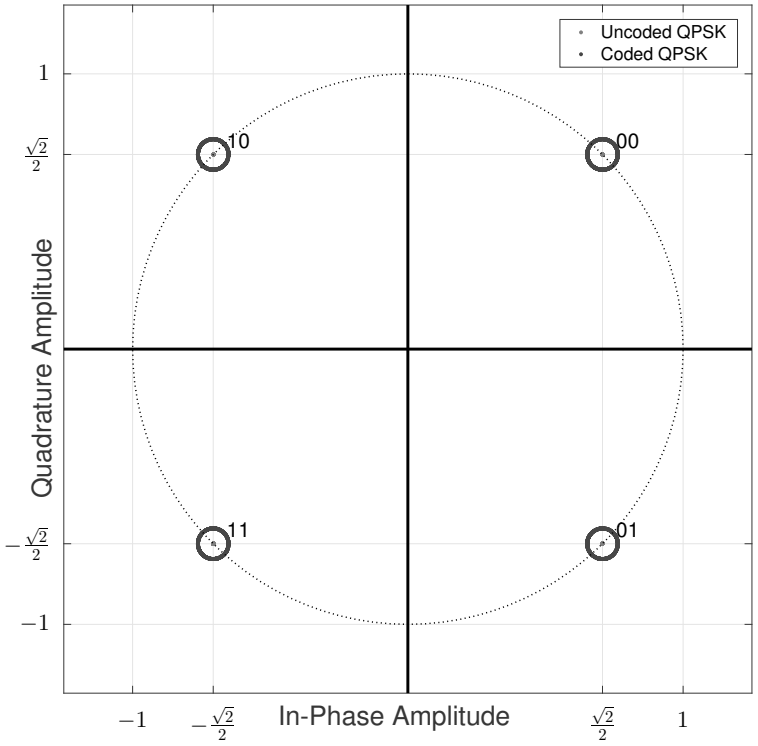

Figure 52. Constellation Projection of the Uncoded QPSK and Coded QPSK signal

The watermark sequences are added onto the modulated QPSK data symbols only. The preamble symbols and packet index symbols are left unaffected, so that the performance of the synchronization and packet reordering process is not degraded. The block diagram of this transmitter design is shown in Figure 51.

Superimposition of Watermark Codes.

An alphabet of $N_{\text{codes}} = 16$ was created to superimpose a hidden watermark onto a QPSK signal. Each of these watermark codes ($\phi_n(t)$) is a $\text{Code}_{\text{length}} = 521$ polyphase sequence on the unit circle. The $\phi_n(t)$ sequences were scaled down by a factor of $\text{Power}_{\text{ratio}} = 18$ to make the average power of the watermark signal comparable to the average power of the QPSK signal. The polyphase sequences were generated by a genetic algorithm with an objective function that provides very good autocorrelation properties and low cross correlation, so that they would be orthogonal to each other. The theoretical In-Phase/Quadrature-Phase (I/Q) projections of the coded QPSK signal and uncoded QPSK signals are illustrated in Figure 52. The coded signal can be modeled as follows:

$$\begin{aligned}
\text{codedSignal}(t) = & A\exp(j2\pi f_c t + \theta_{n(1)}) + \frac{\phi_{m(1)}(t)}{18} + \\
& A\exp(j2\pi f_c t + \theta_{n(2)}) + \frac{\phi_{m(1)}(t)}{18} + \\
& \ldots \\
& A\exp(j2\pi f_c t + \theta_{n(521)}) + \frac{\phi_{m(1)}(t)}{18} + \\
& A\exp(j2\pi f_c t + \theta_{n(522)}) + \frac{\phi_{m(2)}(t)}{18} + \\
& \ldots
\end{aligned} \tag{54}$$

Receiver.

A burst-mode QPSK receiver was implemented to project the received symbols in constellation space. The constellation points obtained from this receiver were used to extract the watermark codes embedded in the QPSK signal. Figure 53 illustrates the burst-mode QPSK receiver implemented in this project. The choice of implementa-

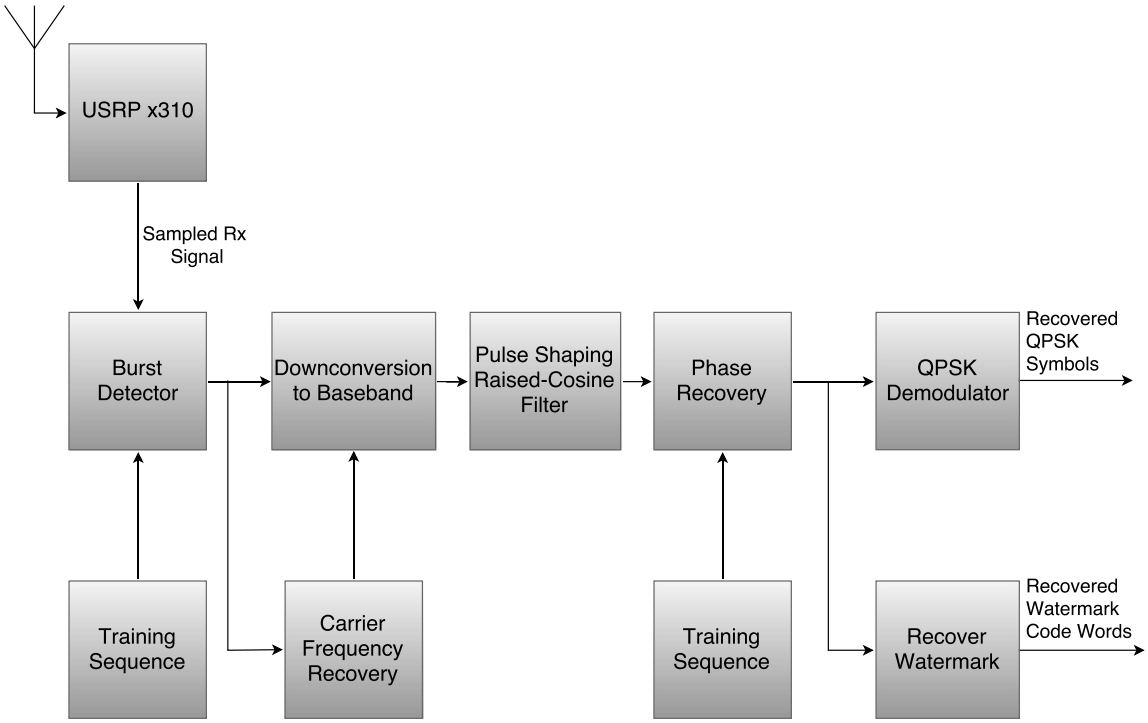

Figure 53. Block Diagram of the QPSK Receiver Implementation and Watermark Extractor

tion for the burst detector, carrier frequency recovery, and phase recovery components can significantly affect the resulting constellation projection. The respective implementations for these components are detailed in this document.

Burst Detector.

Burst detection is normally implemented using an energy detection algorithm. Using this scheme, the beginning of a burst is detected by computing when the input signal power exceeds a specified threshold. However, this research cross-correlates the received signal with the known preamble sequence to detect the presence of a burst. Using this technique, it is possible to estimate symbol boundaries, since the peak of the cross-correlation aligns with the beginning of the preamble. This technique only works when the preamble has very good correlation properties, and the center frequency offset between the transmitter and receiver is relatively small.

Intermediate Carrier Recovery.

The carrier frequency of a M-Phase Shift Keying (PSK) signal can be estimated by raising the sampled M-PSK signal to the M power in order to remove the modulation. Raising the signal to the M power creates a significant tone at M times the carrier frequency, revealing the suppressed carrier [49]. In the specific case of QPSK the tone at four times the carrier frequency is evident in the following expression:

$$\begin{aligned}\boldsymbol{R}^4(n) =& S^4\mathbf{a}^4(n)\exp(j8\pi f_c t)+ \\ & 4S^3\mathbf{a}^3(n)\exp(j6\pi f_c t)\boldsymbol{\omega}(n)+ \\ & 6S^2\mathbf{a}^2(n)\exp(j4\pi f_c t)\boldsymbol{\omega}^2(n)+ \\ & 4S\mathbf{a}(n)\exp(j2\pi f_c t)\boldsymbol{\omega}^3(n)+\boldsymbol{\omega}^4(n).\end{aligned} \tag{55}$$

This research estimated the intermediate carrier frequency in a burst-by-burst basis by computing $\hat{F}\text{Carr} = (\arg\max_n(|\mathcal{F}\{R^4(n)\}|)/4$. This technique produces reliable intermediate frequency estimates when the Signal to Noise Ratio (SNR) is $E_b/N0 > 4$ dB. It is not possible to synchronize the receiver when the SNR is $E_b/N0 \leq$ 4 dB because the intermediate frequency estimates obtained are unreliable. These limitations of intermediate frequency estimates is consistent with the Cramer-Rao Lower Bound (CRLB) for QPSK signals [50, 51].

Phase Recovery.

Typical implementations of QPSK receivers use a Phase-Locked Loop (PLL) to reconstruct the suppressed carrier. PLL algorithms use feedback to detect and compensate for phase errors [52]. For simplicity, this research implements a phase detection algorithm that rotates the received constellation points from 0 radians to $\pi/2$ radians in $N = 100$ increments, and finds the phase angle that projects symbols closer

to ideal locations. The pseudo-code for this algorithm is presented in Algorithm 2.

Algorithm 2 Phase Angle Estimator

Require: Received Constellation Projections(rxConstProj)
  rotationVariances $\leftarrow \infty$
  for $N = 1$ to 100 **do**
    $\theta \leftarrow \frac{N\pi}{2\times 100}$
    rotatedCProj $\leftarrow$ rxConstProj $\cdot e^{j\theta}$
    temp $\leftarrow$ |real(rotatedCProj)| + j |imaginary(rotatedCProj)|
    rotationVariances(N)$\leftarrow$ variance(temp)
  end for
  $N \leftarrow \arg\min_N$(rotationVariances)
  return rxConstProj $\cdot e^{\frac{jN\pi}{2\times 100}}$

There are four different phase angle ambiguities after derotating the constellation. This research resolves these ambiguities by comparing the four possible phase angles with the known preamble. Finally, the constellation projection is normalized by scaling each constellation point as follows:

$$constPoint = \frac{constPoint}{mean\left(|rxConstProj|\right)}. \tag{56}$$

Watermark Extraction.

One of the advantages of this watermark implementation is that the synchronization of the received QPSK signal does not have to be performed separately on the watermark and QPSK symbols, since the watermark's phase angle, frequency and symbol boundaries are synchronized with the QPSK signal. A single phase angle ($\hat{\theta}$) estimate and carrier frequency offset ($\hat{F}_{\text{carr}}$) estimate are computed in a burst-by-burst basis. These estimates are used for QPSK demodulation and watermark extraction, as seen in Figure 53.

The watermark extractor shown in Figure 54 has two main components: signal normalizer and code estimator. First, normalization is required because the sequences

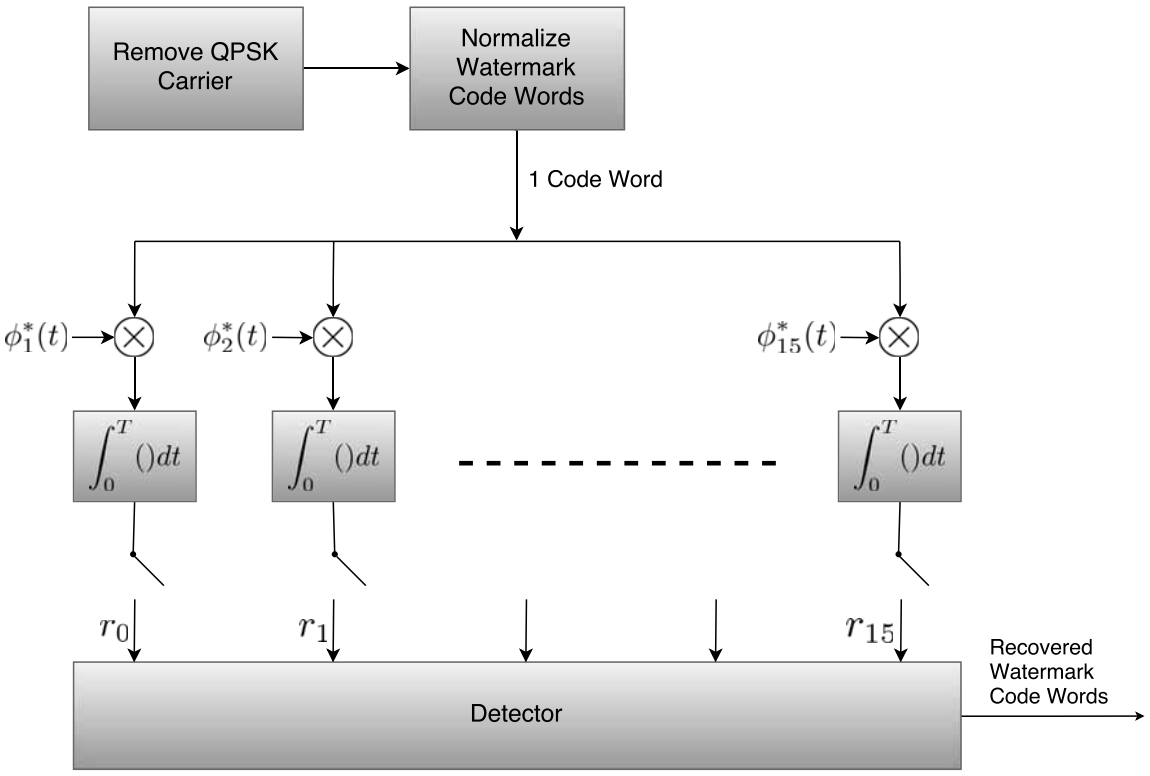

Figure 54. Block Diagram of the Watermark Extractor Implementation

were centered at the origin when they were created, then a QPSK signal was applied as a carrier. To normalize the watermark, four vectors were created and then the constellation points were sorted into these vectors based on the quadrant in which they were located. After the constellation points were sorted into their respective vector, each vector had the mean of its real components and the mean of its imaginary components subtracted from the constellation points in the vector to bring the sequences back to the origin.

The second component of the watermark extractor is the code estimator. Codes were estimated by computing the integral with respect to time of the received signal ($Rx(t)$) dotted with the complex conjugate of the reference signals ($\phi_n^*(t)$). Once the received signal was integrated with all possible $N_{\text{codes}} = 16$, the received watermark code was determined from the code that provides the maximum integration value. There were four bits of data stored in each watermark code, since there were a total

of $N_{\text{codes}} = 16$ watermark codes.

4.4 Experimental Results

This section presents and analyzes the results of the coded Quadrature Phase Shift Keying (QPSK) signal and contrasts the performance with uncoded QPSK signal and theoretical results. The objectives of these tests are as follows:

- Measure performance of information transmitted via the concealed communication channel (watermark).
- Quantify signal degradation of QPSK modulation due to embedded watermark.

Generation of Orthonormal Watermark Codes.

An evolutionary algorithm was utilized to compute a set of polyphase orthogonal signals. The objective of evolutionary algorithms is to minimize a given fitness function [60]. The pseudo-code for the fitness function that the genetic algorithm optimized is shown in Algorithm 3.

Algorithm 3 Genetic Algorithm Fitness Function

```
maxCorrelationValue ← 0
for IDX1 = 1 to 15 do
  for IDX2 = IDX1 to 16 do
    if IDX1 ≠ IDX2 then
      maxCorrelationValue= maxCorrelationValue+∑ |φ(IDX1) · φ(IDX2)|
    end if
  end for
end for
return maxCorrelationValue
```

The resultant signals were used as the reference codes (ϕ_n) that formed the orthonormal signaling system. The cross-correlation of all polyphase codes is shown in Figure 55.

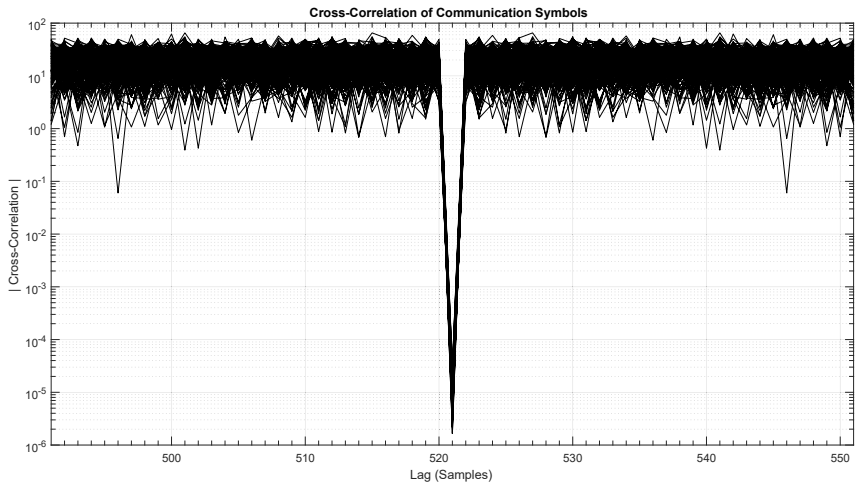

Figure 55. Cross-Correlation of $N_{\mathbf{symbols}}$=16 Orthogonal Polyphase Communication Symbols of Length $\mathbf{Symbol}_{\mathbf{length}} = 521$

The codes generated using the genetic algorithm had very good autocorrelation properties, even though the fitness function did not intentionally optimize these properties. These sequences had very good autocorrelation properties because they were obtained using random numbers and were very long. The autocorrelation of all sequences is shown in Figure 56.

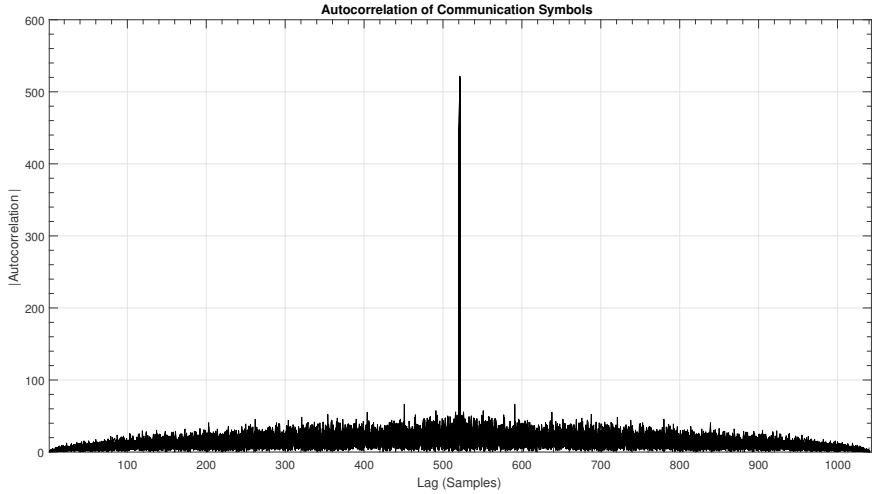

Figure 56. Autocorrelation of $N_{\mathbf{symbols}}$=16 Orthogonal Polyphase Communication Symbols of Length $\mathbf{Symbol}_{\mathbf{length}} = 521$

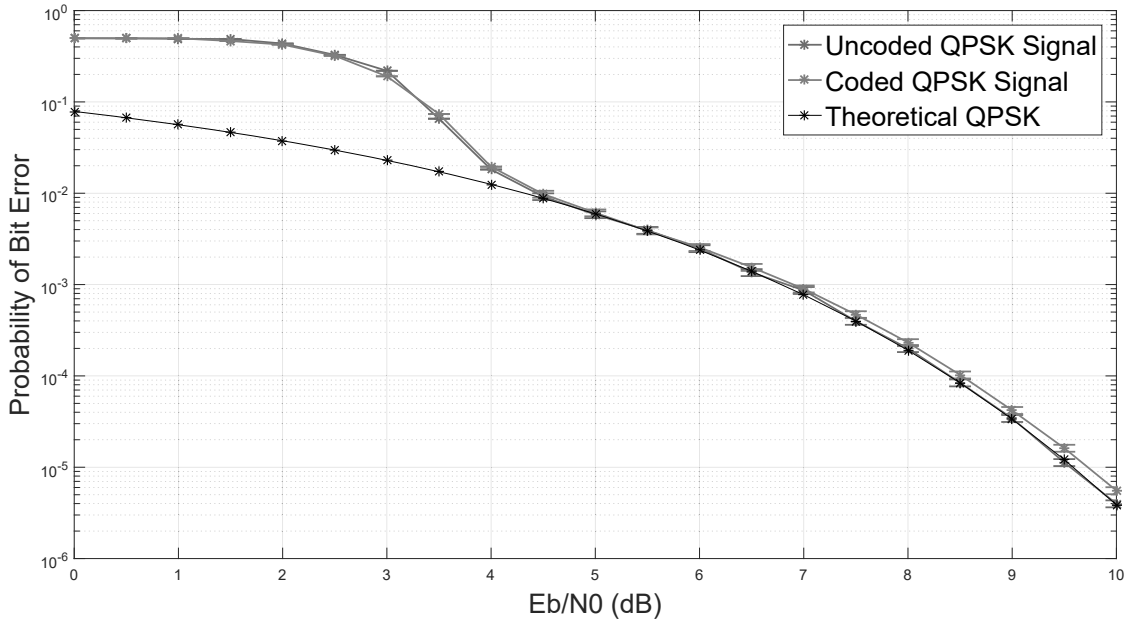

Figure 57. Performance of QPSK Receiver for Coded Signals and Uncoded Signals Showing the 99% Confidence Intervals

Coded QPSK Performance.

The implementation of the QPSK receiver did not need to be modified to account for the embedded watermark. This behavior was tested by simulating the system with a signal in which the embedded watermark codes $\phi_m = 0$ as described in (54). The Bit Error Rate (BER) performance of the communication system was only marginally affected by the embedded signal. The performance of the QPSK receiver was consistent with theory for $E_b/N_0 \geq 5$. The receiver did not achieve synchronization for $E_b/N0 \leq$ 4 dB because the intermediate frequency estimates ($\hat{F}_{\text{carr}}$) obtained were unreliable as illustrated in Figure 57. These limitations in the computation of intermediate frequency estimates ($\hat{F}_{\text{carr}}$) is consistent with the Cramer-Rao Lower Bound (CRLB) for QPSK signals [50, 51].

There is no statistical difference in the performance of the QPSK receiver between coded and uncoded for $E_b/N_0 \leq 9$ dB. The performance of the uncoded signal at $E_b/N_0 = 10$ dB was 3.99×10^{-6}, while the performance of the system at $E_b/N_0 = 10$

dB for coded signal was 5.55×10^{-6}.

Performance of Watermark Codes Extraction.

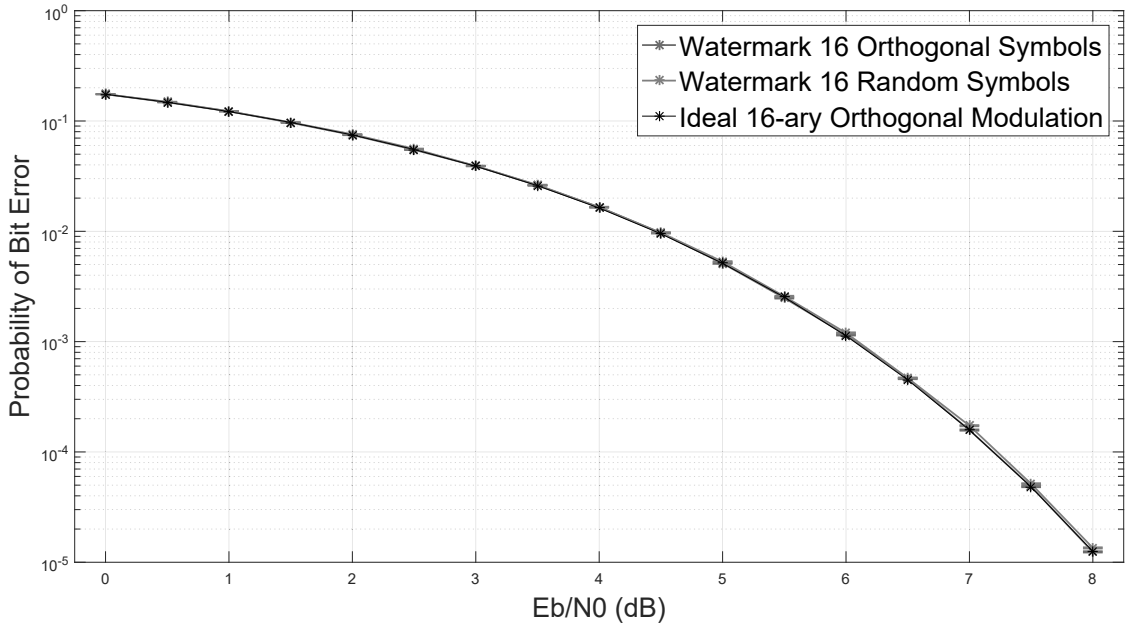

Figure 58. BER for Watermark with Symbols of Length $\text{Symbol}_{\text{length}} = 521$ Indicating the 99% Confidence Interval

The performance of the watermark extraction was tested by simulating the system with a signal in which the amplitude $A = 0$ of the signal as modeled in (54). The system was tested with two sets of codes: orthonormal codes, and random sequences. The performance of the system was compared with theoretical performance of M-ary orthogonal signaling system over an Additive White Gaussian Noise (AWGN) channel.

The performance of the watermark extraction system was consistent with theoretical values. It was also observed that there was no statistical difference between codes with orthonormal sequences and codes with random sequences for $E_b/N_0 < 7$ as illustrated in Figure 58. Even for $E_b/N_0 \geq 7$ the difference in performance was negligible.

Performance of QPSK Receiver and Watermark Extraction.

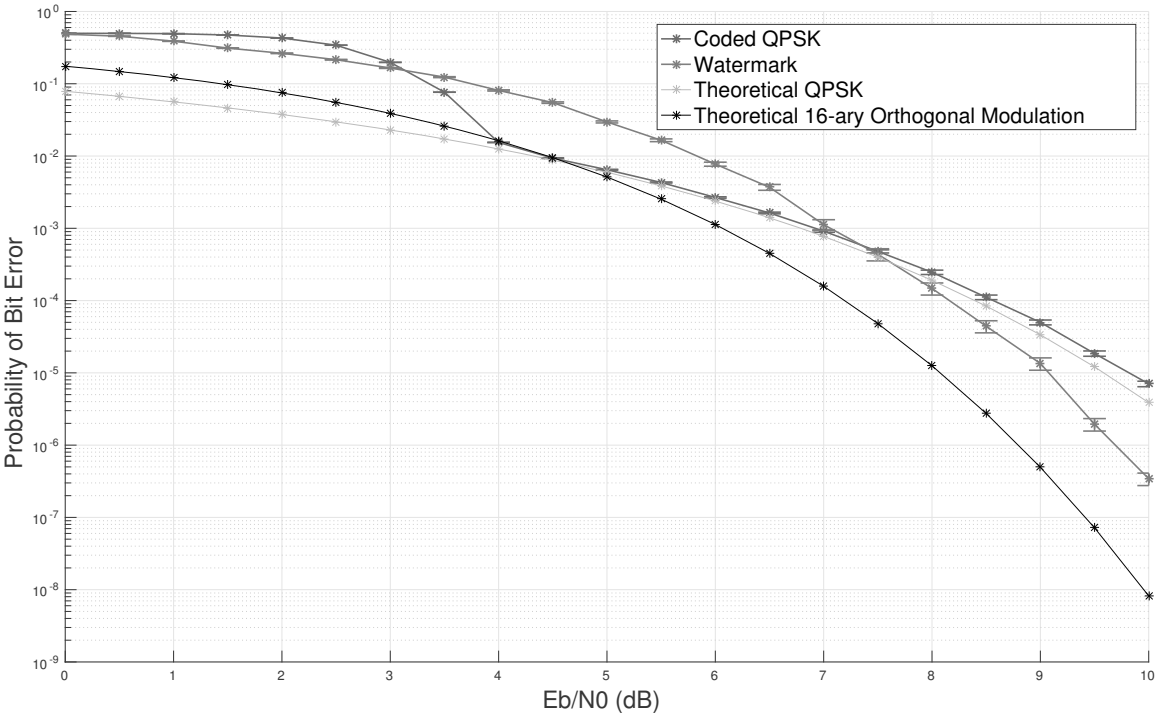

Figure 59. BER for Coded QPSK signal and Watermark Extraction Showing the 95% Confidence Interval

The performance of the QPSK receiver and watermark extraction is shown in figure 59. The BER for watermark codes outperforms the QPSK BER for $E_b/N_0 > 7$ dB. This behavior is desirable because the bits used for authentication had very low probability of error. This difference in performance was due to the different data rates between the two signals. The data rate ratio between the watermark signal and the QPSK signal is 1:260 bits.

4.5 Conclusions

Software-Defined Radios (SDRs) are essentially arbitrary waveform generators, capable of emulating the Radio Frequency (RF) emissions for any given transmitter. This research explains a method that establishes a concealed communication channel,

which can be used to exchange credentials to authenticate the Primary User (PU). The concealed communication channel was added to the signal as a watermark, minimizing the impact to the primary signal. Watermark extraction was very easy to implement, minimizing the processing power required to authenticate the user. Additionally, Secondary Users (SUs) not equipped to process the watermark are able to retrieve the information contained in the primary signal. The Bit Error Rate (BER) of the main signal at a Signal to Noise Ratio (SNR)=8 E_b/N_0 dB was 2.46×10^{-4} while the theoretical value was 1.9×10^{-4}. The BER performance of the extracted watermark at an SNR=8 E_b/N_0 dB was 1.47×10^{-4}.

4.6 Appendix

Additional Results.

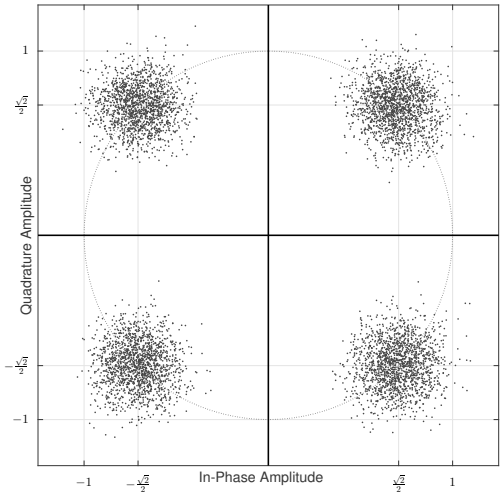

Figure 60. Constellation Projection of Uncoded QPSK Signal at E_b/N_0=15 dB. Signal transmitted over-the-air using a Blade-RF SDR transmitter and received with a NI X310 SDR.

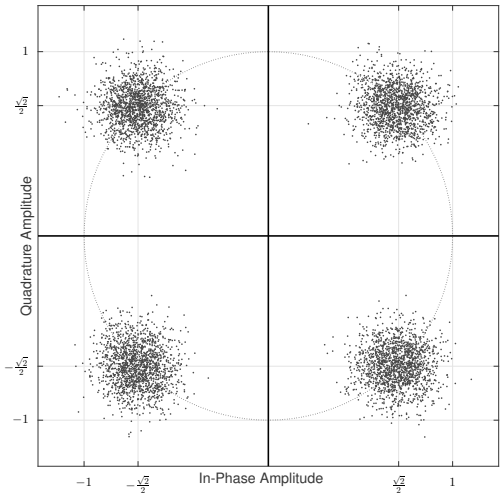

Figure 61. Constellation Projection of Coded QPSK Signal at E_b/N_0=15 dB. Signal transmitted over-the-air using a Blade-RF SDR transmitter and received with a NI X310 SDR.

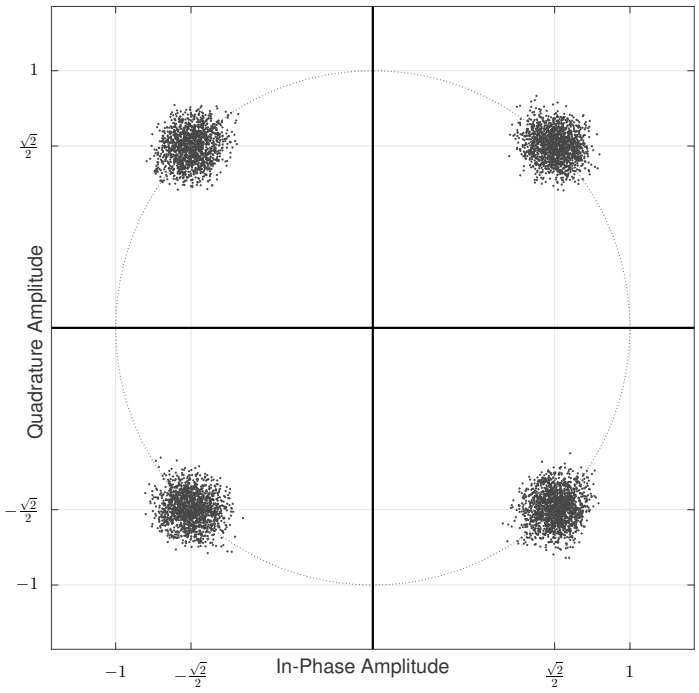

Figure 62. Constellation Projection of Uncoded QPSK Signal at E_b/N_0=25 dB. Signal transmitted over-the-air using a Blade-RF SDR transmitter and received with a NI X310 SDR.

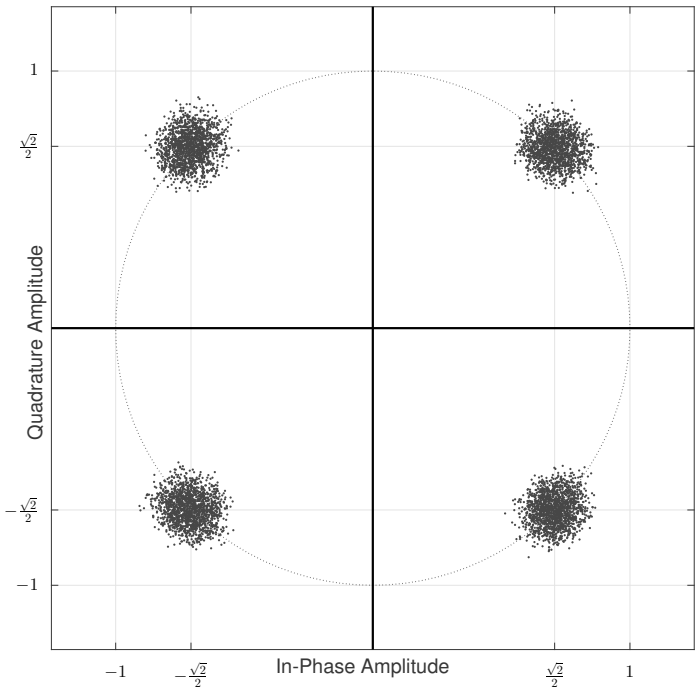

Figure 63. Constellation Projection of Coded QPSK Signal at E_b/N_0=25 dB. Signal transmitted over-the-air using a Blade-RF SDR transmitter and received with a NI X310 SDR.

V. Conclusions

The use of communication systems based on wireless links has been growing exponentially for the last couple of decades. Some portions of the spectrum are currently saturated in an attempt to accommodate the recent surge of spectrum users. The spectrum scarcity problem is exacerbated by the fixed spectrum allocations mandated by current laws. Cognitive Radio (CR) is an idea proposed by researchers that mitigates spectrum scarcity by defining two types of users: Primary Users (PUs) and Secondary Users (SUs). PUs are licensed users that have priority for the part of the spectrum that they own. SUs are unlicensed users of the spectrum with equal access rights whenever the PU is not transmitting. Therefore, any SU transmission needs to be generated in a way that minimizes interference with PU.

There is potential to abuse the spectrum sharing scheme as defined by the CR concept. Malicious users can create a Primary User Emulation Attack (PUEA) by generating signals that mimic PU's Radio Frequency (RF) radiations. There are two main reasons to launch a PUEA: illegally obtain exclusive spectrum access and Denial of Service (DOS). Previous research methods to mitigate PUEAs fall into three main ideas: Naive detection, localization-based and Physical Layer (PHY) coding. Naive detection methods estimate the mean and variance of the PU's transmissions, and use future measurements for authentication. Localization based methods authenticate PU transmissions by estimating the location of the RF emanations and comparing them to known PU's locations. PHY coding methods estimate the location of the source of emissions by letting a reference signal interfere with the PU's emissions, and analyze the results from the point of view at multiple receivers.

With the exception of naive detection, these methods rely on a network of nodes sharing RF measurements to authenticate the source of transmission. Additionally, the computation of location estimates requires a lot of processing power. This dis-

sertation describes three methods to authenticate the source of a RF emission by inspecting signals at PHY: device discrimination using Radio Frequency Distinct Native Attribute (RF-DNA) fingerprinting, device discrimination using Constellation-Based Distinct Native Attribute (CB-DNA) fingerprinting, and signal watermarking.

RF-DNA fingerprints were generated by computing statistics of a portion of the received signal that remains constant in all transmissions. Burst-mode wireless Modulator/Demodulators (MODEMs) normally add known sequences in fixed portions of the signal (i.e., preambles, postambles, midambles, pilot tones, etc.) to aid the receiver during the synchronization process. This dissertation generated RF-DNA fingerprints for $N_d = 15$ devices with mixed configurations: 8 like-model Blade-RF Software-Defined Radios (SDRs) devices and 7 National Instruments (NI) X310 SDRs. The mean correct classification rate using RF-DNA fingerprints was %C=78%.

CB-DNA fingerprints were generated by projecting the received signal into a constellation space. The resulting constellation projections are grouped based on the previous, current, and next estimated symbol. The CB-DNA identifying features are obtained by computing the statistics (variance, skewness, kurtosis, etc.) on each conditional projection. The effectiveness of CB-DNA fingerprints to thwart a PUEA was analyzed experimentally. This dissertation generated CB-DNA fingerprints for $N_d = 15$ devices with mixed configurations: 8 like-model Blade-RF SDRs devices and 7 NI X310 SDRs. The algorithm correctly classified BladeRF devices with %C$\geq$95% and the X310 devices with %C$\geq$83% using CB-DNA. The mean classification rate for BladeRF devices was %C$\approx$ 99%, X310 devices was %C$\approx$ 90%, and for all devices was %C$\approx$ 95% using CB-DNA.

The watermark method establishes a side-channel that enables the exchange of Hash Based Message Authentication Code (HMAC) that authenticates the PU. The proposed signal watermarking implementation derives synchronization parameters

from the main communication channel, minimizing the required processing power. The established communication link provides reliable Bit Error Rate (BER) performance even at a low Signal to Noise Ratio (SNR). For example, the BER in an Additive White Gaussian Noise (AWGN) channel was 1.47×10^{-4} at an SNR=8 E_b/N_0 dB.

Although results contained in this research are very promising, there is much work that can be done to further refine the methods specified in this document. Specifically, future work should include:

- This research evaluated the performance of CB-DNA Multiple Discriminant Analysis / Maximum Likelihood (MDA/ML) for Quadrature Phase Shift Keying (QPSK) signals. The methods described in this document are applicable for any In-Phase/Quadrature-Phase (I/Q) modulation scheme such as M-ary Quadrature Amplitude Modulation (M-QAM). An interesting research topic will be to implement the CB-DNA methods described in this document to a higher order modulation scheme (i.e. 16-QAM, 32-QAM, 8-PSK, etc.).

- Implement the CB-DNA MDA/ML classification algorithm to discriminate a well defined waveform such as: ZigBee, Z-Wave, etc.

- Near real time computation of CB-DNA fingerprints and MDA/ML classification using GNU-Radio and/or Field Programmable Gate Array (FPGA) implementation

- Signal watermarking implementation for higher order modulation schemes (i.e. 16-QAM, 32-QAM, 8-PSK, etc.).

The performance of the CB-DNA classification algorithm was tested in four worst-case scenarios for PUEAs: like-model devices, like-model passband components, like-model baseband components, and large number of like-model devices. The tests ex-

ceeded a mean of %C=90% correct classification rate for all test cases using CB-DNA fingerprints when $E_b/N0 \geq$24 dB. Additionally, CB-DNA fingerprints outperformed RF-DNA fingerprints in all test cases.

These experiments consider the most-challenging case because all SDR devices, baseband components, and passband components are brand new with the same manufacturer and model number. Classification results are expected to improve for SDR devices that are of a different brand or model number.

Bibliography

1. Advanced Television Systems Committee, "ATSC Digital Television Standard - Part 2: RF Transmission Systems Characteristics," 2011.

2. M. Lukacs, "Classification of antennas with mismatched loads using multiple discrimant analysis and general learning vector quantization and an ultra-wideband noise interrogation signal," *Air Force Institute of Technology*, 2014.

3. D. R. Reising, M. A. Temple, and M. J. Mendenhall, "Improving intra-cellular security using air monitoring with RF fingerprints," in *IEEE Wireless Communications and Networking Conference (WCNC)*, pp. 1–6, 2010.

4. W. M. Lowder, "Real-time RF-DNA fingerprinting of ZigBee devices using a software-defined radio with FPGA processing," Master's thesis, Air Force Institute of Technology, 2015.

5. W. E. Cobb, E. W. Garcia, M. A. Temple, R. O. Baldwin, and Y. C. Kim, "Physical layer identification of embedded devices using RF-DNA fingerprinting," in *Military Communications Conference (MILCOM)*, pp. 2168–2173, Oct 2010.

6. T. J. Carbino, M. A. Temple, and T. J. Bihl, "Ethernet card discrimination using unintentional cable emissions and constellation-based fingerprinting," in *International Conference on Computing, Networking and Communications (ICNC)*, pp. 369–373, Feb 2015.

7. R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification, 2nd edition*. Wiley-Interscience, 2000.

8. S. Pagadarai and A. M. Wyglinski, "A quantitative assessment of wireless spectrum measurements for dynamic spectrum access," in *4th International Conference on Cognitive Radio Oriented Wireless Networks and Communications (CROWNCOM)*, pp. 1–5, June 2009.

9. M. A. McHenry, P. A. Tenhula, D. McCloskey, D. A. Roberson, and C. S. Hood, "Chicago spectrum occupancy measurements & analysis and a long-term studies proposal," in *Proceedings of the First International Workshop on Technology and Policy for Accessing Spectrum (TAPAS)*, ACM, 2006.

10. R. Chen, J. M. Park, and J. H. Reed, "Defense against Primary User Emulation Attacks in Cognitive Radio Networks," *IEEE Journal on Selected Areas in Communications*, vol. 26, pp. 25–37, Jan 2008.

11. Z. Chen, T. Cooklev, C. Chen, and C. Pomalaza-Raez, "Modeling Primary User Emulation Attacks and Defenses in Cognitive Radio Networks," in *28th IEEE International Performance Computing and Communications Conference (IPCCC)*, pp. 208–215, 2009.

12. X. Xie and W. Wang, "Detecting Primary User Emulation Attacks in Cognitive Radio Networks via Physical Layer Network Coding," *Procedia Computer Science*, vol. 21, pp. 430 – 435, 2013.

13. A. Mody and G. Chouinard, "IEEE 802.22 Wireless Regional Area Networks," *IEEE 802.22-10/0073r03*, 2010.

14. S. Chen, K. Zeng, and P. Mohapatra, "Hearing is believing: Detecting mobile primary user emulation attack in white space," in *IEEE INFOCOM*, pp. 36–40, Apr 2011.

15. S. Anand, Z. Jin, and K. Subbalakshmi, "An analytical model for primary user emulation attacks in cognitive radio networks," in *3rd IEEE Symposium on New Frontiers in Dynamic Spectrum Access Networks (DySPAN)*, pp. 1–6, 2008.

16. Z. Jin, S. Anand, and K. Subbalakshmi, "Detecting Primary User Emulation Attacks in Dynamic Spectrum Access Networks," in *IEEE International Conference on Communications (ICC)*, pp. 1–5, 2009.

17. Z. Jin and S. Anand, "Mitigating primary user emulation attacks in dynamic spectrum access networks using hypothesis testing," *ACM SIGMOBILE Mobile Computing and Communications Review*, vol. 13, no. 2, pp. 74–85, 2009.

18. Z. Jin, S. Anand, and K. Subbalakshmi, "Robust Spectrum Decision Protocol against Primary User Emulation Attacks in Dynamic Spectrum Access Networks," in *IEEE Global Telecommunications Conference (GLOBECOM)*, pp. 1–5, 2010.

19. A. Alahmadi, M. Abdelhakim, J. Ren, and T. Li, "Defense Against Primary User Emulation Attacks in Cognitive Radio Networks Using Advanced Encryption Standard," vol. 9, no. 5, pp. 772–781, 2014.

20. O. Ureten and N. Serinken, "Wireless security through RF fingerprinting," *Canadian Journal of Electrical and Computer Engineering*, vol. 32, pp. 27–33, Winter 2007.

21. A. V. Oppenheim, R. W. Schafer, J. R. Buck, *et al.*, *Discrete-time signal processing*, vol. 3. Prentice Hall Englewood Cliffs, NJ, 2009.

22. D. G. Morrison, "On the Interpretation of Discriminant Analysis," *Journal of Marketing Research*, vol. 6, no. 2, pp. 156–163, 1969.

23. S. Manel, J. Dias, and S. J. Ormerod, "Comparing discriminant analysis, neural networks and logistic regression for predicting species distributions: a case study with a Himalayan river bird," *Ecological Modelling*, vol. 120, no. 2, pp. 337–347, 1999.

24. ATMEL Corporation, *AVR2015: RZRAVEN Quick Start Guide*. ATMEL Corporation, 2008.

25. T. Fawcett, "ROC graphs: Notes and practical considerations for researchers," *Machine learning*, vol. 31, pp. 1–38, 2004.

26. T. Yucek and H. Arslan, "A survey of spectrum sensing algorithms for cognitive radio applications," *IEEE Communications Surveys Tutorials*, vol. 11, no. 1, pp. 116–130, 2009.

27. R. K. Sharma and D. B. Rawat, "Advances on Security Threats and Countermeasures for Cognitive Radio Networks: A Survey," *IEEE Communications Surveys Tutorials*, vol. 17, no. 2, pp. 1023–1043, 2015.

28. R. Chen, J. Park, Y. T. Hou, and J. H. Reed, "Toward secure distributed spectrum sensing in cognitive radio networks," *IEEE Communications Magazine*, vol. 46, pp. 50–55, April 2008.

29. H. Li and Z. Han, "Dogfight in Spectrum: Combating Primary User Emulation Attacks in Cognitive Radio Systems, Part I: Known Channel Statistics," *IEEE Transactions on Wireless Communications*, vol. 9, pp. 3566–3577, November 2010.

30. H. Li and Z. Han, "Dogfight in Spectrum: Combating Primary User Emulation Attacks in Cognitive Radio Systems x2014 Part II: Unknown Channel Statistics," *IEEE Transactions on Wireless Communications*, vol. 10, pp. 274–283, January 2011.

31. S. U. Rehman, K. W. Sowerby, and C. Coghill, "Radio-frequency fingerprinting for mitigating primary user emulation attack in low-end cognitive radios," *IET Communications*, vol. 8, pp. 1274–1284, May 2014.

32. C. Zhao, W. Wang, L. Huang, and Y. Yao, "Anti-PUE Attack Base on the Transmitter Fingerprint Identification in Cognitive Radio," in *5th International Conference on Wireless Communications, Networking and Mobile Computing*, pp. 1–5, Sept 2009.

33. B. Naqvi, S. Murtaza, and B. Aslam, "A mitigation strategy against malicious Primary User Emulation Attack in Cognitive Radio networks," in *International Conference onEmerging Technologies (ICET)*, pp. 112–117, Dec 2014.

34. T. N. Le, W. L. Chin, and Y. H. Lin, "Non-cooperative and cooperative PUEA detection using physical layer in mobile OFDM-based cognitive radio networks," in *International Conference on Computing, Networking and Communications (ICNC)*, pp. 1–5, Feb 2016.

35. W. R. Ghanem, M. Shokair, and M. I. Desouky, "An improved primary user emulation attack detection in cognitive radio networks based on firefly optimization algorithm," in *33rd National Radio Science Conference (NRSC)*, pp. 178–187, Feb 2016.

36. M. Haghighat and S. M. S. Sadough, "Cooperative spectrum sensing in cognitive radio networks under primary user emulation attacks," in *6th International Symposium on Telecommunications (IST)*, pp. 148–151, Nov 2012.

37. O. R. Afolabi, K. Kim, and A. Ahmad, "On Secure Spectrum Sensing in Cognitive Radio Networks Using Emitters Electromagnetic Signature," in *Proceedings of 18th Internatonal Conference on Computer Communications and Networks (ICCCN)*, pp. 1–5, Aug 2009.

38. K. Kim, C. M. Spooner, I. Akbar, and J. H. Reed, "Specific Emitter Identification for Cognitive Radio with Application to IEEE 802.11," in *IEEE Global Telecommunications Conference (GLOBECOM)*, pp. 1–5, Nov 2008.

39. P. K. Harmer, D. R. Reising, and M. A. Temple, "Classifier selection for physical layer security augmentation in Cognitive Radio networks," in *IEEE International Conference on Communications (ICC)*, pp. 2846–2851, June 2013.

40. C. Zhao, L. Xie, X. Jiang, L. Huang, and Y. Yao, "A PHY-layer Authentication Approach for Transmitter Identification in Cognitive Radio Networks," in *International Conference on Communications and Mobile Computing (CMC)*, vol. 2, pp. 154–158, 2010.

41. M. D. Williams, M. A. Temple, and D. R. Reising, "Augmenting Bit-Level Network Security Using Physical Layer RF-DNA Fingerprinting," in *IEEE Global Telecommunications Conference (GLOBECOM)*, pp. 1–6, Dec 2010.

42. Y. Huang and H. Zheng, "Radio frequency fingerprinting based on the constellation errors," in *18th Asia-Pacific Conference on Communications (APCC)*, pp. 900–905, Oct 2012.

43. V. Brik, S. Banerjee, M. Gruteser, and S. Oh, "Wireless Device Identification with Radiometric Signatures," in *Proceedings of the 14th International Conference on Mobile Computing and Networking (MobiCom)*, pp. 116–127, ACM, 2008.

44. B. Sklar, *Digital communications*, vol. 2. Prentice Hall NJ, 2001.

45. S. Tyler and J. Loftsson, "Periodic binary sequences with very good autocorrelation properties," *Telecommunications and Data Acquisition Progress Report*, vol. 82, pp. 143–158, 1985.

46. S. U. Rehman, K. Sowerby, and C. Coghill, "Analysis of receiver front end on the performance of RF fingerprinting," in *IEEE 23rd International Symposium on*

Personal, Indoor and Mobile Radio Communications (PIMRC), pp. 2494–2499, Sept 2012.

47. H. Patel, M. A. Temple, and B. W. Ramsey, "Comparison of High-end and Low-end Receivers for RF-DNA Fingerprinting," in *IEEE Military Communications Conference*, pp. 24–29, Oct 2014.

48. N. S. Alagha, "Cramer-Rao bounds of SNR estimates for BPSK and QPSK modulated signals," *IEEE Communications Letters*, vol. 5, pp. 10–12, Jan 2001.

49. F. G. Stremler, *Introduction to communication systems*, vol. 3. Addison-Wesley Publishing Company, Reading, MA, 1990.

50. F. Rice, B. Cowley, B. Moran, and M. Rice, "Cramer-Rao lower bounds for QAM phase and frequency estimation," *IEEE Transactions on Communications*, vol. 49, pp. 1582–1591, Sep 2001.

51. M. Luise and R. Reggiannini, "Carrier frequency recovery in all-digital modems for burst-mode transmissions," *IEEE Transactions on Communications*, vol. 43, pp. 1169–1178, Feb 1995.

52. C. R. Johnson Jr, W. A. Sethares, and A. G. Klein, *Software receiver design: build your own digital communication system in five easy steps*. Cambridge University Press, 2011.

53. I. J. Cox, M. L. Miller, and A. L. McKellips, "Watermarking as communications with side information," *Proceedings of the IEEE*, vol. 87, pp. 1127–1141, Jul 1999.

54. A. Abduvaliev, S. Lee, and Y. K. Lee, "Simple hash-based message authentication scheme for wireless sensor networks," in *9th International Symposium on Communications and Information Technology (ISCIT)*, pp. 982–986, Sept 2009.

55. X. Tan, K. Borle, W. Du, and B. Chen, "Cryptographic link signatures for spectrum usage authentication in cognitive radio," in *Proceedings of the 4th Conference on Wireless Network Security (WiSec)*, pp. 79–90, ACM, 2011.

56. F. C. Huang, B. C. Wang, Y. L. Tsai, and T. H. Lin, "An energy-efficient QPSK demodulation scheme with injection-locking technique for green radio communication," in *IEEE International Conference on Internet of Things (iThings), Green Computing and Communications (GreenCom), and Cyber-Physical-Social Computing (CPSCom)*, pp. 614–617, Sept 2014.

57. X. Zhang, J. H. Lee, and M. H. Sunwoo, "Phase recovery for qpsk transmission without using complex multipliers," in *Proceedings of the 6th International Conference on Ubiquitous Information Management and Communication (ICUIMC)*, pp. 125:1–125:4, ACM, 2012.

58. H. H. Nguyen and E. Shwedyk, *A First Course in Digital Communications*. Cambridge University Press, 2009.

59. K. M. Borle, B. Chen, and W. Du, "A physical layer authentication scheme for countering primary user emulation attack," in *IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 2935–2939, May 2013.

60. A. Auger and N. Hansen, "Evolution strategies and related estimation of distribution algorithms," in *Proceedings of the 10th Annual Conference Companion on Genetic and Evolutionary Computation (GECCO)*, pp. 2727–2740, ACM, 2008.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704–0188

1. REPORT DATE *(DD–MM–YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From — To)*
15–09–2016	Doctoral Dissertation	Oct 2013 — Sep 2016

4. TITLE AND SUBTITLE

Physical Layer Defenses Against Primary User Emulation Attacks

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Betances, Joan Addison, Major, USAF

5d. PROJECT NUMBER

16G178

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Air Force Institute of Technology
Graduate School of Engineering and Management (AFIT/EN)
2950 Hobson Way
WPAFB OH 45433-7765

8. PERFORMING ORGANIZATION REPORT NUMBER

AFIT-ENG-DS-16-S-005

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Air Force Research Lab Information Directorate (RI)
525 Brooks Road
Rome Lab AFB NY 13441
DSN 587-4478
Email: michael.gudaitis@us.af.mil

10. SPONSOR/MONITOR'S ACRONYM(S)

AFRL/RI

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

DISTRIBUTION STATEMENT A:
APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

Current Cognitive Radio (CR) spectrum sensing research efforts tend to focus on the development of new mechanisms to detect Primary User (PU) or improve existing ones. However, previous researchers have identified that a Primary User Emulation Attack (PUEA) can disrupt the operation of a CR system by significantly reducing the spectrum available to unlicensed users. This dissertation presents three methods to counteract PUEAs: Radio Frequency Distinct Native Attribute (RF-DNA), Constellation-Based Distinct Native Attribute (CB-DNA), and signal watermarking. RF-DNA fingerprinting extract identifying features from Radio Frequency (RF) signals using a Region of Interest (ROI) that remains constant for all transmissions such as preambles, midambles, pilot tones, etc. CB-DNA fingerprinting uniquely identifies emissions from a radio by computing statistical features of the received signal projected into a constellation space. Finally, the signal watermarking method establishes a side-channel that enables the exchange of a Hash Based Message Authentication Code (HMAC) that authenticates the source of a signal.

15. SUBJECT TERMS

Cognitive Radio (CR), Primary User Emulation Attacks (PUEA), Radio Frequency Distinct Native Attribute (RF-DNA), Constellation Based Distinct Native Attribute (CB-DNA), Wireless Security

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Dr. Kenneth Hopkinson, AFIT/ENG
U	U	U	U	130	19b. TELEPHONE NUMBER *(include area code)* (937)255-3636; kenneth.hopkinson@afit.edu