2010

Application of Wavelet Denoising to Improve OFDM-based Signal Detection and Classification

Randall W. Klein

Michael A. Temple
Air Force Institute of Technology

Michael J. Mendenhall
Air Force Institute of Technology

Recommended Citation

Klein, R. W., Temple, M. A., & Mendenhall, M. J. (2010). Application of wavelet denoising to improve OFDM-based signal detection and classification. Security and Communication Networks, 3(1), 71–82. https://doi.org/10.1002/sec.115

SECURITY AND COMMUNICATION NETWORKS
Security Comm. Networks. 2010; **3**:71–82
Published online 24 June 2009 in Wiley InterScience
(www.interscience.wiley.com) DOI: 10.1002/sec.115

Application of wavelet denoising to improve OFDM-based signal detection and classification[‡,§]

Randall W. Klein, Michael A. Temple[*,†] and Michael J. Mendenhall
Air Force Institute of Technology, Wright-Patterson AFB, OH 45433, U.S.A.

Summary

The developmental emphasis on improving wireless access security through various OSI PHY layer mechanisms continues. This work investigates the exploitation of RF waveform features that are inherently unique to specific devices and that may be used for reliable device classification (manufacturer, model, or serial number). Emission classification is addressed here through detection, location, extraction, and exploitation of RF 'fingerprints' to provide device-specific identification. The most critical step in this process is burst detection which occurs prior to fingerprint extraction and classification. Previous variance trajectory (VT) work provided sensitivity analysis for burst detection capability and highlighted the need for more robust processing at lower signal-to-noise ratio (SNR). The work presented here introduces a dual-tree complex wavelet transform (DT-$\mathbb{C}$WT) denoising process to augment and improve VT detection capability. The new method's performance is evaluated using the instantaneous amplitude responses of experimentally collected 802.11a OFDM signals at various SNRs. The impact of detection error on signal classification performance is then illustrated using extracted RF fingerprints and multiple discriminant analysis (MDA) with maximum likelihood (ML) classification. Relative to previous approaches, the DT-$\mathbb{C}$WT augmented process emerges as a better alternative at lower SNR and yields performance that is 34% closer (on average) to 'perfect' burst location estimation performance. Published in 2009 by John Wiley & Sons, Ltd.

KEY WORDS: RF fingerprints; wavelet denoising; OFDM; 802.11a; dual tree wavelet transform

1. Introduction

Considerable research has been conducted on detecting and/or mitigating spoofing within the medium access control (MAC) layer of the open systems interconnection (OSI) stack [1,2]. There has been a recent shift toward providing added security at the OSI Physical (PHY) layer by exploiting RF features that are inherently unique to a specific device and that are difficult to replicate by an unintended party. For example, some efforts have investigated received signal strength (RSS) (a power-based metric) for detecting and/or locating a spoofing node [1,2]. Both of these efforts demonstrated some success at detecting spoofing using experiments conducted with different hardware and in different physical environments.

[*]Correspondence to: Michael A. Temple, AFIT/ENG Air Force Institute of Technology, Wright-Patterson AFB, OH 45433, U.S.A.
[†]E-mail: michael.temple@afit.edu
[‡]The views expressed in this paper 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.

RF fingerprinting is one alternative PHY layer approach that is readily dismissed in Reference [2] for 'scale' reasons. For applications where size constraints are less restrictive, RF fingerprinting remains a viable alternative and is considered here as in previous works [3,4]. Collectively, related works in RF fingerprinting, electromagnetic signatures, intra-pulse modulation, and unintentional modulation [5–13] form a solid basis for developing techniques that may be applicable to commercial communication devices. If the inherent RF fingerprints are repeatedly extractable and sufficiently unique, they are potentially useful for determining the specific make, model, and/or serial number of a given device.

Previous work highlighted signal structure uniqueness and attributed inter-device differences to various manufacturing, aging, and environmental factors [5]. While several processing steps are required to effectively exploit the unique RF fingerprints, burst location is arguably the most important [8,10]. In this context, burst location includes determining both the burst start time and the subsequent signal region(s) from which fingerprints are extracted. Both of these factors are important given that improper selection of either can unduly bias the processing to favor channel noise effects or steady-state signal effects [5]. With the exception of more recent work in References [3,4,14], these previous efforts lack a detailed sensitivity analysis of burst detection and fingerprint classification performance under varying channel noise conditions.

Noise sensitivity analysis is imperative for determining the minimum acceptable signal-to-noise ratio (SNR) that provides consistent and reliable classification results. This minimum acceptable SNR also allows determination of maximum transmitter–receiver separation distances which can aide in establishing the geometric layout of physical hardware to improve overall network security. Noise sensitivity performance also provides a good discriminant for comparing various detection and classification techniques. For the work presented here, burst location performance is conducted for a combined channel noise and burst-to-burst variability effect using multiple 802.11a bursts and multiple independent noise realizations for each burst.

Related burst detection work in Reference [14] provides preliminary results using variance trajectory (VT) of 802.11a instantaneous amplitude response. The choice of using OFDM-based signals for demonstration, and in particular the 802.11a signal, was driven by two factors, including (1) consistency with previous related 802.11a work that has been extensively published [3,4,14–17], and (2) the continued emergence of OFDM-based signals as envisioned for 3G/4G (IMT/IMT-Advanced) radio communications. While the transient detection and classification techniques used in this work 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.

The impact of burst detection error on signal classification performance is addressed here using RF fingerprints and multiple discriminant analysis (MDA) with maximum likelihood (ML) classification. While VT burst detection and MDA-ML classification performance in earlier work [14] was shown to be consistent with 'perfect' burst estimation performance at higher SNRs in the range of $10 \leq \text{SNR} \leq 30\,\text{dB}$, performance diverged at SNRs in the range of $-3 \leq \text{SNR} \leq 10\,\text{dB}$. These previous results demonstrated a margin for improvement only at the lower SNRs and highlighted the need for a more robust technique, providing the impetus for the work presented in this paper. More specifically, a dual-tree complex wavelet transform (DT-$\mathbb{C}$WT) is introduced to denoise the signal prior to VT calculation to improve performance at lower SNRs ($-3 \leq \text{SNR} \leq 10\,\text{dB}$). It is envisioned that this technique would be activated only at the operator's discretion, since there is no gain at higher SNRs.

2. Background

2.1. Signal Characteristics

Device emissions can be exploited using various signal characteristics. However, instantaneous amplitude and instantaneous phase characteristics are perhaps the most extensively investigated [5,8–10]. More recently, these two characteristics have been augmented with instantaneous frequency and successfully exploited for device classification [3,4,14]. The instantaneous amplitude, $a(k)$, instantaneous phase, $\theta(k)$, and instantaneous frequency, $f(k)$, responses of a complex sampled signal $s(k)$ are given by [18]

$$s(k) = i(k) + jq(k) \tag{1}$$

$$a(k) = \sqrt{i^2(k) + q^2(k)} \tag{2}$$

$$\theta(k) = \tan^{-1}\left[\frac{q(k)}{i(k)}\right] \tag{3}$$

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

where $i(k)$ and $q(k)$ are the instantaneous in-phase and quadrature-phase components of $s(k)$.

2.2. Variance Trajectory (VT)

The work in Reference [8] analyzed Bluetooth signals using the VT process with instantaneous phase. The process windows an input signal and calculates the signal variance for each window, creating a variance vector. The difference between consecutive variance values is calculated to form the VT. The work presented here generates VT sequence $\{\mathrm{VT}_a(i)\}$ using instantaneous amplitude sequence $\{a(k)\}, k = 1, 2, \ldots, N_a$, to estimate the burst start. The ith element of sequence $\{\mathrm{VT}_a(i)\}$ is given by [3]

$$\mathrm{VT}_a(i) = |W_a(i) - W_a(i+1)| \qquad i = 1, 2, \ldots, L_w - 1 \tag{5}$$

$$W_a(m) = \frac{1}{N_w}\sum_{k=1+(m-1)N_s}^{1+(m-1)N_s+N_w} [a(k) - \mu_w]^2 \qquad m = 1, 2, \ldots, L_w \tag{6}$$

where N_w is the window extent, and N_s is the number of samples the window advances between calculations. The μ_w factor in Equation (6) is the sample mean of $\{a_w(k)\}$ which is the subsequence of consecutive elements from $\{a(k)\}$ contained in the window. Figure 1 shows a representative amplitude response and corresponding VT response for two different analysis SNRs. As shown, there is a distinct VT peak response corresponding to the burst start which becomes less discernable as SNR decreases.

2.3. Discrete Wavelet Transform Denoising

Signal denoising is accomplished using a discrete Wavelet transform (DWT) by exploiting differences in the distribution of signal burst energy and the additive white gaussian noise (AWGN) in which it is embedded. The noise channel is uniformly distributed in the wavelet domain. Signal bursts, however, are

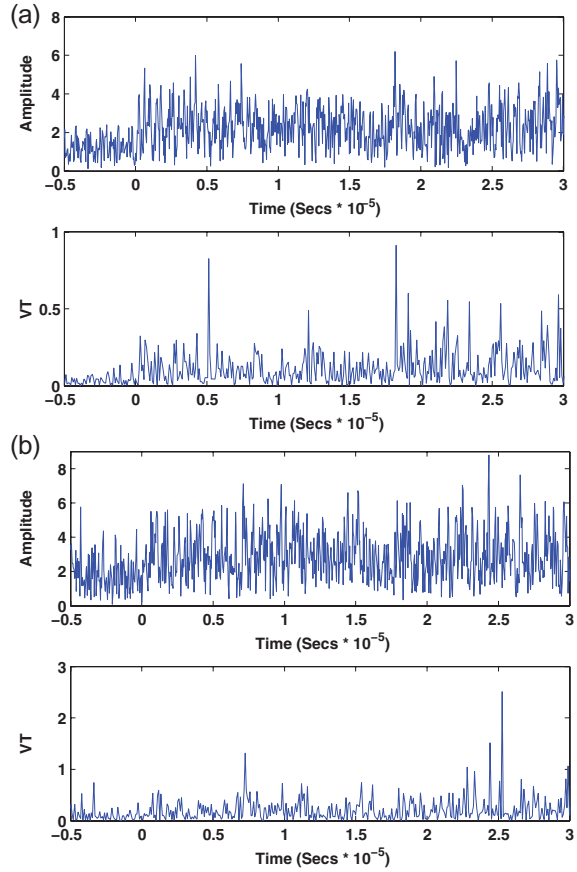

Fig. 1. Instantaneous signal amplitude and corresponding VT response for (a) SNR = 6 dB and (b) SNR = 0 dB.

non-uniformly distributed in the wavelet domain and significant signal information manifests in the large magnitude wavelet coefficients. It is common to threshold the magnitude of the wavelet coefficients in wavelet denoising applications where coefficients larger than the threshold contain significant signal contribution [19–28]. Those wavelet coefficients with magnitude less than the threshold are understood to be noise. Due to the compaction property of the wavelet transform, there are relatively few large magnitude coefficients.

One distinct disadvantage of the DWT is the lack of shift invariance. i.e., if the signal is shifted in time by some amount, the transformation of that signal yields a different set of coefficients. In the application addressed here, this problem has the consequence of complicating the computation of reasonable thresholds for signal denoising. This problem is mitigated by using the DT-ℂWT [29,30].

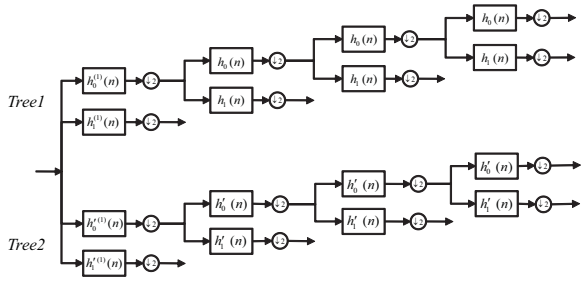

Fig. 2. Representative dual-tree complex wavelet transform (DT-ℂWT) structure [30].

2.4. Dual-tree Complex Wavelet Transform

The dual-tree complex wavelet transform (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 inter-scale and intra-scale neighborhoods [29]. Furthermore, the DT-ℂWT magnitude response exhibits reduced ringing in the wavelet domain due to high-frequency noise and sharp discontinuities, which make the denoising algorithms more reliable. The DT-ℂWT is commonly implemented as two real-valued filter banks as shown in Figure 2.

For real-valued input signals, the Tree1 and Tree2 filter banks yield real-valued coefficients representing real and imaginary components of complex coefficients, respectively [29]. For complex input signals as used in this work, each filter bank yields complex coefficients which are annotated as Tree1 and Tree2 for discussion. 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 [29,30]

$$\psi(t) = \sqrt{2}\sum_n h_1(n)\phi(2t-n) \tag{7}$$

where

$$\phi(t) = \sqrt{2}\sum_n h_0(n)\phi(2t-n) \tag{8}$$

Ideally, the corresponding functions for the Tree2 path are the Hilbert transforms of Equations (7) and (8) and given by

$$\psi'(t) = \sqrt{2}\sum_n h'_1(n)\phi'(2t-n) \tag{9}$$

where

$$\phi'(t) = \sqrt{2}\sum_n h'_0(n)\phi'(2t-n) \tag{10}$$

3. Methodology

3.1. DT-ℂWT Denoising

The process for denoising with the DT-ℂWT is illustrated in Figure 3. The complex input signal $f(n)$ is transformed *via* the DT-ℂWT (Section 2.4) producing two complex-valued sets of wavelet coefficients, Tree1 and Tree2, one from each filter bank. The two complex sets are combined into one set of real-valued coefficients $d(n)$ according to

$$d(n) = \sqrt{|\text{Tree1}(n)|^2 + |\text{Tree2}(n)|^2} \tag{11}$$

The resultant $d(n)$ coefficients from Equation (11) are compared with threshold t_{denoise} and all coefficients less than t_{denoise} are set to zero. That is, $\forall n'$ where $d(n') < t_{\text{denoise}}$, $\text{Tree1}(n') = 0$ and $\text{Tree2}(n') = 0$ and the 'punctured' set of coefficients $d'(n)$ produced. The inverse DT-ℂWT (IDT-ℂWT) is then applied to $d'(n)$ to create the denoised complex signal $g(n)$. The denoised coefficients are subsequently processed using the VT technique described in Section 2.2 to generate Denoised VT results. Figure 4 shows representative amplitude and corresponding VT responses for two different analysis SNRs of a denoised signal. As compared to the Traditional VT signal responses in Figure 1, there is a more distinct peak associated with the burst start at Time = 0 s.

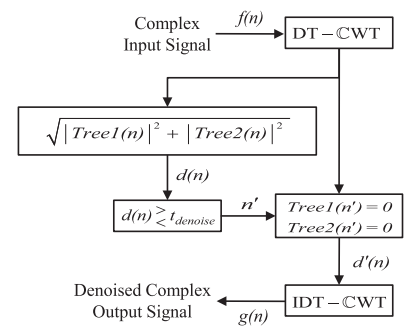

Fig. 3. Illustration of DT-ℂWT denoising process.

Published in 2009 by John Wiley & Sons, Ltd.

DOI: 10.1002/sec

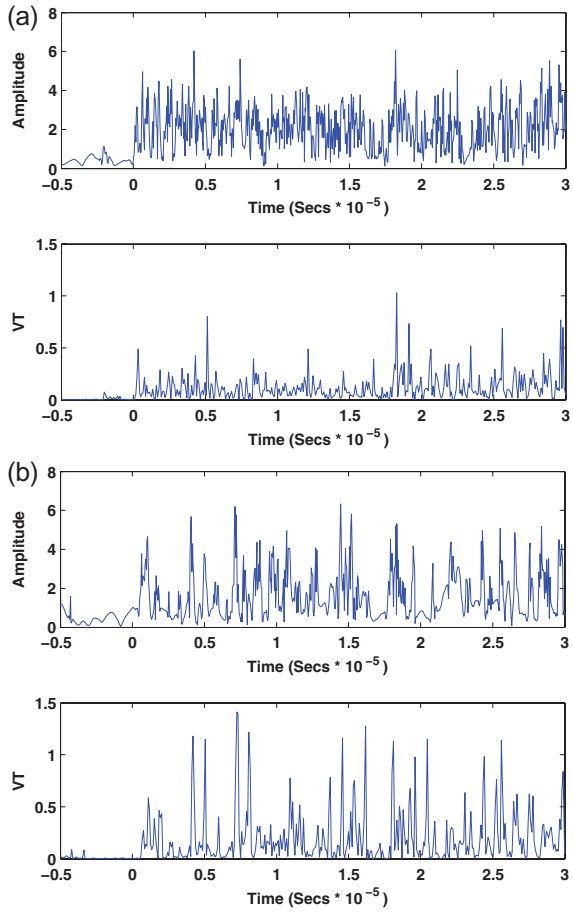

Fig. 4. Denoised instantaneous signal amplitude and corresponding VT response for (a) SNR = 6 dB and (b) SNR = 0 dB.

3.2. Overall Demonstration Process

The sequential burst detection, transient location, and classification process are implemented relative to what commonly occurs in an operational collection, i.e., a real-time system samples the received signal, detects the 'presence' of a burst, locates a given starting point (sample number) within the burst turn-on transient region, and classifies the burst using specific extracted features. If a burst is received and its 'presence' not declared, it is an undetected burst. If a burst is received and its 'presence' declared, it is a detected burst. The focus of this work is on detected bursts with subsequent algorithmic processing used for transient location. For those cases where the transient location algorithm does not converge in accordance with prescribed criteria (number of iterations, parametric tolerance, etc.), the detected bursts are designated as 'non-convergent' and a default transient location value assigned. When algorithm convergence occurs, the bursts are designated as 'convergent' and the estimated location assigned. When algorithm convergence occurs for identical bursts with both location techniques, the bursts are designated as 'dual convergent' bursts. This operational methodology is implemented here using a four step process comprised of: (1) burst detection, (2) burst start location, (3) RF fingerprint extraction, and (4) device classification.

(1) Step 1—burst detection: The first step represents coarse burst detection. This step monitors the RF environment to detect the *presence* of RF bursts. The received signal is segmented into a consecutive series of subsections using relatively wide, non-overlapping windows, $N_s = N_w$ from Equation (6). While not a requirement, non-overlapping windows are used to minimize processing time. This has the disadvantage of producing coarser estimates of where the actual burst response starts, while at the same time capturing more signal power within each window and improving detectability. For all results presented in this paper, a window size of $N_w = 512$ signal samples (21.6 μs) is used for Step 1.

The two burst detection methods (VT and Denoised VT as described previously in Sections 2.2 and 3.1, respectively) are applied to the windowed signal data and an *a priori* threshold t_{detect} used to declare detection. Once detection occurs, the corresponding segment of windowed signal data is passed to Step 2 where it is assumed that an actual burst start occurs within the window. However, as with all signal detection approaches false alarms can occur with bursts falsely declared present. Results of a performance comparison between the two detection methods for this step are provided in Section 4.1.

(2) Step 2—burst start location: This step is similar to Step 1 in that the two detection methods are reapplied. However, the objective here is to accurately and precisely locate the *time* at which an abrupt change occurs in the VT_a response of Equation (5). The effectiveness of this approach is based on the implicit assumption that the 802.11a OFDM signal can be modeled as having a step change response in the burst transient region. This assumed response is consistent with 802.11a specifications [31] and has been successfully exploited in change point estimation research [15–17].

The segment of windowed data passed from Step 1 is further sub-segmented using much narrower and highly overlapped windows. The overlapping windows allow for better location accuracy at the expense of increased processing time. For this work, a window size of $N_w = 20$ samples (0.84 μs) is used with a shift of $N_s = 2$ samples (84.2 ns) between consecutive windows. A different *a priori* detection threshold t_{locate} (compared to Step 1) is used to automatically estimate the burst start location based on a significant peak response occurring in $\text{VT}_a(i)$ of Equation (5).

If a significant peak is located, the signal is passed to Step 3. It is possible that no significant peak is found and the algorithm therefore does not converge to a solution. This could be the case if Step 1 passed along a false alarm and no signal burst is present or if the threshold is just too high for that particular burst. When this happens, there are two options available: the signal can be discarded or a default start location can be assigned. In an operational environment, where the algorithm has exposure to numerous bursts, discarding signals may be a good choice. However, for this work, the probability of detect in Step 1 is 100%, since the data is manually extracted from the RF environment. With that assumption and the limited amount of data to process, a default location is assigned to those signals which had no significant peak. As per Section 3.2, these bursts are denoted as 'non-convergent' bursts. The default location is chosen to be the last sample number of the window. Only those bursts that converge to a solution are used in remaining steps and are denoted as 'convergent' bursts. A performance comparison for this step between the two methods is documented in Section 4.2.

(3) Step 3—RF fingerprint extraction: After locating the burst start, statistical waveform feature data is extracted from the next 16.0 μs of the burst to represent the RF Fingerprint. This 16.0 μs region of the burst corresponds to the 802.11a preamble [31]. More specifically, statistics are calculated for the three instantaneous waveform characteristics given by Equations (2), (3), and (4) across three distinct regions within the 16.0 μs preamble, including the short symbol region, the long symbol region, and the combined short-long symbol region (entire preamble). The three statistical features considered for this work include the variance, skewness, and kurtosis.

As a result of this process, the RF fingerprint (feature vector) used for device classification consists of 27 features per burst (3 preamble regions × 3 waveform characteristics × 3 statistics). After RF fingerprints have been extracted from each detected burst, the feature vectors are passed to Step 4.

(4) Step 4—device classification: The impact of detection error on signal classification performance is illustrated using the extracted RF fingerprints from Step 3 and MDA with ML classification. MDA is an extension of Fisher's Linear Discriminant (FLD) process for more than two classes [32]. Classification is demonstrated here using an MDA-ML process [33]. For the 3-class problem, the MDA process projects higher-dimensional data onto a two-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 dimensionality greater than 3. Using this lower-dimensional data, ML decision boundaries are determined assuming normally distributed input data, equal costs or risk, and uniform prior probabilities. To discriminate c classes using d-dimensional input data, the input vector $\mathbf{x}$ is linearly projected onto a $(c-1)$-dimensional space using

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

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 two-dimensional trained decision boundaries. The process classifies each unknown input data set by projecting it onto the trained 'Fisher plane' using Equation (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 performance comparison for this step between the two methods is documented in Section 4.3.

3.3. Data Collection and Noise Simulation

The process for collecting 802.11a signals is shown in Figure 5. The Agilent-based RF Signal Intercept

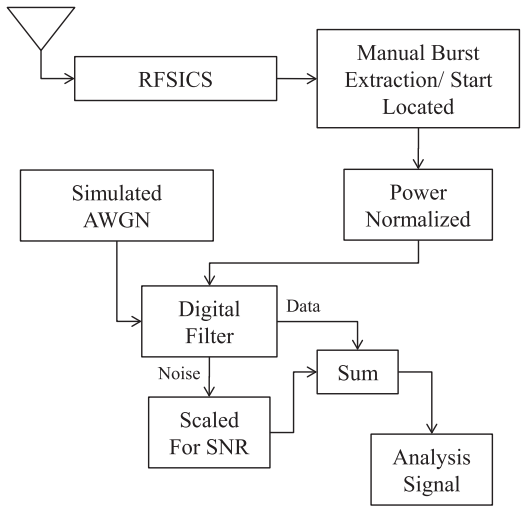

Fig. 5. Process for RFSICS 802.11a signal collection and post-collection processing, to include noise generation, digital filtering, and scaling for desired analysis SNR.

and Collection System (RFSICS) was used to collect approximately 200 bursts from three different 802.11a devices. For all collections, the devices under test and RFSICS were co-located in a typical wireless office environment, i.e., a room containing 20–25 active workstations with various wireless peripherals, personnel with wireless personal communication devices, a wireless network access node, composite cubicle partitions, metal book cases, metal filing cabinets, etc. As such, all collected signals inherently include the desired line-of-sight (LOS) signal component as well as in-band interfering components commonly found in such an environment, e.g., non-LOS multipath, intra-system multiple access, inter-system coexistence, etc.

Basic functionality of the RFSICS is provided by Agilent's E3238S system [34]. This includes an RF front-end collection range of 20.0 MHz to 6.0 GHz from which a 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 4, for a final sample rate of $f_s = 23.75$ MSPS and corresponding sample interval of $T_s = 1/f_s \approx 42.1$ ns per sample.

The near-baseband RFSICS data were further post-processed in a MATLAB environment with each burst visually analyzed to accurately identify the sample number corresponding to the burst start. Starting from this sample number, a sufficiently long portion of the signal is extracted to capture the 802.11a preamble response and a small portion of pseudo-randomly modulated data. The extracted burst response is then normalized to equal power with respect to the other collected bursts. Next, the bursts responses are digitally filtered and their power at the filter output is calculated. A 6th-order baseband Chebyshev digital filter was used with a -3 dB bandwidth of 9.2 MHz. At this point, the final baseband signals used for burst detection, transient location, and classification are sampled at frequency of $f_s = 23.75$ MSPS and are effectively oversampled by a factor of approximately 1.3 times Nyquist. The typical *collected* SNR at this point in the process is on the order of SNR $= 40$ dB. Provided that RFSICS collection and subsequent post-processing is identical for all signals, it is reasonable to assume that 'coloration' (variation in amplitude, phase and/or frequency characteristics) induced prior to burst detection, transient location, and burst classification is approximately identical as well. 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. To simulate varying SNR conditions, like-filtered noise is added prior to analysis. This is done by generating random complex AWGN that is filtered using the same digital filter as used for the signal. The filtered noise power is then calculated and used to scale the filtered AWGN to achieve the appropriate analysis SNR when added to the filtered signal. A representative 802.11a RF burst is shown in Figure 6 at the collected SNR and analysis SNRs of SNR $= 10$ dB and SNR $= 0$ dB.

3.4. Threshold Determination

Three distinct threshold values are required, one each for burst detection, burst location, and denoising when the Denoised VT process is employed. All SNR-dependent threshold values are determined *a priori* based on noise-only analysis using 100 000 AWGN realization. When evaluating the Denoised VT technique, the DT-$\mathbb{C}$WT denoising threshold t_{denoise} in

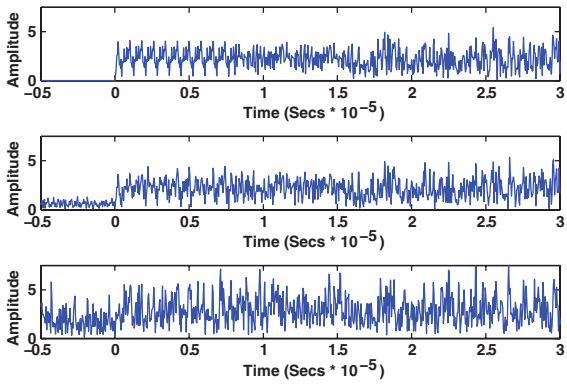

Fig. 6. Instantaneous amplitude: (Top) *collected* signal, (Middle) *filtered* signal-plus-AWGN at SNR = 10 dB, and (Bottom) *filtered* signal-plus-AWGN at SNR = 0 dB.

Section 3.1 is established using random realizations of noise that are generated, filtered, scaled (for appropriate SNR), DT-ℂWT transformed and coefficients retained for threshold determination. The remaining burst detection threshold (t_{detect}) and burst location threshold (t_{locate}) are determined using a similar noise-only analysis with appropriate window parameters for the given technique. In all cases, results from 100 000 iterations are histogrammed and the threshold value empirically chosen.

The DT-ℂWT denoising threshold value (t_{denoise}) is empirically chosen and corresponds to the histogram bin value below which 95% of the noise-only values occur. The burst detection threshold value (t_{detect}) is chosen using conventional noise-only analysis of probability of false alarm (P_{fa}) and probability of detection (P_{d}) as represented on a receiver operating characteristic (ROC) curve. Results of this analysis are reported in Section 4.1.

The burst start location threshold value (t_{locate}) is chosen such that there is a trade-off between the number of early burst location estimates (Time < 0 s) *versus* the number of algorithm non-convergent solutions. The actual values chosen were selected to ensure that both of these conditions are present and observable in the data. When comparing the Traditional VT and Denoised VT processes, the threshold is constrained to provide a similar number of early locates (10%) for both techniques to illicit a more fair comparison.

4. Results

Burst detection, burst start location, and device classification results are presented for both the Traditional VT and Denoised VT techniques for $-3 \leq \text{SNR} \leq 10$ dB.

4.1. Burst Detection

ROC curves were calculated as described in Step 2 of Section 3.2 to characterize performance differences between the two burst detection techniques. Results in Figure 7 show that at both SNR = 6 dB and SNR = 0 dB, the Denoised VT technique provides a higher probability of detection (P_{d}) for a given probability of false alarm (P_{fa}). With a higher P_{d} for a given P_{fa}, the Denoised VT technique passes more bursts to the next step when compared with Traditional VT. With more bursts passing, it is possible to correctly classify the device in less time and have a higher confidence in the classification.

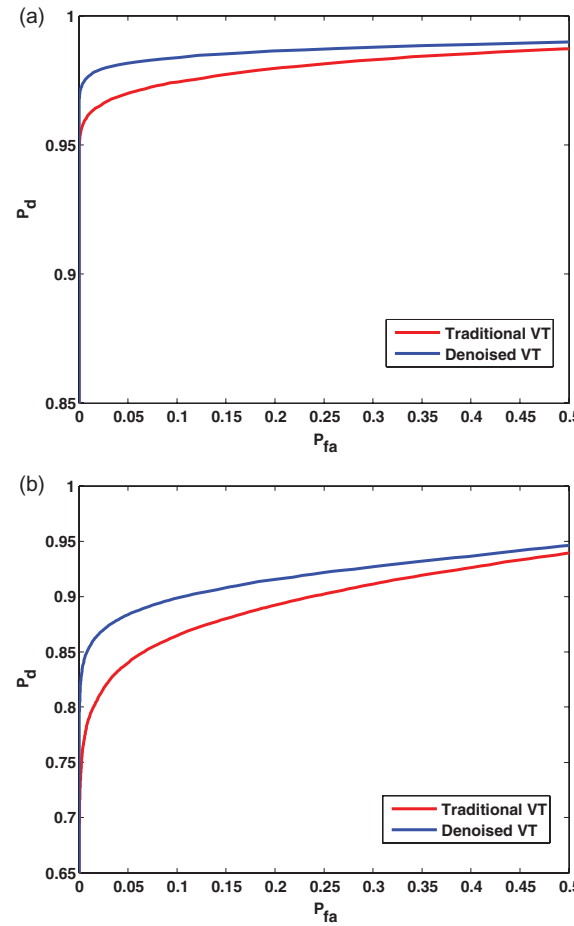

Fig. 7. Probability of false alarm (P_{fa}) *versus* probability of detection (P_{d}) ROC curves for Traditional VT and Denoised VT techniques at (a) SNR = 6 dB and (b) SNR = 0 dB.

Published in 2009 by John Wiley & Sons, Ltd.

DOI: 10.1002/sec

4.2. Burst Location

To isolate the effects of burst location accuracy from burst detection error, the RF bursts were manually detected prior to burst location analysis. Thus, there is no noise-only data input to this process to generate false alarms and $P_d = 100\%$. Plots in this section share two common attributes, including (1) the correct burst locations occur at Time = 0 s and (2) the default non-convergent solutions occur at Time = 1.5 μs (see Step 2 of Section 3.2). For all $-3 \leq \text{SNR} \leq 10$ dB, the Denoised VT technique outperforms the Traditional VT technique by maintaining a more precise and accurate burst start location estimate while converging to a solution more often (fewer non-convergent solutions).

These results illustrate the combined effects of noise and burst-to-burst signal variability. In this case, 500 AWGN realizations were scaled for each SNR and added to each of the 200 collected bursts—a total of 100 000 unique AWGN realizations per SNR. Results for Traditional VT and Denoised VT estimation are shown in Figure 8. Analysis of these results indicates that the Denoised VT technique outperforms the Traditional VT technique by (1) correctly locating 74.7% more of the burst start locations while (2) experiencing 30.2% fewer non-convergent solutions, i.e., fewer estimated starts at the default value of Error = 1.5 μs.

In an operational implementation, only those bursts causing location convergence according to Step 2 of Section 3.2 would be used for further processing. Therefore, for comparing classification performance of the two techniques this research only used bursts that resulted in a converged location solution with *both* techniques. As per Section 3.2, these bursts are denoted as being 'dual convergent'. All other bursts that resulted in a converged location solution with only one of the two technique are excluded from subsequent classification given they could unduly bias results towards the technique having more converged solutions. The burst start location error probability distribution for 'dual convergent' bursts is shown in Figure 9 for two analysis SNRs. The distribution differences (and their associated 'fingerprints') account for the only differences between the two techniques being processed by the classifier.

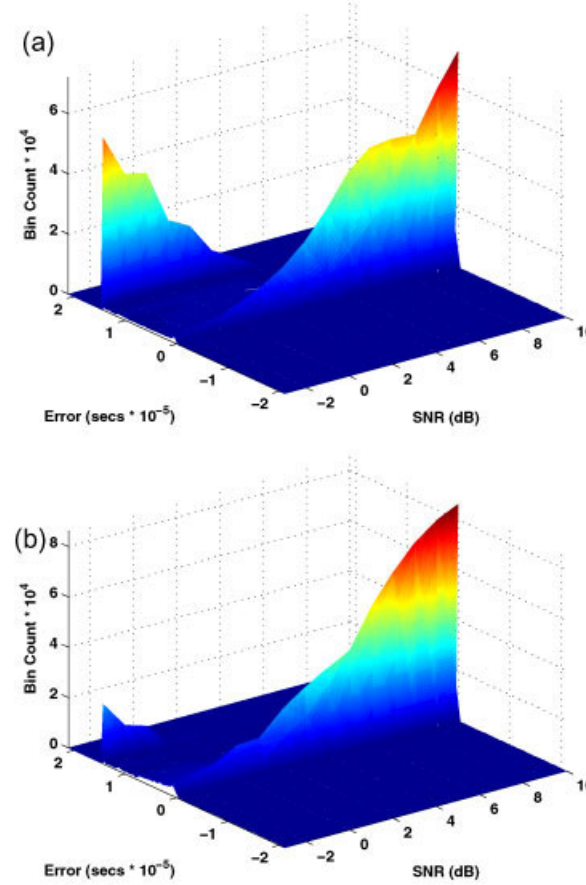

Fig. 8. Burst location histograms showing impact of *Combined* variation across bursts and AWGN realizations using 200 bursts and 500 independent noise realizations. Non-convergent algorithm solutions located at default value of Error = 1.5 μs. (a) Traditional VT and (b) Denoised VT.

4.3. Burst Classification

A total of 200 802.11a bursts were collected from three different devices and used to demonstrate the impact of burst estimation on MDA-ML classification performance. The multi-dimensional MDA-ML input data represents the signal 'fingerprint' which consists of variance, kurtosis, and skewness statistics calculated over the instantaneous amplitude, instantaneous frequency, and instantaneous phase responses of the 802.11a preamble region.

A K-fold cross validation and Monte Carlo process was used to ensure statistical significance in simulated results. While the actual required value of K to produce validated results can vary as a function of data 'behavior', the values of $K = 5$ and $K = 10$ are common choices for K-fold cross validation [35]. As an example, for 200 'dual convergent' bursts per device with $K = 5$ cross validation, the input data set is partitioned into five equal subsets (40 bursts each),

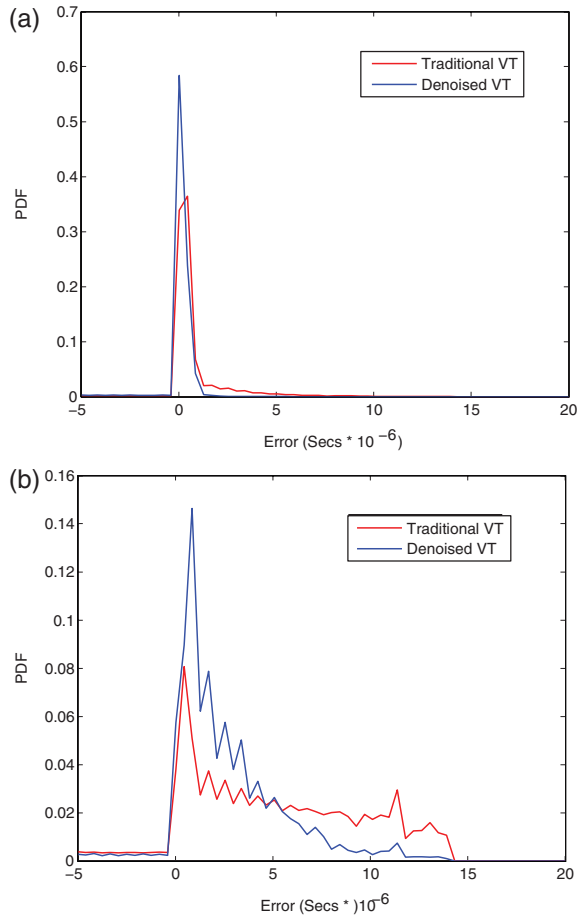

Fig. 9. Burst location error probability for distribution for 'dual convergent' bursts: (a) SNR = 6 dB and (b) SNR = 0 dB.

with four subsets (160 bursts) used for training and the remaining 'held out' subset (40 bursts) used for classification [35].

The overall *K*-fold cross validation and Monte Carlo simulation process included (1) generating, filtering, and scaling AWGN to achieve the desired SNR; (2) estimating the burst start location using the method being evaluated; (3) determining which bursts are 'dual convergent', as per Section 4.2, and determining the minimum 'dual convergent' bursts for the three devices to ensure an equal number of samples for each device; (4) generating projection matrix **W** as per Step 4 of Section 3.2 using the first 80% 'dual convergent' bursts from each device for training the projection matrix **W** for MDA and computing the parameters for the likelihoods used for the ML classifier; (5) transforming the remaining 20% 'dual convergent' bursts from each device as 'unknown' input data using **W** and classifying each per ML criteria; (6) storing/accumulating classification results; (7) circularly shifting (re-ordering) the collected bursts by 20%; (8) repeating Step 4 through Step 7 a total of four more times; (9) repeating Step 1 through Step 8 a total of 500 times using different independent AWGN realizations for each iteration; (10) averaging accumulated classification results from Step 6 to obtain overall classification performance.

A similar process was used in Reference [14] to show that MDA-ML classification performance with Traditional VT location estimation approaches that of Perfect location estimation for $10 \leq \text{SNR} \leq 30$ dB, with notably poorer performance achieved for $-3 \leq \text{SNR} \leq 10$ dB. Therefore, this work specifically proposes the use of denoising in the lower SNR region where performance improvement is realizable. This is done by iteratively applying the above process for $-3 \leq \text{SNR} \leq 10$ dB in 1.0 dB steps. In practice, the decision on whether or not to employ denoising could be based on SNR estimates.

Results in Figure 10 shows average MDA-ML classification performance including burst detection error effects for Perfect, Traditional VT, and Denoised VT burst detection methods. In this case, 'Perfect' results were obtained using a manual burst detection and location process based on visual inspection of instantaneous amplitude responses from each burst. As shown, the Perfect results provide an upper bound on achievable performance. Results for the VT and Denoised VT methods are similar for SNR $>$ 6 dB and SNR < -2 dB. For $-1 <$ SNR < 5 dB, the

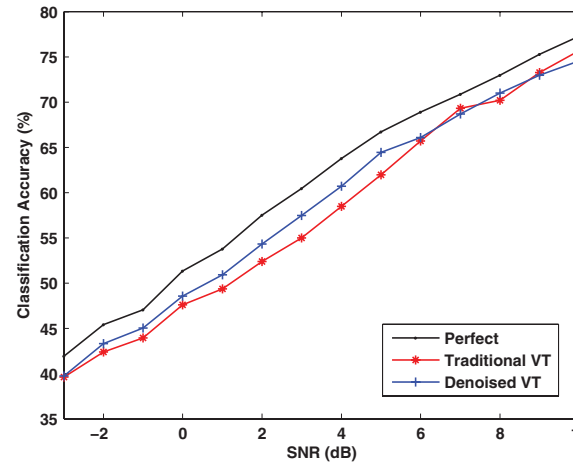

Fig. 10. Average MDA-ML classification accuracy for 'dual convergent' bursts including burst detection error effects for various burst detection methods.

Table I. MDA/ML classification confusion matrix for various burst detection methods at SNR = 3 dB.

Input class	Class estimate A	B	C
	Perfect		
A	**68%**	21%	11%
B	31%	**44%**	25%
C	14%	17%	**69%**
	Traditional VT		
A	**67%**	22%	11%
B	31%	**42%**	27%
C	22%	21%	**57%**
	Denoised VT		
A	**67%**	21%	12%
B	30%	**43%**	27%
C	18%	19%	**63%**

Denoised VT technique outperforms the Traditional VT technique and provides an average improvement in classification accuracy of 1.75%, which is a 34% improvement towards the Perfect case upper bound.

Classification performance is commonly illustrated using a confusion matrix that shows the percentage of time a particular input class is estimated as one of the possible classes, with the bold diagonal entries representing correct classification performance for a given device. Classification results for the SNR = 3 dB data points in Figure 10 are shown in Table I. As indicated in the table, the greatest improvement that Denoised VT provided when compared with Traditional VT is in correctly classifying *Class C*. Comparing Perfect to Traditional VT, there is only a 1–2% margin for improvement in correctly classifying *Class A* and *Class B*, respectively. However, there is a 12% margin for improvement for *Class C*. Here, the Denoised VT provides a 6% gain which represents a 50% improvement towards achieving Perfect transient location performance.

5. Conclusion

Previous work provided noise sensitivity analysis for VT burst detection and highlighted the need for more robust processing at lower SNRs. A new technique is presented here that uses a DT-ℂWT to denoise signals and improve overall VT capability. Instantaneous amplitude responses from collected 802.11a signals were used to demonstrate performance of the Denoised VT technique at varying SNR. As implemented with DT-ℂWT processing, the Denoised VT technique outperforms the Traditional VT technique in all areas, including burst detection, burst start location, and device classification. For burst detection, the denoising technique provides more positive detections for a given false alarm rate. For burst start location, the denoising technique is 74.7% more precise in finding burst start locations and experiences 30.2% fewer non-convergent solutions. Finally, device classification performance was demonstrated using extracted RF fingerprints and MDA with ML classification. Relative to the Traditional VT technique, the Denoised VT process emerges as a better alternative at lower SNRs and yields a classification performance increase of 1.75% (on average) or a 34% improvement towards achieving 'perfect' burst location estimation performance.

Acknowledgments

This research was supported by the Sensors Directorate, Air Force Research Laboratory, and the Tactical SIGINT Technology (TST) Program.

References

1. Chen Y, Trappe W, Martin R. Detecting and localizing wireless spoofing attacks. *IEEE Conference on Sensor, Mesh and AdHoc Comm and Nets (SECON)*, June 2007; 193–202.
2. Sheng Y, Tan K, Chen G, Kotz D, Campbell A. Detecting 802.11 MAC layer spoofing using received signal strength. *IEEE 27th Annual Conference on Computer Communications*, April 2008.
3. Suski WC, Temple MA, Mendenhall MJ, Mills RF. Using spectral fingerprints to improve wireless network security. In *Proceedings of the 2008 IEEE Global Communications Conference*, November–December 2008.
4. Suski WC, Temple MA, Mendenhall MJ, Mills RF. Radio frequency fingerprinting commercial communication devices to enhance electronic security. *International Journal Electronic Security and Digital Forensics* 2008; **1**(3): 301–322.
5. Ureten O, Serinken N. Wireless security through RF fingerprinting. *Canadian Journal of Electrical and Computer Engineering* 2007; **32**(1): 27–33.
6. Langley LE. Specific emitter identification (SEI) and classical parameter fusion technology. *IEEE Western Electronics Show and Conference (WESCON)*, September 1993; 277–381.
7. Hall J, Barbeau M, Kranakis E. Using transceiverprints for anomaly based intrusion detection. *3rd IASTED International Conference on Communication, Internet and Information Technology (CIIT)*, November 2004.
8. Hall J, Barbeau M, Kranakis E. Detection of transient in radio frequency fingerprinting using signal phase. *IASTED International Conference on Wireless and Optical Communications*, May 2003.

9. Ureten O, Serinken N. Detection of radio transmitter turn-on transients. *IEE Electronics Letters* 1999; **35**(23): 1996–1997.
10. Ureten O, Serinken N. Bayesian detection of WiFi transmitter RF fingerprints. *IEE Electronics Letters* 2005; **41**(6): 373–374.
11. Serinken N, Ureten O. Generalised dimension characterization of radio transmitter turn-on transients. *IEE Electronics Letters* 2000; **36**(12): 1064–1064.
12. Dudczyk J, Matuszewski J, Wnuk M. Applying the radiated emission to specific emitter identification. *International Conference on Microwaves, Radar and Wireless Communications*, May 2004; 431–434.
13. Kawalec A, Rapacki T, Wnuczek S, Dudczyk J, Owczarek R. Mixed method based on intrapulse data and radiated emission to emitter sources recognition. *15th International Conference on Microwaves, Radar and Wireless Communications*, May 2004.
14. Klein RW, Temple MA, Mendenhall MJ, Reising DR. Sensitivity analysis of burst detection and rf fingerprinting classification performance. In *Proceedings of the 2009 IEEE International Conference on Communications*, June 2009.
15. Ureten O, Serinken N. Improved coarse timing for burst mode ofdm. *Global Telecommunications Conference, 2007. GLOBECOM '07. IEEE*, November 2007; 2841–2846.
16. Ureten O, Pacheco RA, Serinken N, Hatzinakos D. Bayesian frame synchronization for 802.11a wlans: experimental results. *Canadian Conference on Electrical and Computer Engineering, 2005*, May 2005; 884–887.
17. Pacheco RA, Ureten O, Hatzinakos D, Serinken N. Bayesian frame synchronization using periodic preamble for ofdm-based wlans. *IEEE Signal Processing Letters* 2005; **12**(7): 524–527.
18. Azzouz E, Nandi A. *Automatic Modulation Recognition of Communication Signals*. Kluwer Academic: Boston, 1996.
19. Carmona RA. Wavelet identification of transients in noisy time series. In *Proceedings of SPIE: Wavelet Applications in Signal and Image Processing*, Vol. 2034, November 1993; 392–400.
20. Frisch M, Messer H. Detection of a transient signal of unknown scaling and arrival time using the discrete wavelet transform. *1991 International Conference on Acoustics, Speech, and Signal Processing (ICASSP-91)*, Vol. 2, April 1991; 1313–1316.
21. Frisch M, Messer H. The use of the wavelet transform in the detection of an unknown transient signal. *IEEE Transactions on Information Theory* 1992; **38**(2): 892–897.
22. Frisch M, Messer H. Detection of a known transient signal of unknown scaling and arrival time. *IEEE Transactions on Signal Processing* 1994; **42**(7): 1859–1863.
23. Del Marco SP, Weiss JE. M-band wavepacket-based transient signal detector using a translation-invariant wavelet transform. *Optical Engineering* 1994; **33**: 2175–2182.
24. Fernandes Da Rocha Pitta J, Hippenstiel R, Fargues M. Transient detection using wavelets. *Master's Thesis*, Naval Post Graduate School, Monterey, CA, March 1995.
25. Del Marco SP, Weiss JE. Improved transient signal detection using a wavepacket-based detector with an extended translation-invariant wavelet transform. *IEEE Transactions on Signal Processing* 1997; **45**(4): 841–850.
26. Thiruvengadam SJ, Chinnadurai P, Thirumalai Kumar N, Abhaikumar V. Wavelet based signal processing algorithms for early target detection. *2003 Conference on Convergent Technologies for Asia-Pacific Region*, Vol. 3, October 2003; 1175–1179.
27. Dutta A, Anand GV. Detection of transient signals by wavelet packet transform and stochastic resonance. *2004 IEEE Region 10 Conference*, Vol. 1, November 2004; 251–254.
28. Fabbroni L, Vannucci M, Cuoco E, Losurdo G, Mazzoni M, Stanga R. Wavelet tests for the detection of transients in the virgo interferometric gravitational wave detector. *IEEE Transactions on Instrumentation and Measurement* 2005; **54**(1): 151–162.
29. Selesnick IW, Baraniuk RG, Kingsbury NC. The dual-tree complex wavelet transform. *IEEE Signal Processing Magazine* 2005; **22**(6): 123–151.
30. Bayram I, Selesnick IW. On the dual-tree complex wavelet packet and *m*-band transforms. *IEEE Transactions on Signal Processing* 2008; **56**(6): 2298–2310.
31. IEEE Computer Society. *IEEE Std 802.11-2007*, June 2007.
32. Fisher RA. The use of multiple measurements in taxonomic problems. *Annals of Eugenics* 1936; **7**: 179–188.
33. Duda RO, Hart PE, Stork DG. *Pattern Classification* (2nd edn). John Wiley & Sons, Inc.: New York, 2001.
34. Agilent Technologies Inc., USA. *Agilent E3238 Signal Intercept and Collection Solutions: Family Overview*, Publication 5989-1274EN, July 2004.
35. Hastie T, Tibshirani R, Friedman J. *Data Mining, Inference, and Prediction*. Springer: New York, USA, 2001.