Application of Wavelet-Based RF Fingerprinting to Enhance Wireless Network Security

Randall W. Klein, Michael A. Temple, and Michael J. Mendenhall

Abstract: **This work continues a trend of developments aimed at exploiting the physical layer of the open systems interconnection (OSI) model to enhance wireless network security. The goal is to augment activity occurring across other OSI layers and provide improved safeguards against unauthorized access. Relative to intrusion detection and anti-spoofing, this paper provides details for a proof-of-concept investigation involving "air monitor" applications where physical equipment constraints are not overly restrictive. In this case, RF fingerprinting is emerging as a viable security measure for providing device-specific identification (manufacturer, model, and/or serial number). RF fingerprint features can be extracted from various regions of collected bursts, the detection of which has been extensively researched. Given reliable burst detection, the near-term challenge is to find robust fingerprint features to improve device distinguishability. This is addressed here using wavelet domain (WD) RF fingerprinting based on dual-tree complex wavelet transform (DT-$\mathbb{C}$WT) features extracted from the non-transient preamble response of OFDM-based 802.11a signals. Intra-manufacturer classification performance is evaluated using four like-model Cisco devices with dissimilar serial numbers. WD fingerprinting effectiveness is demonstrated using Fisher-based multiple discriminant analysis (MDA) with maximum likelihood (ML) classification. The effects of varying channel SNR, burst detection error and dissimilar SNRs for MDA/ML training and classification are considered. Relative to time domain (TD) RF fingerprinting, WD fingerprinting with DT-$\mathbb{C}$WT features emerged as the superior alternative for all scenarios at SNRs below 20 dB while achieving performance gains of up to 8 dB at 80% classification accuracy.**

Index Terms: **Complex wavelet transform (CWT), dual-tree, intrusion detection, multiple discriminant analysis (MDA), physical layer, RF fingerprinting, wavelet transform, wireless security.**

I. INTRODUCTION

The work in [1]–[3] is representative of research that has been conducted to address network safeguards spanning the application and data link layers of the open systems interconnection (OSI) model. There has also been significant earlier work [4]–[10] and more recent work [11]–[16] where researchers have focused on the OSI physical (PHY) layer to exploit inherent RF features that are device dependent and difficult for unauthorized users to replicate. The methods employed for these PHY-based approaches is as varied as the signal attributes they have exploited. That is, there are differences in 1) operational standards (802.11 and 802.15) which dictate different signal modulations (DSSS, OFDM, etc.), 2) signal regions exploited (transient and non-transient), 3) collection system capability (bandwidth, sample rate, collection interval, etc.), and 4) experimental set-up and/or environmental conditions (anechoic chamber, typical office, etc.). These differences collectively impact the degree to which inherent RF features will vary and make PHY-based security much more challenging than other bit-level approaches being considered in other OSI model layers. Furthermore, these differences make it very difficult to perform reliable quantitative cross-comparison of PHY-based methods being investigated. Thus, a brief discussion follows on activities that are most related to the work presented here to enable proper assessment of the proof-of-concept investigation.

The statistics of power-based received signal strength (RSS) measurements have been considered and provide some anti-spoofing protection based on PHY layer attributes [11], [12]. While the authors in [12] dismiss RF fingerprinting as a viable alternative for "scale" reasons, it remains a viable alternative for less constrained air monitoring applications and has been successfully demonstrated using *transient* responses from DSSS-based IEEE 802.15 compliant signals [13], [16] and with *non-transient* responses from OFDM-based IEEE 802.11 compliant signals [14], [15]. The works in [13]–[16] do use a common Fisher-based approach for feature generation and/or device discrimination. Collectively, these earlier works have shown that inherent signal features can be used to form RF fingerprints that are repeatedly extractable and sufficiently unique to enable device specific identification, to include distinguishing between manufacturers, model numbers and/or serial numbers.

Signal structure uniqueness is generally attributable to differences in device manufacturing processes, component tolerances, material properties and environmental factors [3], [4], [10]. If sufficiently unique, this structure may be exploited to uniquely identify devices based on their RF fingerprints. Burst detection is arguably the most important step in the fingerprinting process and has been extensively researched [6], [8], [17], [18]. Subsequent signal region(s) selection for fingerprint extraction is of near equal importance given that fingerprint robustness is desired amidst burst detection error and imprudent signal region selection can adversely bias processing in favor of channel noise or undesired signal features [4]. Fingerprint classification sensitivity to burst detection error and channel noise variation has been previously addressed for 802.11a signals [14], [15], [17], [18]. These works showed that variance trajectory (VT) burst detection enables classification performance that is consistent with "perfect" burst estimation and that denoising low SNR signals with a dual-tree complex wavelet transform (DT-$\mathbb{C}$WT) enhanced overall performance.

Manuscript received April 18, 2009.

This work was supported by the US Air Force Research Laboratory and the Tactical SIGINT Technology (TST) Program.

The authors are with the Air Force Institute of Technology, Wright-Patterson AFB, Ohio 45433, USA, email: randyklein99@gmail.com, {michael.temple, michael.mendenhall}@afit.edu.

Given demonstrated detection capability, the near-term challenge is to improve classification performance by finding improved fingerprint features. This challenge is addressed here using wavelet domain (WD) RF fingerprinting based on DT-ℂWT features. The impact of using WD fingerprints is first addressed using "perfect" burst estimation and Fisher-based multiple discriminant analysis (MDA) with maximum likelihood (ML) classification. Earlier work in [14], [15], [17], [18] addressed *inter-manufacturer* device discrimination using time domain (TD) fingerprinting with experimentally collected 802.11a signals from Cisco and Dell devices. The work presented here addresses a fundamentally more difficult problem, i.e., *intra-manufacturer* (serial number) discrimination using identical model devices manufactured by Cisco.

The choice of using OFDM-based signals for a proof-of-concept investigation, and specifically the 802.11a signal, was driven by several important factors, including: 1) Consistency with previous related work [14], [15], [17], [19]–[21], 2) the ability to directly compare TD processing and results from some of these earlier works with new WD results, and 3) the continued emergence of OFDM-based signals as envisioned for future 4G software defined/cognitive radio (SDR/CR) communications [22]–[24]. While the fingerprint and classification techniques used here are likely applicable to other signal types, and may actually perform better with some of them, the challenges posed by OFDM-based signals must be addressed.

II. Background

A. RF Fingerprint Classification

There has been considerable work in previous years involving the exploitation of RF signal characteristics to classify signals and identify the devices producing them [4], [6]–[8], [14], [15]. Collectively, these works embody the field of RF fingerprint classification which fundamentally requires two processes, including: 1) Fingerprint generation and 2) fingerprint classification. Fingerprint generation requires the selection and extraction of features that enable signal/device discrimination. Desirable properties of the selected feature set include: 1) Reduced dimensionality to minimize processing and storage requirements, 2) intra-device repeatability, and 3) inter-device uniqueness. For this work, the classification features are statistics of instantaneous signal characteristics per the details provided in Sections II-A.1 and II-A.2. The resultant RF statistical fingerprints are then used for signal/device classification per the details provided in Section II-A.3.

A.1 Fundamental Signal Characteristics

Various signal characteristics can be exploited to provide device identification, with some of the earlier works predominantly focusing on instantaneous amplitude and instantaneous phase [4], [6]–[8]. More recently, subsequent research has successfully exploited instantaneous frequency as well [14], [15], [17], [18]. A complex sampled time domain (TD) signal of the form

$$s_{TD}(n) = I_{TD}(n) + jQ_{TD}(n) \tag{1}$$

has instantaneous amplitude, $a(n)$, phase, $\phi(n)$, and frequency, $f(n)$, characteristics given by

$$a\left(n\right) = \sqrt{I_{TD}^2\left(n\right) + Q_{TD}^2\left(n\right)}, \tag{2}$$

$$\phi\left(n\right) = \tan^{-1}\left[\frac{Q_{TD}\left(n\right)}{I_{TD}\left(n\right)}\right], \tag{3}$$

$$f(n) = \frac{1}{2\pi}\frac{\phi(n) - \phi(n-1)}{\Delta n} \tag{4}$$

where $I_{TD}(n)$ and $Q_{TD}(n)$ are the instantaneous in-phase and quadrature-phase components of $s_{TD}(n)$.

In practice, each characteristic response is "centered" (mean removed) to remove collection system biases that may unduly influence subsequent processing. The instantaneous amplitude and frequency responses are simply centered using

$$a_c(n) = a(n) - \mu_a, \tag{5}$$

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

where $n = 1, 2, \cdots, N_M$, N_M is the total number of samples in the collected signal, and μ_a and μ_f are amplitude and frequency means calculated across N_M samples of (2) and (4), respectively.

Given the phase response in (3), a linear component is first removed prior to centering. This component may be due to collection receiver coloration or result from inexact frequency estimation during post-collection down-conversion. The resultant *non-linear phase* response is given by

$$\phi_{nl}(n) = \phi(n) - 2\pi\mu_f(n)\Delta_t \tag{7}$$

where μ_f is the same frequency mean used in (6) and Δ_t is the time sample spacing. As a final step, the mean of ϕ_{nl} is removed to yield the desired *centered non-linear* phase which is given by

$$\phi_{cnl}(n) = \phi_{nl}(n) - \mu_{\phi_{nl}} \tag{8}$$

where $\mu_{\phi_{nl}}$ is the mean across N_M samples of $\phi_{nl}(n)$ in (7). The centering of signal characteristics in (5)–(8) is consistent with previous device classification work that successfully employed similar procedures [14], [15], [17], [18].

A.2 Basic Statistical Features

Direct use of the fundamental signal characteristics in Section II-A.1 as classification features can be prohibitive in terms of data storage and computational processing time. The computational burden can be eased by reducing the feature space dimensionality used for device classification. This approach was used in [14], [15], [17], and [18] where the inherent statistical properties of the fundamental signal characteristics were exploited for device classification. The statistics of interest in this earlier work included the variance (σ^2), skewness (γ), and kurtosis (κ), defined as:

$$\sigma_x^2 = \frac{1}{N_x}\sum_{k=1}^{N_x}\left[x(k) - \bar{x}\right]^2, \tag{9}$$

$$\gamma_x = \frac{1}{\sigma_x^3 N_x}\sum_{k=1}^{N_x}\left[x(k) - \bar{x}\right]^3, \tag{10}$$

$$\kappa_x = \frac{1}{\sigma_x^4 N_x} \sum_{k=1}^{N_x} [x(k) - \bar{x}]^4 \tag{11}$$

where $\bar{x}$ is the mean of $\{x(k)\}$. The final RF fingerprints are formed by calculating these statistics for the appropriate centered instantaneous signal characteristic(s) in Section II-A.1, i.e., setting $\{x(k)\}$ equal to $\{a_c(n)\}$ with elements from (5), setting $\{x(k)\}$ equal to $\{f_c(n)\}$ with elements from (6), and/or setting $\{x(k)\}$ equal to $\{\phi_{cnl}(n)\}$ with elements from (8).

A.3 MDA/ML Classification

While many different techniques are available for classification, they all employ two fundamental processes: Training and classification. That is, they *train* the classifier using a subset of the input data and then *classify* using the remaining data. For the most part, these techniques are oblivious to what the input data actually represents and their performance is predominantly driven by the statistical behavior of the data. With regard to RF fingerprint classification, there has been little novelty in developing specialized classification techniques and most researchers have opted for well-known techniques. The predominant techniques of choice have been based on neural networks [9], [20], [25]–[31], with some limited additional work based on Kalman filtering and/or a Hotelling statistic [32], [33].

Multiple discriminant analysis (MDA) with maximum likelihood (ML) estimation is a viable classification alternative that has been successfully used for TD RF fingerprint classification [15], [17], [18]. MDA is an extension of Fisher's linear discriminant (FLD) process for more than two classes [35]. For a 3-class problem, the Fisher-based MDA process projects higher-dimensional data onto a 2-dimensional "Fisher plane" that maximizes inter-class distances while simultaneously minimizing intra-class distances. In principle, this method cannot improve classification potential. However, it provides good class separation and visualization of data having input dimensionality greater than three. Using this lower-dimensional data, decision boundaries calculated from ML distributions are determined assuming normally distributed input data, equal costs and uniform prior probabilities. In general, to discriminate c classes using d-dimensional input data, the input vector $\mathbf{x}$ is linearly projected onto a $(d-1)$-dimensional space using

$$\mathbf{y} = \mathbf{W}^{\mathbf{T}}\mathbf{x} \tag{12}$$

where $\mathbf{y}$ is the vector of projected values and $\mathbf{W}$ is a $d \times (c-1)$ projection matrix. Classification is performed using unknown data and the trained 2-dimensional decision boundaries calculated from ML distributions. The process classifies each "unknown" input data set by projecting it onto the trained Fisher plane according to (12). Projected points falling within the correct region are correctly classified while those falling outside the correct region are misclassified. The percentage of correct classification is determined based on the total number of unknown trials. A more complete description of the MDA/ML process is provided in [36].

B. Dual-Tree Complex Wavelet Transform (DT-ℂWT)

Device classification can be performed using a discrete wavelet transform (DWT), with one popular method using a subset of the largest DWT coefficient magnitudes as the classification features [37]. One distinct disadvantage of DWT-based approaches is that the DWT is not shift invariant. As with signal denoising, this presents a problem for RF fingerprinting applications given that robust classification performance relies on the fingerprint features being unique, repeatable and stable. These properties cannot be assured if the underlying features (DWT coefficients) vary dramatically throughout the processing interval of interest. For example, variation in burst detection and start location error generally translates to greater variation in fingerprint features. The DT-ℂWT is used to address the lack of shift invariance in DWT processing.

The DT-ℂWT is a DWT extension that is "nearly shift-invariant," i.e., the DT-ℂWT coefficients are independent of time domain shift and more strongly dependent on interscale and intrascale neighborhoods [38]. This shift invariance has been previously exploited to improve classification performance for hyperspectral images [39]. Furthermore, the DT-ℂWT magnitude response exhibits reduced ringing that is generally induced by high-frequency noise and sharp discontinuities [38].

The DT-ℂWT is commonly implemented using two real-valued filter banks. These are denoted as *Tree1* and *Tree2* in Fig. 1 which shows one common architecture for DT-ℂWT implementation [34]. The scaling and wavelet functions for *Tree1* are symmetric (even functions) while *Tree2* has scaling and wavelet functions that are anti-symmetric (odd functions). The wavelet and scaling functions, $\psi(t)$ and $\phi(t)$ respectively, for the *Tree1* filter bank are given by [34], [38]

$$\psi(t) = \sqrt{2}\sum_n h_1(n)\phi(2t-n), \tag{13}$$

$$\phi(t) = \sqrt{2}\sum_n h_0(n)\phi(2t-n) \tag{14}$$

where the filter coefficients $h_1(n)$ and $h_0(n)$ are implemented directly as the analysis filters (AF) given in [40]. Ideally, the corresponding functions for the *Tree2* filter bank are the Hilbert transforms of (13) and (14), expressed as

$$\psi'(t) = \sqrt{2}\sum_n h'_1(n)\phi'(2t-n), \tag{15}$$

$$\phi'(t) = \sqrt{2}\sum_n h'_0(n)\phi'(2t-n) \tag{16}$$

where the filter coefficients $h'_1(n)$ and $h'_0(n)$ are implemented directly as the analysis filters (AF) given in [40]. As shown in Fig. 1, the first stage filters for both *Tree1* and *Tree2* have different coefficients when compared to the later stage filters and are denoted as $h_1^{(1)}(n)$, $h_0^{(1)}(n)$, $h_1'^{(1)}(n)$, and $h_0'^{(1)}(n)$, respectively. The first stage filter coefficients are implemented directly as the first analysis filters (FAF) given in [40].

For real-valued input signals, the filter bank outputs in Fig. 1 are real-valued wavelet domain (WD) coefficients representing real (I_{WD}^l) and imaginary (Q_{WD}^l) components of complex coefficients [38]. These components can be functionally combined in a form similar to (1) and expressed as

$$s_{WD}^l(n) = I_{WD}^l(n) + jQ_{WD}^l(n)\ . \tag{17}$$

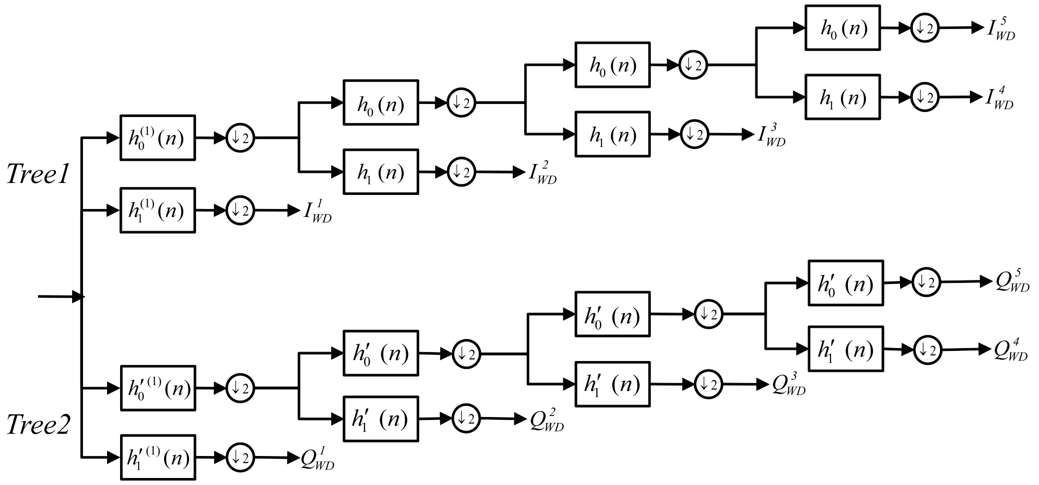


Fig. 1. Four stage (five level) dual-tree complex wavelet transform (DT-ℂWT) [34].

Using $s^l_{WD}(n)$ elements from (17), the sequence $\{s_{WD}(n)\}$ of all elements can be interpreted as what may be called a "complex sampled WD signal." Given the similar structure of this WD signal and the TD signal in (1), WD fingerprint classification can be performed using the process in Sec. II-A. In this case, the WD signal in (17) can be used in (2)–(8) to generate WD signal characteristics and statistics calculated per (9)–(11) to form statistical WD fingerprints.

III. Methodology

A. Overall Demonstration Process

All results presented in Section IV were generated using the overall demonstration process illustrated in Fig. 2. The dashed boundaries delineate primary hardware and software processes. The "Signal Collection" hardware process consisted of placing communication devices (source and destination laptops with 802.11a PCMCIA cards) and the agilent-based RF signal intercept and collection system (RFSICS) in a chamber and making free-space signal collections. The collected signal data (a series of complex valued samples) was passed along for subsequent Post-Collection Processing which was accomplished exclusively in a MATLAB® environment. The implementation and functionality of various processes in Fig. 2 are discussed in the following sections.

B. Signal Collection Process

Collections were made with the source laptop (containing the device under test), destination laptop and the RFSICS in an anechoic chamber. Chamber collections were specifically chosen for proof-of-concept investigation to help isolate device specific hardware effects from channel/propagation effects and thereby enable appropriate attribution of classification differences to device differences. Thus, the collections included minimal to no multipath propagation effects and the impact of varying source-to-destination distance (range) is adequately captured and incorporated through SNR variation.

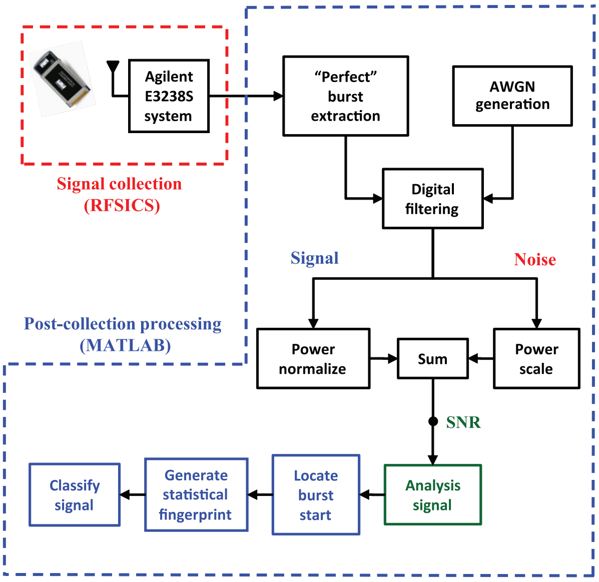


Fig. 2. Overall process for signal collection, analysis signal generation, burst detection and start location, fingerprint extraction, and classification.

A given device under test (Cisco PCMCIA card) was placed in the source laptop and a continuous file transfer initiated between it and the destination laptop. The destination laptop was configured such that it transmitted at minimum power to enable reliable post-collection separation of desired source and undesired destination bursts. For subsequent collections, the positions and orientations of the laptops and RFSICS were maintained and alternate devices inserted into the source laptop.

Basic RFSICS functionality is provided by Agilent's E3238S system [41] and includes an RF collection range of 20.0 MHz to 6.0 GHz. The band of interest is selected using a tunable RF filter with fixed bandwidth of 36.0 MHz. The selected RF band is down-converted to an intermediate frequency (IF) of 70.0 MHz and passed to a digitizer. The digitizing process consists of down-conversion (near baseband), 12-bit analog-to-digital conversion at 95 M samples-per-second (sps), digital filtering (user defined bandwidth), Nyquist compliant sub-sampling, and data storage as complex in-phase (I) and quadrature (Q) components. A digital filter bandwidth of 18.56 MHz was selected for all 802.11a signals collected for this work. This resulted in the RFSICS automatically applying a sub-sampling factor of four, for a final sample rate of $f_s = 23.75$ Msps and corresponding sample interval of $T_s = 1/f_s \approx 42.1$ nsec per sample. The typical *collected* SNR for the chamber collected signals is on the order of $SNR = 40$ dB.

C. Post-Collection Processing

C.1 Analysis Signal Generation

The first post-collection process of "perfect" burst extraction uses the near-baseband, complex I-Q data from the RFSICS collections. Extraction is accomplished through a combination of automated amplitude threshold detection followed by visual analysis and manual alignment to accurately identify the sample number corresponding to the burst start. The extracted burst responses are digitally filtered using a baseband filter and power-normalized. A 6th-order Chebyshev digital filter was implemented having a –3 dB bandwidth of 7.7 MHz. This bandwidth was experimentally chosen as it provided a maximum observed classification performance for TD fingerprints but a near-maximum performance for WD fingerprints, ensuring a bias towards the TD technique. This particular bandwidth choice gives the TD technique an approximate 2% advantage in device classification. This will be considered when presenting, comparing and analyzing subsequent results.

At this point, the sample frequency of the filtered signal is $f_s = 23.75$ Msps which effectively represent oversampling by a factor of approximately 1.5 times Nyquist. Provided that the RFSICS collection and subsequent post-processing is identical for all signals, it is reasonable to assume that "recording coloration" (variation in amplitude, phase and/or frequency characteristics) induced by the RFSICS and post-processing prior to burst start location, statistical fingerprint generation and signal classification is approximately identical. This is important in the overall process and ensures that final results are based on as received signal characteristics and features versus being unduly influenced by signal-dependent collection and post-processing coloration.

The desired "analysis signal" is intended to simulate varying SNR conditions that typically exist in an operational environment and is analogous to varying collection range and/or simulating intentional/unintentional jamming. This signal is generated by adding like-filtered, power-scaled noise to the digitally filtered, power-normalized signal. This is done by generating random complex AWGN that is filtered using the same digital filter as used for the signal. The filtered noise signal is then power-scaled to achieve the desired analysis SNR when added to the filtered signal.

C.2 Burst Detection and Start Location

To isolate the effects of using different feature sets from the effects of burst detection and location error, the RF bursts were visually detected and their "perfect" starting location (sample number) determined. This number was used to locate the preamble region for fingerprint extraction.

To further investigate TD and WD feature sensitivity to burst detection error, a random error was introduced into the perfect starting locations on a burst-by-burst basis by comparing sample numbers of the -3 dB threshold detected bursts and the corresponding manually detected perfect bursts. This produces what is referred to here as randomly "jittered" burst detection data. In this case, this error was generated using statistics from the histogrammed observed location error. Based on statistics (mean, standard deviation, skewness, and kurtosis), a four-parameter discrete Beta distribution generator was created to provide simulated detection error similar to what was observed. The random jitter error was applied to perfect burst location data prior to extracting the fingerprints used for both training and classification. This was functionally implemented in Step 2 of the classification process described in Section III-D.

C.3 Statistical Fingerprint Generation

Following burst detection and start location, the RF statistical fingerprints are generated using the process shown in Fig. 3. As indicated within the dashed lines, the Characteristics and Statistics generating functions are identical for both the time domain (TD) and wavelet domain (WD) techniques. A signal region of interest is selected from the input analysis signal and parsed into a predefined number of subregions for fingerprint generation. As illustrated in Fig. 4, the region of interest here is the 802.11a preamble response which is parsed into $N_r = 3$ subregions. This choice was based on 1) previous works which successfully exploited the preamble [14], [15], and 2) the preamble bit sequence being identical for all bursts per the 802.11 standard [42]. Fig. 4 shows that the standard modulated preamble response is comprised of 10 short and 2 long OFDM symbols. Fingerprint features were extracted from three different defined regions that included: 1) The first 8.0 μsec (10 short OFDM symbols), 2) the last 8.0 μsec (2 long OFDM symbols), and 3) the entire 16.0 μsec preamble (all short and long symbols).

For TD classification, the centered subregion characteristics are calculated using (2)–(8) and statistical classification features calculated using (9), (10), and (11) for each resultant characteristic response. The resultant TD RF fingerprint (feature vector) consists of 27 total features per collected burst (3 subregions × 3 signal characteristics × 3 statistics). The TD fingerprint for burst b, from device (class) c, in subregion r is given by

$$\begin{aligned}\mathbf{F_r^{b,c}} = [\, &\sigma_r^2(a),\ \sigma_r^2(\phi),\ \sigma_r^2(f),\\ &\gamma_r(a),\ \gamma_r(\phi),\ \gamma_r(f),\\ &\kappa_r(a),\ \kappa_r(\phi),\ \kappa_r(f)\,]\end{aligned} \tag{18}$$

where $b = 1, 2, \cdots, N_b$ (N_b total bursts), $r = 1, 2, \cdots, N_r$ (N_r total subregions), and $c = 1, 2, 3$ is the class index. Considering $N_r = 3$ subregions as used here, the composite TD feature vector (1 × 27) is formed using (18) and given by

$$\mathbf{F_{TD}^{b,c}} = \left[\mathbf{F_1^{b,c}}\ \mathbf{F_2^{b,c}}\ \mathbf{F_3^{b,c}}\right]\ . \tag{19}$$

For WD classification, the processing is identical to TD processing except that a DT-$\mathbb{C}$WT decomposition is performed in

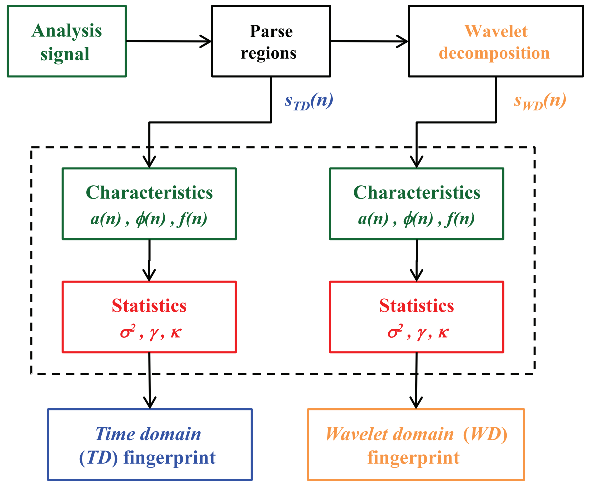

Analysis signal
Parse regions
Wavelet decomposition
$s_{TD}(n)$
$s_{WD}(n)$
Characteristics
$a(n)$, $\phi(n)$, $f(n)$
Characteristics
$a(n)$, $\phi(n)$, $f(n)$
Statistics
σ^2, γ, κ
Statistics
σ^2, γ, κ
Time domain (TD) fingerprint
Wavelet domain (WD) fingerprint

Fig. 3. Generation process for statistical RF fingerprints. The characteristics and statistics generating functions are identical for both the TD and WD techniques and implemented using (2)–(8) and (9)–(11), respectively [18].

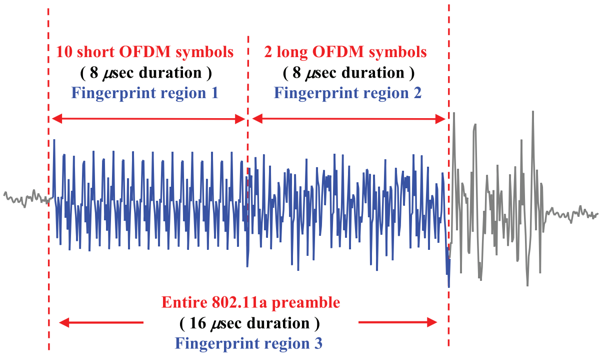


Fig. 4. 802.11a preamble response showing fingerprint regions.

each subregion. As depicted in Fig. 1, the DT-ℂWT decomposes each subregion into five levels associated with different wavelet scales. The "complex WD signal" samples are calculated using (17), followed by characteristic generation and centering using (2)–(8). The statistical classification features are calculated using (9), (10), and (11). The resultant WD RF fingerprint (feature vector) consists of 135 total features per collected burst (3 subregions × 5 DT-ℂWT decomposition levels per subregion × 3 signal characteristics × 3 statistics). Paralleling the TD development, the WD fingerprint for burst b, from device c, in subregion r which has been decomposed into l DT-ℂWT levels is given by

$$\mathbf{F}^{\mathbf{b,c}}_{\mathbf{r,l}} = [\,\sigma^2_{r,l}(a),\ \sigma^2_{r,l}(\phi),\ \sigma^2_{r,l}(f),\ \gamma_{r,l}(a),\ \gamma_{r,l}(\phi),\ \gamma_{r,l}(f),\ \kappa_{r,l}(a),\ \kappa_{r,l}(\phi),\ \kappa_{r,l}(f)\,] \tag{20}$$

where $l = 1, 2, \cdots, N_l$ with N_l being the total number of DT-ℂWT decomposition levels per subregion. Considering $N_r = 3$ subregions with $N_l = 5$ levels as used here, the composite WD classification feature vector (1 × 135) is formed using (20) and

Table 1. Cisco device serial numbers and permutations used for MDA/ML classification.

	Serial number			
Perm	N4U9	N4UD	N4UW	N4PX
1	×	×	×	
2		×	×	×
3	×		×	×
4	×	×		×

is given by

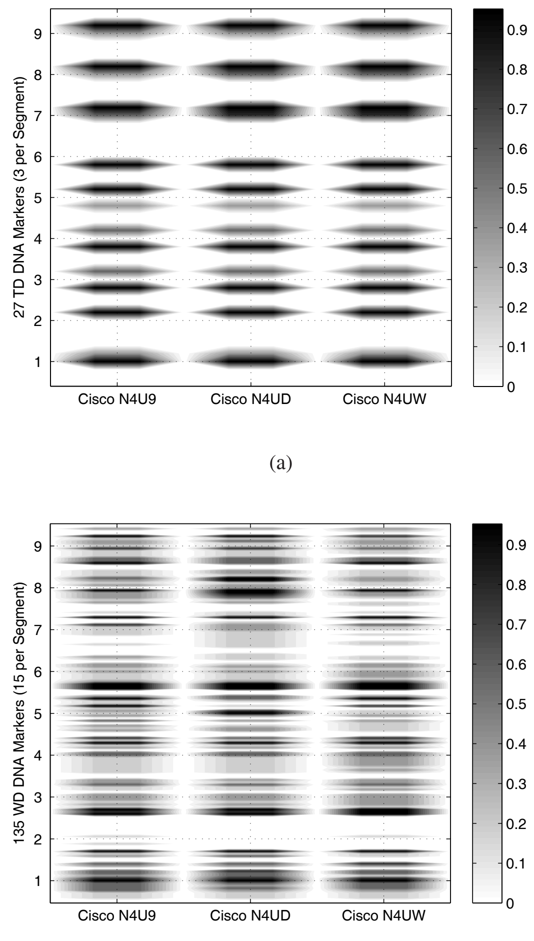
$$\mathbf{F}^{\mathbf{b,c}}_{\mathbf{WD}} = \Big[\mathbf{F}^{\mathbf{b,c}}_{\mathbf{1,1}}\,\mathbf{F}^{\mathbf{b,c}}_{\mathbf{1,2}}\,\mathbf{F}^{\mathbf{b,c}}_{\mathbf{1,3}}\,\mathbf{F}^{\mathbf{b,c}}_{\mathbf{1,4}}\,\mathbf{F}^{\mathbf{b,c}}_{\mathbf{1,5}}\ \mathbf{F}^{\mathbf{b,c}}_{\mathbf{2,1}}\,\mathbf{F}^{\mathbf{b,c}}_{\mathbf{2,2}}\,\mathbf{F}^{\mathbf{b,c}}_{\mathbf{2,3}}\,\mathbf{F}^{\mathbf{b,c}}_{\mathbf{2,4}}\,\mathbf{F}^{\mathbf{b,c}}_{\mathbf{2,5}}\ \mathbf{F}^{\mathbf{b,c}}_{\mathbf{3,1}}\,\mathbf{F}^{\mathbf{b,c}}_{\mathbf{3,2}}\,\mathbf{F}^{\mathbf{b,c}}_{\mathbf{3,3}}\,\mathbf{F}^{\mathbf{b,c}}_{\mathbf{3,4}}\,\mathbf{F}^{\mathbf{b,c}}_{\mathbf{3,5}}\Big]. \tag{21}$$

The uniqueness of fingerprint statistical features can be illustrated using so called "distinct native attributes" (DNA) in RF DNA Fingerprint plots such as shown in Fig. 5. These plots were generated by randomly selecting 250 collected bursts for each device, scaling them to achieve SNR = 20 dB, and averaging the corresponding statistical fingerprints from (19) or (21) as appropriate. For visual clarity, the average fingerprint features are normalized within each of the nine segments (3 signal characteristics × 3 statistics). Each TD fingerprint segment includes 3 markers (one for each signal subregion shown in Fig. 4). Each WD fingerprint segment includes 15 markers (3 signal subregions × 5 DT-ℂWT levels). Given that normalization has been applied within markers, caution must be exercised when comparing 1) across TD and WD responses in Fig. 5, and 2) across markers of a given device for a given technique. It is reasonable to compare behavior across devices using a given technique and a given marker. In this case, the cross-device differences are indicative of potential discriminability with greater differences corresponding to increased class separability using MDA/ML processing.

D. MDA/ML Signal Classification

Device classification is performed using statistical fingerprints with the Fisher-based MDA/ML process described in Section II-A.3. For all MDA/ML classification results presented, a total of $N_b = 2000$ bursts were used from $N_d = 3$ different Cisco devices according to the permutations in Table 1. For each permutation in Table 1, $N_b = 2000$ fingerprints for each device were used to form a composite fingerprint matrix for MDA/ML classification. Each row of the composite matrix represents one statistical fingerprint generated using (19) for TD fingerprinting or (21) for WD fingerprinting. Thus, the resultant composite fingerprint matrix for each device has dimension 2000 × 27 for TD fingerprinting and 2000 × 135 for WD fingerprinting.

Monte Carlo simulation and K-fold cross validation processes are used with MDA/ML signal classification. Monte Carlo simulation is used to ensure statistical significance and K-fold cross validation is used to generalize the prediction error to an independent data set [43]. While the required value of K can vary as a function of data "behavior," values of $K = 5$



(a)

(b)

Fig. 5. Average RF DNA fingerprints for (a) TD (27-Feature TD fingerprinting) and (b) WD processing based on 250 randomly selected bursts at SNR = 20 dB (135-Feature WD fingerprinting).

and $K = 10$ are common choices for cross validation [43]. Using $K = 5$ with $N_b = 2000$ bursts (fingerprints) per device, the input fingerprints are partitioned into $K = 5$ equal subsets (400 each), with $K - 1 = 4$ subsets (1600 fingerprints) used for training and the remaining "held-out" subset (400 fingerprints) used for classification [43].

The overall process for MDA/ML classification with K-fold cross validation is shown in Fig. 6. Accounting for a total of N_{MC} independent Monte Carlo noise realizations, the process for generating average classification results includes the following steps. Note that the Fold Iteration Accumulator in Fig. 6 is cleared prior to the start of this process.

1. Generating the analysis signal for a given SNR per Section III-C.1
2. Performing burst detection and start location per Section III-C.2
3. Generating statistical fingerprints per Section III-C.3 for the technique under evaluation (TD or WD)
4. Generating projection matrix $\mathbf{W}$ per (12) using $K - 1 = 4$ subsets (80% of the fingerprints) from each device for training and ML classifier parameter calculation

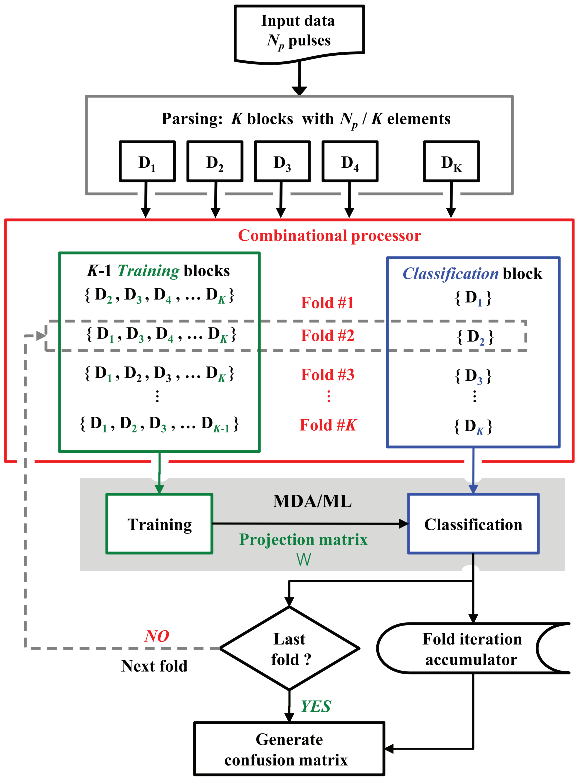
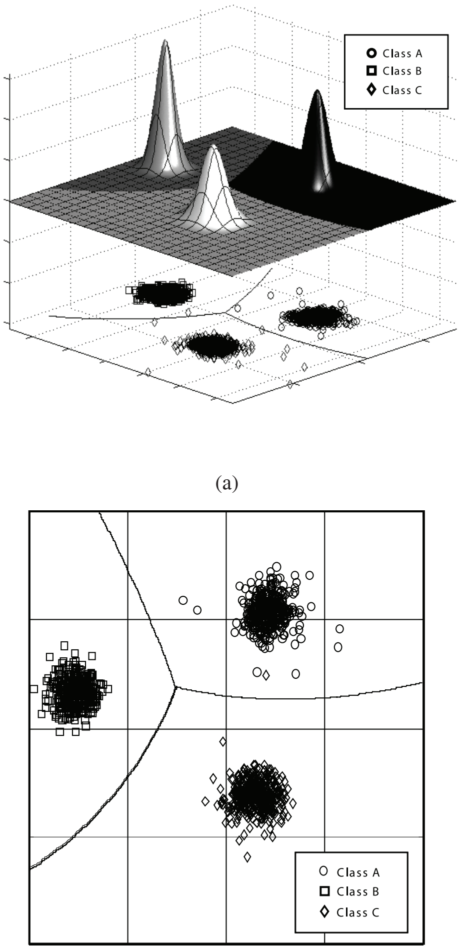


Fig. 6. MDA/ML classification process with K-fold cross validation.

5. Transforming the "held-out" subset (20% of the fingerprints) from each device as "unknown" inputs using $\mathbf{W}$ and classifying each per ML criteria
6. Accumulating the current fold classification results
7. Selecting the next $K - 1 = 4$ blocks for the next fold
8. Repeating Step 4–Step 7 for $K - 1 = 4$ additional folds
9. Repeating Step 1–Step 8 a total of N_{MC} times using different independent AWGN realizations for each iteration (fold iteration accumulator not cleared)
10. Averaging fold iteration accumulator results to obtain average classification performance (Accounting for all factors, the final average is based on a total of $N_{MC} \times N_b \times 3$ independent classification decisions.)
11. Repeating the process for each desired analysis SNR

Representative MDA-transformed training fingerprints and trained decision boundaries calculated from ML distributions are shown in Fig. 7(a) for 802.11a signals at $SNR = 40$ dB. The corresponding projection of "unknown" MDA-transformed fingerprints are shown in Fig. 7(b) overlayed with trained decision boundaries from Fig. 7(a). Note that even under these high SNR conditions incorrect classification is possible. For example, one of the Class C ($*$ markers) fingerprints is clearly projected into the Class A (× markers) ML decision region and would be incorrectly classified.

Confidence intervals provide one means for declaring statistically significant differences and/or similarities when comparing alternatives, e.g., TD versus WD performance for a given scenario. All comparative conclusions drawn in Section IV are based on estimated classification accuracy $\hat{p}$ with CI = 95% con-



(a)

(b)

Fig. 7. Representative MDA/ML (a) *Training* and (b) *Classification* for 802.11a signals at $SNR = 40$ dB: (a) MDA/ML training: ML decision boundaries and (b) MDA/ML classification: Projected fingerprints.

fidence intervals given by [44]

$$\text{CI} = \pm\, 1.96 \sqrt{\frac{\hat{p}\,(1-\hat{p})}{N_p}} \tag{22}$$

where $\hat{p}$ is calculated as the number of correct classification decision divided by N_p independent trials. All results in Section IV were generated using $N_b = 2000$ bursts per device and $N_{MC} = 50$ Monte Carlo iterations of the MDA/ML process. Thus, there are a total of $N_p = 2000 \times 50 = 100000$ independent classification decisions made per device, i.e., percentages in each row of the MDA/ML classification confusion matrices are based on 100000 trials. Note that resultant confidence intervals for given scenarios are intentionally omitted from tabular and plotted results in Section IV given that 1) they are not typically provided in confusion matrix representations (convention), 2) they are very small for a majority data points in a given scenario and tend to obscure/blurr marker discrimination (visual clarity), and 3) the focus of this work is on general revelation and demonstration versus precise assessment for a particular set of conditions and/or parameters (reliable trend analysis is sufficient). For completeness, numerical CI values are provided in Table 2 to enable detailed assessment if desired.

Table 2. Numerical CIs as a function of classification accuracy $\hat{p}$ based on $N_p = 100000$ independent trials.

Classification Accuracy $\hat{p}$	± CI ($\times 10^3$)
0.0	0.00
0.1	1.90
0.2	2.53
0.3	2.90
0.4	3.10
0.5	3.16
0.6	3.10
0.7	2.90
0.8	2.53
0.9	1.90
1.0	0.00

IV. RESULTS AND ANALYSIS

For comparative analysis, results were generated using TD and WD fingerprints generated from *identical* collected signals with *identical* Monte Carlo noise realizations that were appropriately filtered and scaled to achieve desired analysis SNRs. This enables reliable one-to-one comparison of TD and WD classification based on 1) CI $= 95\%$ confidence intervals and 2) a performance "gain" metric which is defined as the difference in required SNR ($SNR^{dB}_{WD} - SNR^{dB}_{TD}$) at a given classification accuracy level. For tracking relative TD and WD performance improvement and/or degradation, the performance gain at 80% classification accuracy was arbitrarily chosen to illustrate trends across the various scenarios considered.

A. MDA/ML Device Classification

Serial number discrimination is demonstrated using all Cisco device permutations in Table 1. Sensitivity to serial number variation is illustrated in Fig. 8 which shows average classification results for all permutations. The mean across all four permutations is shown by the filled markers. The results for both TD and WD techniques show that Permutation #1 and Permutation #3, which both include Cisco devices with serial numbers N4U9 and N4UW, yield the poorest results for nearly all SNR values considered. Permutation #1 is the "most stressing" case for MDA/ML classification and per Table 1 includes three devices with serial numbers that differ in only the last digit. While not verifiable, it is assumed that these devices have been manufactured using identical components, from identical lots, with identical processes, under identical environmental conditions. Thus, discriminating between these devices presents the most stressing case for classification.

The mean classification results in Fig. 8 are presented again in Fig. 9 for closer inspection. Based on CI $= 95\%$ confidence intervals, both techniques perform statistically similar for SNR ≥ 25 dB and the WD fingerprinting technique is superior for $-2 <$ SNR < 24 dB. The WD fingerprints achieve 80%

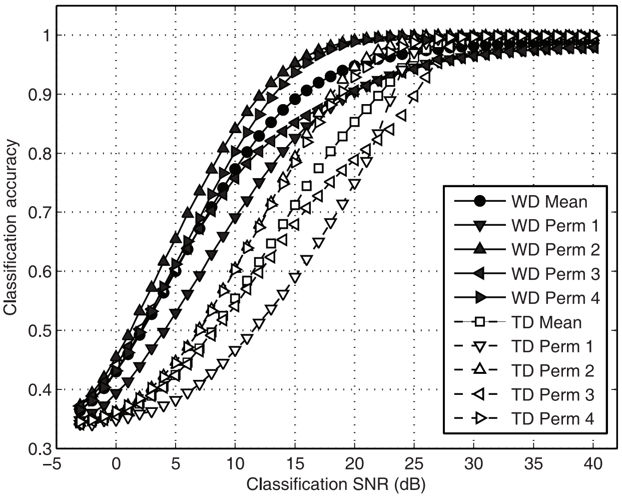


Fig. 8. Average MDA/ML classification results for all Cisco device permutations in Table 1. Mean across all permutations shown with filled markers.

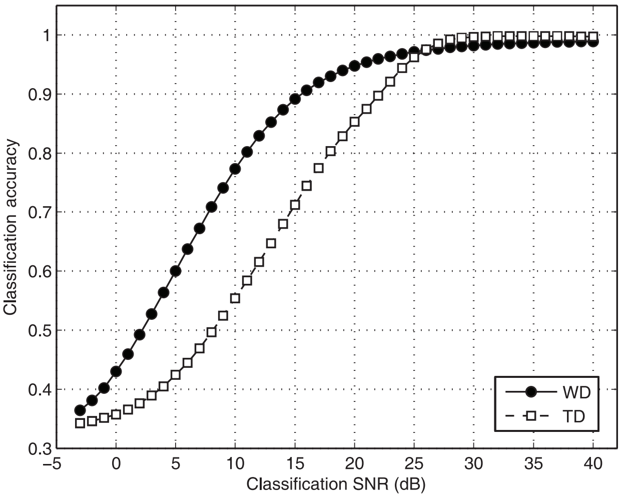


Fig. 9. Mean classification results from Fig. 8. Based on $\mathrm{CI} = 95\%$ confidence intervals, WD performance is superior for $-2 < \mathrm{SNR} < 24$ dB.

classification accuracy at SNR ≈ 11 dB. This represents a gain of approximately 7 dB with respect to equivalent TD fingerprinting performance.

Classification confusion matrices are presented in Table 3 for Permutation #1 of the Cisco devices for signals at $SNR = 11$ dB. Note that all values shown are statistically different based on $\mathrm{CI} = 95\%$ confidence intervals. As indicated in the lower comparison matrix, WD fingerprinting provides improved classification performance across all three classes, with the greatest improvement of 28.1% obtained in correctly classifying Class B. One common result with both fingerprinting techniques is that Class A and Class C devices are more confused with each other and confused less often with Class B. With respect to the device serial numbers, Class A and Class C are closer to each other than either one is to Class B.

Table 3. Confusion matrices for TD and WD fingerprinting: Permutation #1 from Table 1 at $SNR = 11$ dB. The WD – TD difference matrix is provided for comparison.

TD	Class Estimate		
Input Class	A	B	C
A	**49.4%**	17.3%	33.3%
B	18.5%	**65.9%**	15.6%
C	34.2%	12.1%	**53.6%**

WD	Class Estimate		
Input Class	A	B	C
A	**69.5%**	5.9%	24.5%
B	5.3%	**94.0%**	0.7%
C	21.5%	1.3%	**77.2%**

WD – TD	Class Estimate		
Input Class	A	B	C
A	**20.1%**	-11.4%	-8.8%
B	-13.2%	**28.1%**	-14.9%
C	-12.7%	-10.8%	**23.6%**

B. Relevant WD Features

Based on the number of classification features, the WD fingerprints represent an approximate 5-fold increase in dimensionality over TD fingerprints. Thus, it is reasonable to ask "Is the superior WD performance in Section IV-A attributable to increased feature dimensionality, more exploitable features, or both?" To help address this question, results were generated using a subset of 27 selected WD features from the larger 135-feature WD fingerprints. The idea was to compare TD and WD performance using an equivalent number of features.

The subset of WD features was selected using a generalized relevance learning vector quantization improved (GRLVQI) classifier [39], [45], [46]. The GRLVQI classifier jointly selects features and classifies in order to optimize features for classification. During this process, the algorithm calculates and outputs a relevance rating for each feature considered, indicating feature importance. The GRLVQI classifier was implemented in the Waikato environment for knowledge analysis (WEKA) environment [47] using WD fingerprints from bursts at $SNR = 40$ dB. The subset of 27 most relevant WD features is provided in Table 4 which shows the relevance ranking (RR), corresponding preamble subregion, WD level (WD LVL), signal characteristic and statistic.

The subset of 27 WD features in Table 4 were used for WD fingerprinting and performance compared with 27-feature TD fingerprinting under the most stressing device Permutation #1. Based on $\mathrm{CI} = 95\%$ confidence intervals, results in Fig. 10 show that 27-feature WD fingerprinting is superior for $0 < \mathrm{SNR} < 20$ dB and achieves 80% classification accuracy at SNR ≈ 19 dB. This represents a performance gain of approximately 2 dB relative to 27-feature TD fingerprinting. Given equal dimensionality, these results suggest a clear increase in exploitable DT-$\mathbb{C}$WT feature information.

Table 4. Subset of 27 most relevant WD features. Relevance ranking (RR) based on GRLVQI classifier output.

RR	Subregion	WD Lvl	Signal Characteristic	Statistic
1	Entire preamble	4	Amplitude	Kurtosis
2	Short symbols	4	Amplitude	Variance
3	Short symbols	5	Frequency	Variance
4	Entire preamble	4	Amplitude	Skewness
5	Short symbols	1	Amplitude	Kurtosis
6	Entire preamble	5	Frequency	Kurtosis
7	Short symbols	3	Amplitude	Kurtosis
8	Long symbols	2	Phase	Kurtosis
9	Entire preamble	3	Phase	Kurtosis
10	Entire preamble	3	Phase	Variance
11	Entire preamble	1	Frequency	Variance
12	Short symbols	3	Amplitude	Variance
13	Long symbols	2	Phase	Skewness
14	Entire preamble	5	Amplitude	Kurtosis
15	Entire preamble	4	Amplitude	Variance
16	Entire preamble	3	Amplitude	Kurtosis
17	Entire preamble	4	Frequency	Kurtosis
18	Short symbols	1	Frequency	Variance
19	Long symbols	1	Amplitude	Kurtosis
20	Entire preamble	5	Phase	Variance
21	Long symbols	5	Amplitude	Variance
22	Short symbols	2	Amplitude	Variance
23	Short symbols	4	Frequency	Kurtosis
24	Entire preamble	1	Phase	Variance
25	Entire preamble	3	Phase	Variance
26	Long symbols	1	Phase	Variance
27	Entire preamble	1	Phase	Kurtosis

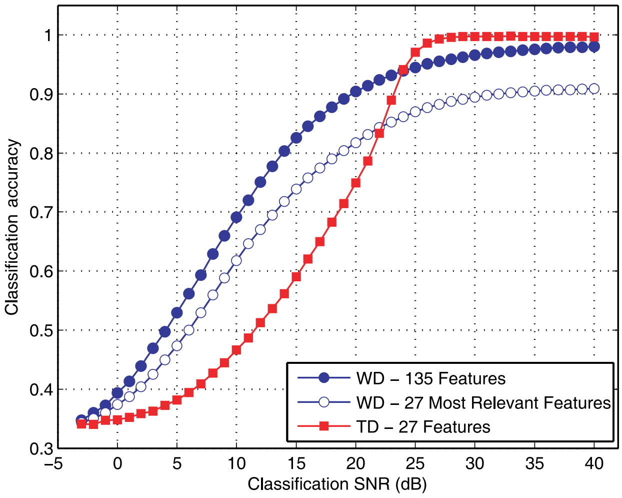


Fig. 10. MDA/ML classification for 27-feature TD and 27-feature WD fingerprinting using most stressing device Permutation #1. Based on $\mathrm{CI} = 95\%$ confidence intervals, WD performance is superior for $0 < \mathrm{SNR} < 20$ dB.

C. Burst Detection Error

The effect of burst detection error is demonstrated for TD and WD fingerprinting using random burst location error. This variation addresses operational situations where 1) dissimilar equipment is used for collecting training and classification data, 2) equipment is not necessarily co-located, or 3) equipment is operating in non-ideal environments. This type of error may also be induced by laboratory equipment, the fidelity of which can impact collected signal coloration and subsequent burst location accuracy. The observed random jitter error was applied to the perfect burst location data per Section III-C.2.

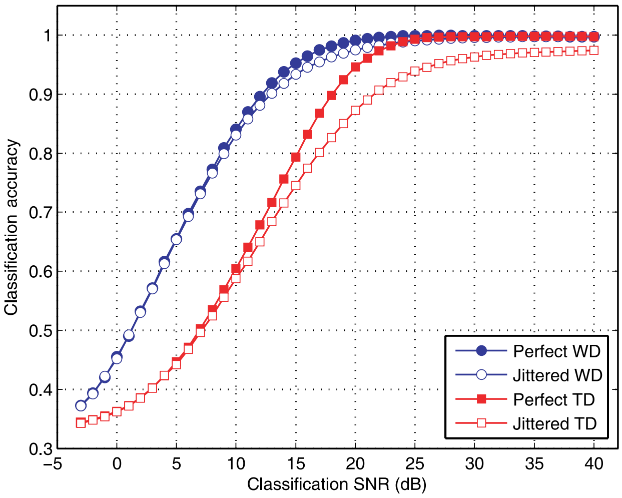


Fig. 11. Average MDA/ML classification accuracy for serial number discrimination using *observed* burst detection error statistics. Based on $\mathrm{CI} = 95\%$ confidence intervals, WD is superior and less sensitive to detection error.

Serial number discrimination results for WD and TD fingerprinting are provided in Fig. 11 for device Permutation #2 using observed jitter statistics for detection error. Performance for perfect detection is provided for comparison and shows that WD fingerprinting is clearly more robust. Based on $\mathrm{CI} = 95\%$ confidence intervals, WD fingerprinting superiority is indicated in three factors: 1) Jittered WD classification performance being better than jittered TD performance for all $-3 < \mathrm{SNR} < 40$ dB, 2) jittered gain at 80% classification accuracy being approximately 2 dB better than non-jittered gain, and 3) the average degradation in classification performance (perfect minus jittered) for a given technique across all SNRs is approximately 0.71% and 3.12% for WD and TD fingerprinting, respectively.

D. Dissimilar MDA/ML Training and Classification SNR

A fingerprinting comparison is made using dissimilar SNRs for MDA/ML training and classification. Specifically, the MDA training bursts were fixed at $\mathrm{SNR} = 40$ dB while the SNR of ML classification bursts varied over $\mathrm{SNR} \leq 40$ dB. These conditions are representative of an air monitor being pre-trained in a high SNR environment to recognize specific authorized devices and then operating in an actual environment with varying channel conditions, device locations, etc. The results presented are for Permutation #1 in Table 1 and are representative of what was obtained for other permutations.

Results in Fig. 12 are for serial number discrimination using both WD and TD fingerprinting with dissimilar SNRs for MDA/ML training and classification. Relative to performance using identical training and classification SNRs (filled mark-

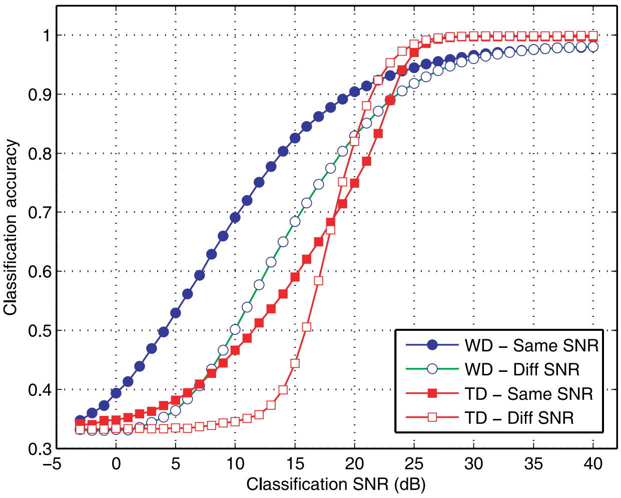


Fig. 12. Average MDA/ML classification accuracy for serial number discrimination using $\mathrm{SNR} = 40$ dB for MDA/ML training with subsequent classification performed at the SNRs indicated.

ers), the WD technique experiences a decrease in accuracy for all $\mathrm{SNR} < 30$ dB while the TD technique actually performs better at $\mathrm{SNR} > 18$ dB and exhibits decreased performance at $\mathrm{SNR} < 19$ dB. Based on $\mathrm{CI} = 95\%$ confidence intervals, WD performance is superior for all $\mathrm{SNR} < 20$ dB and achieves 80% classification accuracy at $\mathrm{SNR} \approx 19$ dB. This represents a modest gain of 1 dB with respect to equivalent TD fingerprinting performance and is approximately 7 dB less when compared with performance obtained when using identical SNRs for training and classification (filled markers).

V. CONCLUSION

The near-term challenge for advancing RF fingerprinting in wireless network security rests in finding robust fingerprint features. This is addressed here using a dual-tree complex wavelet transform (DT-$\mathbb{C}$WT) and wavelet-based fingerprints. Wavelet domain (WD) fingerprinting effectively exploits the nearly shift-invariant property of the DT-$\mathbb{C}$WT and provides improved classification relative to previous time domain (TD) approaches. Given that reliable serial number discrimination is achieved for 802.11a signals under both perfect and observed burst detection error conditions, RF fingerprinting remains a viable alternative for less constrained air monitoring applications.

TD and WD fingerprinting were compared using Fisher-based MDA/ML device classification under identical scenarios (device combinations, SNR, etc.). Sensitivity to varying channel SNR, dissimilar feature dimensionality, burst detection error and dissimilar SNRs for MDA/ML training and classification was considered as well. For all scenarios and sensitivities analyzed, WD fingerprinting emerged as the superior alternative for obtaining robust serial number discrimination at SNRs below 20 dB, with performance gains of up to 8 dB demonstrated at 80% classification accuracy.

"The views expressed in this article are those of the author(s) and do not reflect official policy of the United States Air Force, Department of Defense or the U.S. Government."

REFERENCES

[1] S. Bratus, C. Cornelius, D. Kotz, and D. Peebles, "Active behavioral fingerprinting of wireless devices," Tech. Report TR2008-610, Institute for Security Technology Studies, Dartmouth College, Mar. 2006.

[2] J. Franklin, D. McCoy, P. Tabriz, V. Neagoe, J. Randwyk, and D. Sicker, "Passive data link layer 802.11 wireless device driver fingerprinting," in *Proc. USENIX Annual Technical Conference*, June 2006, pp. 1–12.

[3] T. Kohno, A. Broido, and K. C. Claffy, "Remote physical device fingerprinting," *IEEE Trans. Dependable and Secure Commun.,* vol. 2, no. 2, pp. 93–108, 2005.

[4] O. Ureten and N. Serinken, "Wireless security through RF fingerprinting," *Canadian J. Elect. Comp. Eng.,* vol. 32, no. 1, pp. 27–33, Winter 2007.

[5] J. Hall, M. Barbeau, and E. Kranakis, "Using transceiverprints for anomaly based intrusion detection," in *Proc. 3rd IASTED Int. Conf. Comm, Internet and Info Technology (CIIT)*, Nov. 2004.

[6] J. Hall, M. Barbeau, and E. Kranakis, "Detection of transient in radio frequency fingerprinting using signal phase," in *Proc. IASTED Int. Conf. Wireless and Optical Comm (WOC)*, May 2003.

[7] O. Ureten and N. Serinken, "Detection of radio transmitter turn-on transients," *IEE Electron. Lett.*, vol. 35, no. 23, pp. 1996–1997, Nov. 1999.

[8] O. Ureten and N. Serinken, "Bayesian detection of WiFi transmitter RF fingerprints," *IEE Electron. Lett.*, vol. 41, no. 6, pp. 373–374, Mar. 2005.

[9] N. Serinken and O. Ureten, "Generalised dimension characterization of radio transmitter turn-on transients," *IEE Electron. Lett.*, vol. 36, no. 12, pp. 1064–1064, June 2000.

[10] K. J. Ellis and N. Serinken, "Characteristics of radio transmitter fingerprints," *Radio Science*, vol. 36, no. 4, pp. 585–597, 2001.

[11] Y. Chen, W. Trappe, and R. Martin, "Detecting and localizing wireless spoofing attacks," in *Proc. IEEE Conf. Sensor, Mesh and Ad Hoc Comm and Nets (SECON)*, June 2007, pp. 193–202.

[12] Y. Sheng, K. Tan, G. Chen, D. Kotz, and A. Campbell, "Detecting 802.11 MAC layer spoofing using received signal strength," in *Proc. IEEE INFOCOM*, Apr. 2008, pp. 1768–1776.

[13] B. Danev and S. Kapkun, "Implications of radio fingerprinting on the security of sensor networks," in *Proc. 3rd Int. Conf. Security and Privacy in Communication Networks (ICST)*, Sept. 2007, pp. 1–5.

[14] W. C. Suski, M. A. Temple, M. J. Mendenhall, and R. F. Mills," Using spectral fingerprints to improve wireless network security," in *Proc. IEEE GLOBECOM*, Nov. 2008, pp. 1–5.

[15] W. C. Suski, M.A. Temple, M. J. Mendenhall, and R. F. Mills, "Radio frequency fingerprinting commercial communication devices to enhance electronic security," *Int. J. Electronic Security and Digital Forensics*, vol. 1, no. 3, pp. 301–322, 2008.

[16] B. Danev and S. Kapkun, "Transient-based identification of wireless sensor nodes," in *Proc. the ACM/IEEE Int. Conf. Information Processing in Sensor Networks (IPSN)*, Apr. 2009, pp. 25–36.

[17] R. W. Klein, M. A. Temple, M. J. Mendenhall, and D. R. Reising, "Sensitivity analysis of burst detection and RF fingerprinting classification performance," in *Proc. IEEE ICC*, June 2009, pp. 1–5.

[18] R. W. Klein, M. A. Temple, and M. J. Mendenhall, "Application of wavelet denoising to improve OFDM-based signal detection and classification," *J. Security and Communication Networks, Special Issue: Security in Next Generation Wireless Networks*, vol. 2, no. 6, 2009.

[19] O. Ureten and N. Serinken, "Improved coarse timing for burst mode OFDM," in *Proc. IEEE GLOBECOM*, Nov. 2007, pp. 2841–2846. [

[20] O. Ureten, R. A. Pacheco, N. Serinken, and D. Hatzinakos, "Bayesian frame synchronization for 802.11a WLANs: Experimental results," in *Proc. Canadian Conf. Electrical and Computer Engineering (CCECE)*, May 2005, pp. 884–887.

[21] R. A. Pacheco, O. Ureten, D. Hatzinakos, and N. Serinken, "Bayesian frame synchronization using periodic preamble for OFDM-based WLANs," *IEEE Signal Process. Lett.*, vol. 12, no. 7, pp. 524–527, July 2005.

[22] S. Haykin, "Cognitive radio: Brain-empowered wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 2, pp. 201–220, Feb. 2005.

[23] W. C. Y. Lee, "CS-OFDMA: A new wireless CDD physical layer scheme," *IEEE Commun. Mag.*, vol. 43, no. 2, pp. 74–79, Feb. 2005.

[24] P. Zhang, X. Tao, J. Zhang, Y. Wang, and Y. Wang, "A vision from the future: Beyond 3G TDD," *IEEE Commun. Mag.*, vol. 43, no. 1, pp. 38–44, Jan 2005.

[25] D. Shaw and W. Kinsner, "Multifractal modelling of radio transmitter transients for classification," in *Proc. Conf. IEEE WESCANEX 97: Communications, Power and Computing*, May 1997, pp. 306–312.

[26] O. Ureten and N. Serinken, *Detection, characterisation and classification of radio transmitter turn-on transients*, Kluwer Academic Publishers: Netherlands, 2000.

[27] O. H. Tekbas and N. Serinken, "Transmitter fingerprinting from turn-on transients," in *Proc. NATO SET Panel Symp. Passive and LPI Radio Frequency Sensors*, Warsaw, Poland, Apr. 2001.

[28] O. H. Tekbas, O. Ureten, and N. Serinken, "Improvement of transmitter identification system for low SNR transients," *IEE Electron. Lett.,* vol. 40, no. 3, pp. 182–183, Feb. 2004.

[29] O. H. Tekbas, N. Serinken, and O. Ureten, "An experimental performance evaluation of a novel radio-transmitter identification system under diverse environmental conditions," *Canadian J. Electrical and Computer Engineering*, vol. 29, no. 3, pp. 203–209, July 2004.

[30] J. Toonstra and W. Kinsner, "Transient analysis and genetic algorithms for classification," in *Proc. IEEE WESCANEX: Communications, Power, and Computing*, May 1995, pp. 432–437.

[31] A. Prochazka and M. Storek, "Wavelet transform use for signal classification by self-organizing neural networks," in *Proc. Int. Conf.Artificial Neural Networks*, June 1995, pp. 295–299.

[32] J. Hall, "Detection of rogue devices in wireless networks," Ph.D. Thesis, School of Computer Science, Carleton University, Aug. 2006.

[33] H. Hotelling, "The generalization of students ratio," *Annals of Mathematical Statistics*, 1931.

[34] I. Bayram and I. W. Selesnick, "On the dual-tree complex wavelet packet and M-band transforms," *IEEE Trans. Signal Process.*, vol. 56, no. 6, pp. 2298–2310, June 2008.

[35] R. A. Fisher, "The use of multiple measurements in taxonomic problems," *Annals of Eugenics*, vol. 7, pp. 179–188, 1936.

[36] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification,* 2nd ed., John Wiley & Sons, Inc.: New York, 2001.

[37] J. A. Pitta, R. D. Hippenstiel, and M. P. Fargues, "Transient detection using wavelets," Masters Thesis, Naval Post Graduate School, Mar. 1995.

[38] I. W. Selesnick, R. G. Baraniuk, and N. C. Kingsbury, "The dual-tree complex wavelet transform," *IEEE Signal Process. Mag.*, vol. 22, no. 6, pp. 123–151, Nov. 2005.

[39] M. J. Mendenhall and E. Merenyi, "Relevance-based feature extraction from hyperspectral images in the complex wavelet domain," in *Poc. IEEE Mountain Workshop on Adaptive and Learning Systems*, July 2006, pp. 24–29.

[40] I. W. Selesnick, Wavelet Software at Polytechnic University, Brooklyn, NY, [Online]. Available. http://taco.poly.edu/WaveletSoftware/

[41] Agilent Technologies Inc., USA. Agilent E3238 Signal Intercept and Collection Solutions: Family Overview, Pub 5989-1274EN, July 2004.

[42] IEEE Computer Society. IEEE Std 802.11-2007, June 2007.

[43] T. Hastie, R. Tibshirani, and J. Friedman, *Data Mining, Inference, and Prediction*, Springer: NY, 2001.

[44] L. M. Leemis and S. K. Park, *Discrete-Event Simulation: A First Course*, Prentice Hall: New Jersey, 2006.

[45] M. J. Mendenhall and E. Merenyi, "Relevance-based feature extraction for hyperspectral images," *IEEE Trans. Neural Netw.*, vol. 19, no. 4, pp. 658–672, Apr. 2008.

[46] M.J. Mendenhall and E. Merenyi, "Generalized relevance learning vector quantization for classification-driven feature extraction from hyperspectral data," in *Proc. American Society for Photogrammetry and Remote Sensing,* May 2006.

[47] Ian H. Witten and Eibe Frank. *Data Mining: Practical Machine Learning Tools and Techniques*, Morgan Kaufmann, 2005.

Randall W. Klein is a Major in the US Air Force. He received his B.S.E.E. degree from the US Air Force Academy in 1999, his M.S.E.E. from the Air Force Institute of Technology (AFIT) in 2001, and his Ph.D. in electrical engineering from AFIT in 2009. His research interests are in RF burst detection and RF fingerprinting techniques to provide hardware specific device discrimination. Randall is a Member of the Tau Beta Pi and Eta Kappa Nu honor societies.

Michael A. Temple is a Professor of Electrical Engineering at the US Air Force Institute of Technology (AFIT). He received his B.S.E. (1985) and M.S.E. (1986) degrees from Southern Illinois University, Edwardsville, and his Ph.D. from AFIT in 1993. He is a Member of AFITs Center for Cyberspace Research (a USAF designated Cyberspace Technical Center of Excellence) and is the principal investigator for physical layer cyber operations research. His research interests are in Spectrally Modulated, Spectrally Encoded (SMSE) waveform design and RF DNA fingerprinting. He is a Member of Eta Kappa Nu, Tau Beta Pi, and a Senior Member of IEEE.

Michael J. Mendenhall is an Assistant Professor of Electrical Engineering at the US Air Force Institute of Technology. He received his B.S. degree in computer engineering from Oregon State University, Corvallis, in 1996, his M.S. degree in Computer Engineering from AFIT in 2001, and his Ph.D. in Electrical Engineering from Rice University, Houston, TX, in 2006. His research interests are in exploitation and processing of hyperspectral images, hyperspectral signature modeling, and the broader topic of computational intelligence.