# Dimensional Reduction Analysis for Constellation-Based DNA Fingerprinting to Improve Industrial IoT Wireless Security

Christopher M. Rondeau
Department of Electrical and Computer Engineering
US Air Force Institute of Technology, WPAFB OH
christopher.rondeau@afit.edu

Michael A. Temple
Department of Electrical and Computer Engineering
US Air Force Institute of Technology, WPAFB OH
michael.temple@afit.edu

J. Addison Betances
Department of Electrical and Computer Engineering
US Air Force Institute of Technology, WPAFB OH
jbetance@afit.edu

## Abstract

*The Industrial Internet of Things (IIoT) market is skyrocketing towards 100 billion deployed devices and cybersecurity remains a top priority. This includes security of ZigBee communication devices that are widely used in industrial control system applications. IIoT device security is addressed using Constellation-Based Distinct Native Attribute (CB-DNA) Fingerprinting to augment conventional bit-level security mechanisms. This work expands upon recent CB-DNA "discovery" activity by identifying reduced dimensional fingerprints that increase the computational efficiency and effectiveness of device discrimination methods. The methods considered include Multiple Discriminant Analysis (MDA) and Random Forest (RndF) classification. RndF deficiencies in classification and post-classification feature selection are highlighted and addressed using a pre-classification feature selection method based on a Wilcoxon Rank Sum (WRS) test. Feature down-selection based on WRS testing proves to very reliable, with reduced feature subsets yielding cross-device discrimination performance consistent with full-dimensional feature sets, while being more computationally efficient.*

## 1. Introduction

Industrial Internet of Things (IIoT) devices are a specialized subset of IoT devices that support sensing and control in areas such as water treatment, power generation and distribution, oil and gas refinement and distribution, and transportation. These critical infrastructure elements are commonly operated through Industrial Control System (ICS) architectures and it is estimated that a few billion consumer devices will be IoT connected by 2020. This is orders-of-magnitude lower than some projections which suggest the number of deployed IIoT devices is rocketing towards reaching 100 billion by 2020 [21].

If reliable, these estimates suggest that some IIoT deployment barriers have been overcome while others of remain. This includes technology-based security factors that contribute to cybersecurity being the top-ranked challenge in IIoT device deployment [19]. When considered in light of ICS applications, improved cybersecurity is absolutely essential and protection remains a national-level priority within both the public and private sectors [8, 16, 30, 31]. The cybersecurity challenges are not unique to IIoT devices, and the need for increasingly secure and reliable communications remains across other commercial applications and automation networks supporting medical, home and building automation, consumer electronics.

Cyberattack mitigation approaches for specifically targetted communication devices (e.g., IIoT access points) have primarily focused on bit-level solutions implemented in upper communication protocol layers. This includes the network and media access control layers with much less emphasis placed on physical (PHY) layer solutions [9, 13, 26]. The underutilized PHY information [32] can be captured in device Distinct Native Attribute (DNA) features that provide human-like discrimination and support a multi-factor authentication framework that benefits from combined first-level "something you have" (device address), second-level "something you know" (network encryption key), and final "something you are" (PHY DNA Fingerprint) checks [5, 22]. The goal is to realize multi-factor authentication benefits and improve device verification reliability [3, 20] by taking advantage of the speed and computational efficiency of biometric-based multi-factor authentication which make it a top-ranked choice for IoT applications [10].

URI: https://hdl.handle.net/10125/60149
ISBN: 978-0-9981331-2-6

### 1.l. Relationship to Prior Work

As a sub-class to the broader Radio Frequency (RF) fingerprinting domain, RF-DNA Fingerprinting has been successfully demonstrated in various applications [14, 18, 23], with specific use for IoT communication devices demonstrated in [7, 29] and IIoT devices demonstrated in [15, 27, 28]. Of greater relevance here is the most recent development and demonstration of Constellation-Based DNA (CB-DNA) Fingerprinting [25] and its use in reliably discriminating ZigBee devices. As with a majority of related RF-DNA works, device discrimination results in [25] are based on a Multiple Discriminant Analysis, Maximum Likelihood (MDA/ML) classification process.

This work expands upon results in [25] using the *same* experimentally collected ZigBee signals, with an alternate Random Forest (RndF) classifier introduced given its broad success using RF-DNA fingerprints and support for *post-classification* Dimensional Reduction Analysis (DRA), i.e., identifying the most relevent sub-set of fingrprint features required for reliable device discrimination. Resultant RndF performance with CB-DNA features was not promising and included 1) a *significant decrease* in average cross-class percent correct classification (%C) when compared with MDA under identical conditions, and 2) the generation of *unreliable* RndF variable importance metrics which voided their use for DRA assessments.

Degraded performance of the usually-comparable-to-MDA RndF classifier and its ineffectiveness for reliable DRA motivated the introduction and first use success of *pre-classification* DRA based on a Wilcoxon Rank Sum (WRS) test. As developed and successfully demonstrated here, the WRS-based DRA method extends the concept of distribution-free assumptions and utilizes nonparametric statistical techniques for cross-class feature comparison.

### 1.2. Paper Organization

The remainder of this paper is organized as follows. Section 2.1 provides background information in sub-sections providing details on DNA-Based Device Discrimination, Experimental ZigBee Signals, CB-DNA Fingerprint Generation, and Classifier Models. Section 3 provides baseline performance results of the Multiple Discriminant Analysis, Maximum Likelihood (MDA/ML) and Random Forest (RndF) classifiers. Section 4 provides details for the Wilcoxon Rank Sum (WRS) test, its use for feature selection using Dimensional Reduction Analysis (DRA), and comparison of WRS vs. RndF DRA results. Section 5 provides the paper summary and conclusion.

## 2. Background

### 2.1. DNA-Based Device Discrimination

Previous works applied DNA fingerprinting methods to both wired and wireless signal PHY-based discrimination [4, 14, 29]. This work focuses on CB-DNA Fingerprinting applied to ZigBee devices that are representative of the 802.15.4 standard class of communication devices supporting ICS applications. The specific focus here is on device ID classification as a means to enhance overall ZigBee network security at the PHY doorway through which a preponderance of malicious cyberattacks occur. The concentration here was on securing Zigbee device operation given that 1) ZigBee devices and related 802.15.4 protocols are deployed world-wide, and 2) ZigBee operation is a representative protocol for broader IIoT applications [7, 29]. The degree of required anti-hacking security varies with ZigBee application criticality and will continue to increase as the number of connected IoT reaches a few billion and the number of connected IIoT devices a 100 billion by 2020 [21]. The increased security risks due to rapid expansion may be offset as the next generation of IIoT hardware technologies are evolving to include multi-protocol 802.15.4/Bluetooth/WiFi operation [24].

The development here was motivated by two key observations of CB-DNA Fingerprinting performance when applied to ZigBee devices [25]. First, there was a desire to consider alternate classification techniques and/or conditional fingerprint features with a goal of reducing complexity (processing time and storage) to better support real-time network security using the multi-factor authentication framework detailed in the introduction. The first action included a quick-look assessment using a RndF ensemble classifier given its previous success in related applications, i.e., it achieved near-equivalent classification accuracy when compared to the MDA classifier [14, 18]. The preliminary RndF CB-DNA findings were not promising and included a *significant decrease* in average cross-class percent correct classification (%C) when compared with MDA using the *same* input CB-DNA fingerprints at the *same* Signal-to-Noise Ratio (SNR); this anomalous degradation has not been previously observed in other MDA and RndF classifier works.

Second, there remains some question regarding the impact of CB-DNA feature number and feature "information" content on performance. For the Atmel ZigBee devices used here, [25] shows that CB-DNA Fingerprinting provided improved classification performance relative to RF-DNA Fingerprinting, with the best overall performance obtained for conditional versus unconditional CB-DNA features. Of note is the

disparity between the fingerprint dimensions which included $N_{Ucnd} = 36$ unconditional and $N_{Cnd} = 120$ conditional features. Thus, it is reasonable to ask if the better conditional fingerprinting performance is due to increased feature "information" or simply a matter of using more features.

Degraded performance of the usually-better-than-MDA RndF classifier and increased classification performance using an increased number of conditional features motivated the use of DRA here for CB-DNA Fingerprinting. The two DRA techniques considered here included 1) the *post-classification* RndF method used for the initial quick-look assessment, with additional analyses conducted for multiple DRA subsets based on rank-ordered Gini variable importance indices, and 2) the first use of a *pre-classification* WRS test that yields a comparative metric reflecting feature relevance and enabling rank-ordered DRA subset feature selection. The WRS-based DRA method here extends the concept of distribution-free assumptions and utilizes nonparametric statistical techniques for cross-class feature comparison. A head-to-head comparison of RndF-based and WRS-based DRA feature selection was completed using MDA/ML classification with a given number of $N_{DRA}$ selected features.

## 2.2. Experimental ZigBee Signals

To enable direct comparison, results here are based on fingerprints generated from the *same* experimentally collected ZigBee signals used in [25]. For compleness, a summary of the experimental collection details are provided here (see [25] for additional details). Emissions were collected from the $N_D = 10$ like-model Atmel devices in a relatively benign (limited multipath) office environment using an Ettus USRP X310 radio with the collection bandwidth set to $W_{Coll} = 10$ MHz and operating at a sample frequency of $f_{Samp} = 10$ MSps per I/Q channel. Post-collection processing included down-conversion and baseband filtering with a 16$^{th}$-order Butterworth filter having a –3.0 dB bandwidth of $W_{BB} = 2$ MHz.

Following post-collection processing, the average estimated SNR across all devices and collections was $SNR_C \approx 40.0$ dB. Prior to CB-DNA fingerprint generation per Section 2.3, additional processing was applied that included 1) constellation de-rotation (phase synchronization) on a burst-by-burst basis, along with 2) SNR-scaling by adding independent, like-filtered ($W_{BB} = 2$ MHz), power-scaled Additive White Gaussian Noise (AWGN) to set the desire analysis range of $SNR \in [4.0, 40.0]$.

Complete details for the Atmel ZigBee devices can be found in [1, 2], with some details summarized here in Table 1 for completeness. the Atmel devices include an AT86RF230 radio transceiver [2] that supports 802.15.4 standard compliant operation in the 2.4 GHz band. As low cost, low power alternatives the Atmel devices are widely used in wireless sensor network, ICS, home and building automation, and many other consumer electronic applications [1]. The general ZigBee PHY layer characteristics for these devices are are shown in Fig. 1. As typical in other wireless protocols, each transmitted ZigBee burst includes a preamble response (first 8.3 mSec) that is the primary region of interest exploited for RF-DNA Fingerprinting. For the conditional CB-DNA fingerprint generation process described in Section 2.3 and subsequent results, statistical features are extracted from *all* received communication symbols within each burst.

**Table 1. ZigBee physical operating characteristics.**

| | |
|---|---|
| **FREQUENCY** | 2.4 GHz, 868 MHz, 915 MHz |
| **BIT RATE** | 20-250 Kbits/s |
| **SECURITY** | PHY - None<br>Network - AES 128 |
| **LATENCY** | ≈ 1000 mSec |
| **RANGE** | 1-75 Meters |
| **MODULATION** | O-QPSK |

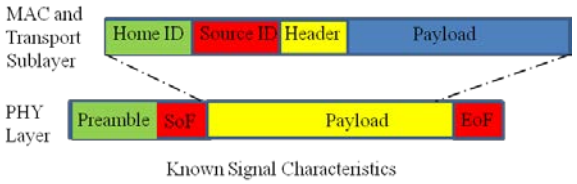



**Figure 1. ZigBee protocol layer components.**

## 2.3. CB-DNA Fingerprint Generation

Previous DNA-based works have used any number of available DNA types, including RF-DNA extracted from both intentional and unintentional radiated emissions, Wired Signal DNA (WS-DNA) [14, 15], and most recently CB-DNA [25]. The CB-DNA fingerprints used here are identical to those in [25] and their development is presented here for completeness. The development is based on an arbitrary complex sequence {X} having $N_X$ elements and the same $N_{Stat} = 14$ fingerprint features (statistics) are calculated for both the 1) *polar* magnitude (*Mag*) and angle (*Ang*) components, and 2) *rectangular* real (*Re*) and imaginary (*Im*) components of {X}.

The specific CB-DNA features used here for polar representation features included variance ($\sigma^2$), skewness ($\gamma$) and kurtosis ($\kappa$) of both the magnitude {*Mag*[X]} and angle {*Ang*[X]} sequences (6 total polar statistics). Rectangular representation features are calculated using the $[Re\{X\}:Im\{X\}]_{2xN_X}$ matrix, with calculated statistics including three unique co-variance $\sigma^2\sigma^2_{(1:3)}$ values, two non-trivial co-skewness moments $\gamma\gamma_{(1:2)}$, and three non-trivial co-kurtosis $\kappa\kappa_{(1:3)}$ moments. Accounting for all possible statistics, the *Statistical Fingerprint* vector for complex sequence {X} is formed as [25],

$$\begin{aligned} \mathbf{F}^{\mathrm{X}} &= \left[\sigma^2_{\mathrm{Mag(X)}} \;\; \gamma_{\mathrm{Mag(X)}} \; \kappa_{\mathrm{Mag(X)}}\right], \\ \mathbf{F}^{\mathrm{X}} &= \mathbf{F}^{\mathrm{X}} \vdots \left[\sigma^2_{\mathrm{Ang(X)}} \;\; \gamma_{\mathrm{Ang(X)}} \; \kappa_{\mathrm{Ang(X)}}\right], \\ \mathbf{F}^{\mathrm{X}} &= \mathbf{F}^{\mathrm{X}} \vdots \left[\sigma\sigma^2_{\mathrm{X}(1:3)} \; \gamma\gamma_{\mathrm{X}(1:2)} \; \kappa\kappa_{\mathrm{X}(1:3)}\right]_{1\times N_{Stat}}, \end{aligned} \quad (1)$$

where $\vdots$ denotes concatenation. For *conditional* CB-DNA Fingerprinting the $\mathbf{F}^X$ in (1) are calculated for $N_{SG}$ selected conditional subgroups of the received signal constellation. The $n$ =1, 2, … $N_{SG}$ subgroup elements for the $m^{th}$ symbol in the M-ary signaling constellation are used to form the $m^{th}$ *Conditional CB-DNA Fingerprint Vector* $\mathbf{F}^{\mathrm{CND}}_{\mathrm{m}}$ according to,

$$\mathbf{F}^{CND}_{m} = \left[\mathbf{F}^{X}_{SG(m,1)} \vdots \mathbf{F}^{X}_{SG(m,2)} \vdots \cdots \vdots \mathbf{F}^{X}_{SG(m,N_{SG})}\right]_{1\times(N_{Stat}\cdot} \quad 2)$$

which are then concatenated to form the *Composite Conditional CB-DNA Fingerprint Vector* as,

$$\mathbf{F}^{CND}_{CB} = [\mathbf{F}^{CND}_{1} \vdots \mathbf{F}^{CND}_{2} \vdots \cdots \vdots \mathbf{F}^{CND}_{M}]_{1\times(N^{CCB}_{F})}, \quad (3)$$

where $N^{CCB}_{F} = N_{Stat} \times N_{SG} \times M$ is the total number of conditional CB-DNA features. In general, both unconditional and conditional CB-DNA fingerprint features can be generated using all or a subset of noted statistics, calculated for all or a subset of available projected groups. Results here are based solely on conditional fingerprints given their demonstrated superiority over unconditional fingerprints using the selected ZigBee signals [25]. The full-dimensional conditional fingerprints include $N_{FD}$ = 270 features and are equally divided into $N_{TRN}$ = 550 training and $N_{TST}$ = 550 testing observations per ZigBee device.

## 2.4. Classifier Models

**2.4.1. Multiple Discriminant Analysis/Maximum Likelihood (MDA/ML).** MDA is a computationally efficient process that has provided reliable device ID discrimination in prior DNA-based works [4, 7, 14, 29]. It is a multi-class extension of Fisher's Linear Discriminant Analysis (LDA) and performs best when the input fingerprint features and their corresponding projections via the eigenvector-based projection matrix **W** are Gaussian distributed. Consistent with the Fisher criterion, the matrix **W** maximizes the ratio of between-class spread (projected class means) to within-class spread (projected class variance). For discrimination of $N_{Cls}$ classes using input fingerprint vectors have $N_F$ features, the projection matrix **W** is of dimension $N_F$ x ($N_{Cls}$-1) and is used to project (1x$N_F$)-dimensional input fingerprints (**F**) into the ($N_{Cls}$-1)-dimensional class estimation space.

Given a trained MDA model that includes matrix **W**, input fingerprint scale factors, projected class training means, and projected class training variances, a 1 vs. $N_{Cls}$ called-class estimate (correct or incorrect) for an "unknown" input testing fingerprint $\mathbf{F}_{\mathrm{Tst}}$ is made by first calculating,

$$\boldsymbol{F}^{\boldsymbol{W}}_{Tst} = \boldsymbol{F}_{Tst}\mathbf{W}, \quad (4)$$

where $\boldsymbol{F}^{\boldsymbol{W}}_{Tst}$ is the projection of $\mathbf{F}_{\mathrm{Tst}}$ in the Fisher space. The classification estimate for $\boldsymbol{F}^{\boldsymbol{W}}_{P}$ is made based on the conditional probability relationship given by,

$$\mathrm{P}\left(c_i|\boldsymbol{F}^{\boldsymbol{W}}_{P}\right) > \mathrm{P}\left(c_j|\boldsymbol{F}^{\boldsymbol{W}}_{P}\right), \quad (5)$$

where $j$ = 1, 2, …, $N_{Cls}$ and $i \neq j$. Assuming $\mathrm{P}(c_i) = 1/N_{Cls}$ for all classes and equal error costs, the relationship in (2) becomes a Maximum Likelihood (ML) estimate that is obtained by maximizing the conditional $\mathrm{P}(\boldsymbol{F}^{\boldsymbol{W}}_{P}|c_i)$ with the class yielding highest probability being the called-class for $\boldsymbol{F}^{\boldsymbol{W}}_{Tst}$ (this is now referred to as an MDA/ML classification process).

**2.4.2. Random Forest (RndF).** The RndF classifier implemented here was based on [6, 14, 17, 18] and includes an ensemble of single decision tree classifiers used to produce a single classification decision. Relative to MDA/ML, RndF benefits include 1) it not being constrained to specific input data distributions and, 2) it provides a measure of feature relevance called variable importance that is available after training [6]. The two fundamental parameters impacting RndF performance include 1) the number of decision trees (classifiers), and 2) the number of predictors (features) sampled at each node. Classifier model development begins by using all observations (i.e., fingerprints containing all features) at a single node. The initial node is then is split into two child nodes, with the split based on a random predictor selection and threshold values for features of each observation at that node.

Not all features are considered at each node and predictor selection is done with replacement, i.e., a given feature may be used as a splitting criterion at multiple nodes. The DNA features selected as the splitting criterion include those producing the largest change in Gini-Index ($G_I$) from a parent node to its children nodes [14]. The index reflects the probability

that a single observation at a given node is from a particular class with $G_I = 0$ occurring when a node contains only one observation from every class. The final grown forest contains final leaf nodes, each of which represents an individual classification decision for a given input fingerprint.

The resultant classifier can be used for post-classification RndF-based DRA feature selection by considering the mean decrease in $G_I$, denoted as $\lambda_{GI(k)}$ [14, 18] and computed for the $k^{th}$ feature by averaging the change in $G_I$ each time the $k^{th}$ feature is used at a splitting decision. The resulting vector of $\lambda_{GI(k)}$ for $k = 1, 2, \ldots, N_F$ is sorted and provides the mechanism feature ranking to form RndF-selected DRA subsets.

## 3. Results: MDA/ML vs. RndF

RndF was introduced to enable 1) one-to-one comparison with prior MDA/ML performance in [25], and 2) assessment of post-classification RndF-based DRA as used successfully in prior work [14]. Results in Fig. 2 shows classification %C vs. SNR performance for MDA/ML and RndF classifiers under selected fingerprinting conditions. The two highest %C curves are MDA/ML classifier performance using $N_{Cond} = 270$ conditional and $N_{Ucnd} = 36$ unconditional CB-DNA fingerprint features as in [25]. These are provided for reference with the highest conditional MDA/ML curve being the baseline for RndF comparisons.

The MDA/ML results in Fig. 2 affirm intuition that adding more features improves accuracy, but a question remains as to whether or not there is a smaller subset of conditional features, closer in number to the unconditional feature set, that could achieve similar accuracy. Given that MDA/ML provides no insight into feature relevance on the final classification decision, post-classification RndF DRA was initially considered. The first DRA step included running the RndF classifier with the full-dimensional $N_{FD} = 270$ conditional CB-DNA fingerprints. These results are provided in Fig. 2 and reflect considerably poorer performance than MDA/ML across all SNR. RndF does achieve the arbitrary benchmark of %C = 90% at SNR = 28 dB but this is approximately $\%C_\Delta \approx 10\%$ poorer than MDA/ML.

As previously stated, the level of degraded RndF %C performance relative to MDA/ML in Fig. 2 is inconsistent with previous works. However, this did not preclude consideration of RndF-based DRA using the rank-ordered Gini indices returned from the RndF classifier. These are shown plotted in Fig. 3 for the SNR = 28 dB model that achieved the arbitrary %C = 90% benchmark. The sorted Gini indices in Fig. 3a are used for DRA feature selection and the

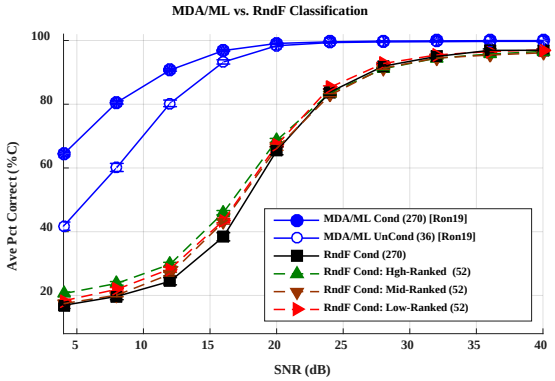



**Figure 2. MDA/ML vs. RndF average classification performance using the indicated features.**

unsorted indices in Fig. 3b showing how the relevant features are distributed within the input fingerprints.

The sorted Gini indices in Fig. 3a reflect the typical highest-to-lowest ascending trend observed with previous signals but the ascension rate is slower with the the least relevant features asymptotically approach near-zero versus zero values; this indicates that nearly all of the conditional CB-DNA features have some relevancy as a classification predictor. The RndF Gini relevancy is investigated by considering three DRA sets comprised of the $N_{DRA} = 52$ highest, middle, and lowest ranked features (the highlighted regions in Fig. 3a).

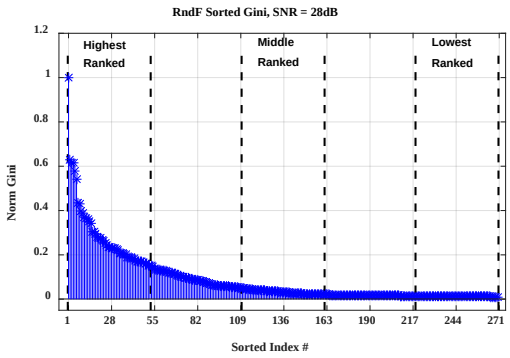
The corresponding classification results for highest (▲), middle (▼), and lowest (►) ranked DRA subsets are shown overlaid in Fig. 2. By comparing all RndF conditional results in Fig. 2, there is minimal impact on RndF classification regardless of the relevancy indicated by Gini indices in Fig. 3. This affirms that selection of features is partly causal to degraded performance, and motivates consideration of an alternate DRA feature selection method that better exploits feature differences. This is addressed using a pre-classification method based on the WRS test.

## 4. Wilcoxon Rank Sum (WRS) Test

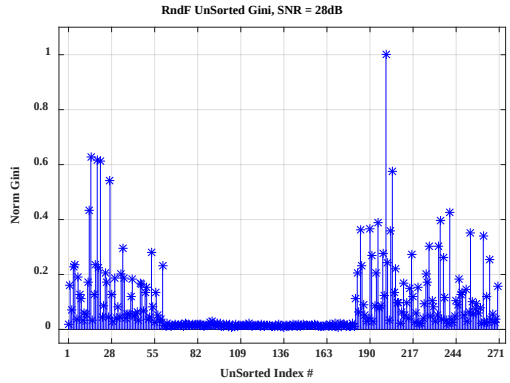
### 4.1. WRS Process Development

There are many nonparametric statistical tests from which to choose, and all of them focus on different aspects of the underlying data distribution. The WRS test is considered to examine the feasibility of using nonparametric statistical analysis as a DRA method where the focus is on the difference in medians of the underlying distributions.

The only assumptions required to employ the WRS test is that the data (input fingerprint features) being




**(a) Sorted Gini indices showing the relationship of the highest, middle, and lowest ranked $N_{DRA}$ = 52 features.**




**(b) Unsorted Gini indices showing the distribution of relevant features witin the input fingerprints.**

**Figure 3. RndF Gini indices for the $N_{FD}$ = 270 conditional fingerprint features that yielded C% ≈ 90% results at SNR = 28 dB in Fig. 2.**

compared are from a continuous distribution and that they are independent (both within their respective population and are mutually independent). Since the comparison for a given fingerprint feature is actually a cross-device comparison, the features are indeed independent given that CB-DNA fingerprints are extracted one-for-one from independently collected device emissions. Moreover, the features are samples from a continuous time domain signal and are therefore from an underlying continuous distribution.

Having satisfied the assumptions, the WRS test considers whether or not the two different classes represented in the feature are from distributions with equal medians [11]. For the purpose of cross-class discrimination and DRA, the goal is to find comparisons that *fail* the WRS test. That is, a failure of the WRS test indicates that, for a given feature and cross-class comparison, regardless of the exact nature of the two underlying distributions, they have different medians. This difference, should it be significant enough, may be exploitable by the MDA/ML or RndF classification processes.

Of note for this class of tests, nonparametric statistical methods in general do not have the same strict assumptions as their equivalent parametric tests. However, to produce a DRA method that is broadly applicable to various DNA Fingerprinting methods, the concept of *asymptotic relative efficiency* of the WRS test is considered. As with many nonparametric statistical tests, the asymptotic relative efficiency of the WRS test suggests that even if the underlying distribution is actually normal, there is a minimal loss of test efficiency (i.e., data required to produce a similar result) and strong agreement between the test outcomes whether using equivalent parametric or nonparametric tests. Therefore, a nonparametric method for DRA can be used in any application if the test's assumptions are met, and regardless of whether or not the equivalent parametric assumptions are met.

The pre-classification DRA methodology here utilizes the WRS metrics of input CB-DNA fingerprints to identify the most relevant features for DRA selection. Given that %C ≈ 90% was achieved at SNR = 28 dB for both the MDA/ML and RndF classifiers in Fig. 2, this is the SNR selected for WRS development and analysis. The WRS test is accomplished for each cross-class comparison and the output is either 1) a value of 1 indicating the test failed to reject the null hypothesis and concluding the distribution medians are not similar, or 2) a value of 0 indicating the null hypothesis test passed and the medians are similar. The results are aggregated into an $N_{Cls} \times N_{Cls}$ upper triangular matrix, where the row/column combination represents the classes being compared. This processes is repeated for each $N_F$ feature and yields an $N_{Cls} \times N_{Cls} \times N_F$ matrix. A simple metric for feature relevance arises by summing the $N_{Cls} \times N_{Cls}$ elements (excluding the diagonal elements) which yields $N_F$ scalars that are denoted by $H_F$ and represent the sum of WRS hypothesis test failures. For example, the maximum value of this summation for a given feature for $N_{Cls}$ = 10 classes is 45 if there is a 1 in all of the upper triangular elements. The final test results are combined to form the vector

$$\boldsymbol{H} = [H_1 \ \cdots \ H_F \ ]_{1\times N_F} . \tag{5}$$

Similar to RndF-based DRA in Section 2.4.2, it was desirable to develop a WRS-based DRA method using a relevance metric that discriminates between seemingly similar features. The hypothesis test presented here outputs a p-value for the WRS test, which may be thought of as how strongly to weigh the decision produced by the test. Since the p-value is isolated to the specific test under which it was conducted, the method of summing the values as previously discussed is

inappropriate. The use of entropy as a RndF Gini-Index alternative [12] motivated consideration of an entropy-based approach here. However, the entropy of an $N_{Cls}$ x $N_{Cls}$ x $N_F$ matrix of p-values provides a scalar representation which can be averaged across the feature to determine the relative weight of the decision criteria at a given feature. Since a low value of entropy is desired, each average value is subtracted from the maximum average value to compute the final average entropy matrix ($\boldsymbol{E}$) of p-values, given by

$$\boldsymbol{E} = \begin{bmatrix} argmax(\boldsymbol{E}) - E_1 \\ \cdots argmax(\boldsymbol{E}) - E_F \end{bmatrix}_{1\times N_F}. \quad (6)$$

To produce a feature relevance metric to score all $N_F$ features, the matrix consisting of the product of the two metrics at the same feature is proposed as,

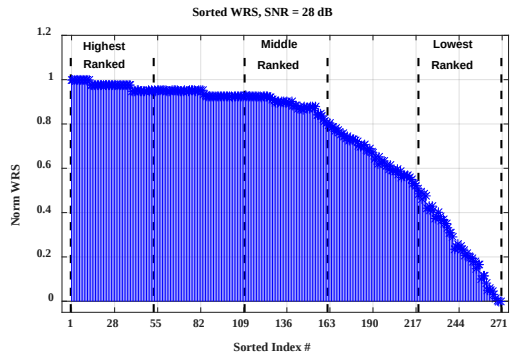
$$\boldsymbol{W_E} = \begin{bmatrix} H_1(argmax(\boldsymbol{E}) - E_1) \\ \cdots H_F(argmax(\boldsymbol{E}) - E_F) \end{bmatrix}_{1\times N_F}, \quad (7)$$

where each element in $\boldsymbol{W_E}$ is the relevance of the feature corresponding to that element according to this entropy-based weighted WRS metric.

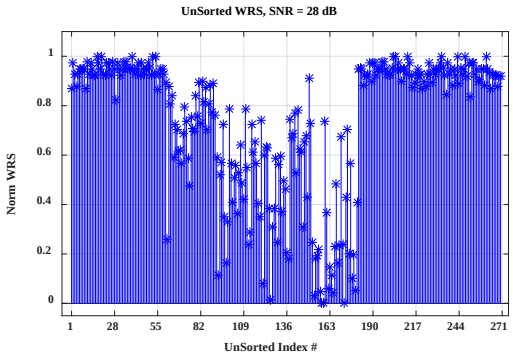
## 4.2. Results: WRS-Based DRA

The normalized weighted WRS metric plots in Fig. 4 for SNR = 28 dB CB-DNA fingerprints include the sorted Fig. 4a metrics used for DRA feature subset selection and unsorted Fig. 4b metrics presented to enable visualization of how relevant features are distributed within the fingerprints. Presentation of assessment results for weighted WRS metric exploitability at SNR = 28 dB is arbitrary and a matter of convenience to enable direct comparison with previously presented RndF results. The Fig. 4 results provide intuition that among the full-dimensional $N_{FD}$ = 270 conditional CB-DNA feature set there exists more relevant proper subsets to be considered for classification DRA assessment.

As with Section 3 RndF assessments, weighted WRS metric relevancy was investigated by considering three DRA sets comprised of the $N_{DRA}$ = 52 highest, middle, and lowest ranked features highlighted in Fig. 4a. The corresponding classification performance for highest (▲), middle (▼), and lowest (►) ranked WRS subsets are shown overlaid in Fig. 5 along with the $N_{FD}$ = 270 full-dimensional performance. The DRA %C trends in Fig. 5 indicate that 1) benefit is realized relative to the less effective RndF-selected DRA performance in Fig. 2, 2) performance is consistent with the highest-to-middle-to-lowest relevancy indicated in Fig. 4b, and 3) pre-classification DRA via rank-ordered weighted WRS metrics is a viable alternative.




**(a) Sorted metrics showing the relationship between the highest, middle, and lowest ranked $N_{DRA}$ = 52 features.**




**(b) UnSorted metrics showing the distribution of relevant features across the input fingerprints.**

**Figure 4. Normalized weighted WRS feature ranking metric for full-dimensional $N_{FD}$ = 270 conditional CB-DNA fingerprints at SNR = 28 dB.**

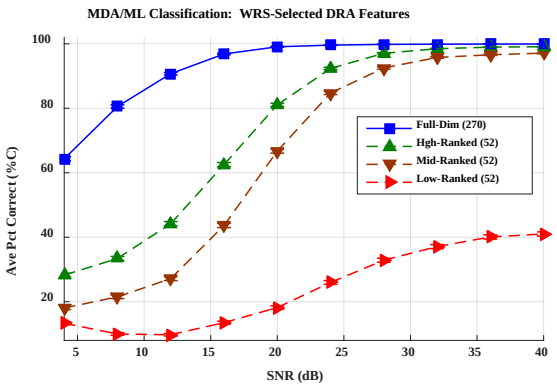



**Figure 5. MDA/ML classification for full-dimensional $N_{FD}$ = 270 CB-DNA features and WRS-selected DRA subsets containing $N_{DRA}$ = 52 highest, middle, and lowest ranked features in Fig. 4a.**

Returning to one of the main goals of reducing the number of required features to achieve a given performance, while minimizing required computation complexity (time, memory, etc.), Fig. 4a weighted WRS rank-ordering was used to analyze MDA/ML classification performance for various DRA subsets. Classification results for the top-ranked $N_{DRA} = 165$, $N_{DRA} = 126$, and $N_{DRA} = 52$ subsets are provided in Fig. 6 and represent an approximate 39%, 53% and 80% fingerprint dimensional reduction, respectively. Of note in Fig. 6 results is that all WRS-selected DRA subsets achieve the arbitrary %C = 90% benchmark for SNR $\leq$ 24 dB, with the $N_{DRA} = 165$ (39% reduced) subset achieving statistical equivalent performance to the full-dimensional set.

### 4.3. Results: WRS vs. RndF DRA

For a final comparison, the RndF Gini relevance ranking in Fig. 3a was revisited and MDA/ML classification results generated for RndF-selected $N_{DRA} = 165$, $N_{DRA} = 126$, and $N_{DRA} = 52$ subsets. These results are presented in Fig. 7 along with the equivalent WRS-selected DRA results from Fig. 6. Considering the performance "gain" ($G_{\Delta dB}$) of DRA selection methods at %C = 90%, calculated as the difference in required SNR for two DRA subsets to achieve the same %C = 90%, the WRS-selected DRA subsets provide $4.0 < G_{\Delta dB} < 8.0$ dB over the RndF-selected subsets.

## 5. Summary and Conclusions

Estimates suggest that the number of deployed IIoT devices could reach 100 billion connected devices by 2020 [21] and technology-based hardware security concerns contribute to making cybersecurity the top-ranked IIoT challenge [19]. Demonstrations here address physical layer based security enhancement using Constellation Based Distinct Native Attribute (CB-DNA) Fingerprinting to reliably discriminate Atmel ZigBee devices that are representative of the 802.15.4 standard class of devices commonly used in IIoT applications. Results here expand upon work in [25] and demonstrate a pre-classification method that provides reliable fingerprint dimensional reduction and identifies feature subsets that 1) achieve discrimination performance of full-dimensional fingerprint sets, while 2) supporting a more efficient implementation of multi-factor device authentication.

Improvements here include the introduction of *pre-classification* Dimensional Reduction Analysis (DRA) based on a Wilcoxon Rank Sum (WRS) test. As developed and demonstrated herein, the WRS-based DRA feature selection process effectively identified a

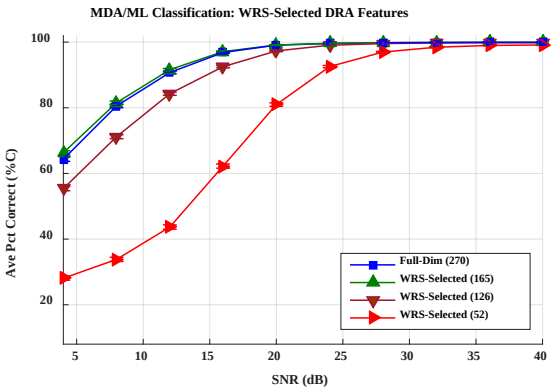



**Figure 6. MDA/ML classification performance for the full-dimensional $N_{FD}$ = 270 CB-DNA feature set and WRS-selected DRA subsets thereof containing $N_{DRA}$ = 165, $N_{DRA}$ = 126, and $N_{DRA}$ = 52, features selected using the sorted ranking in Fig. 4a.**

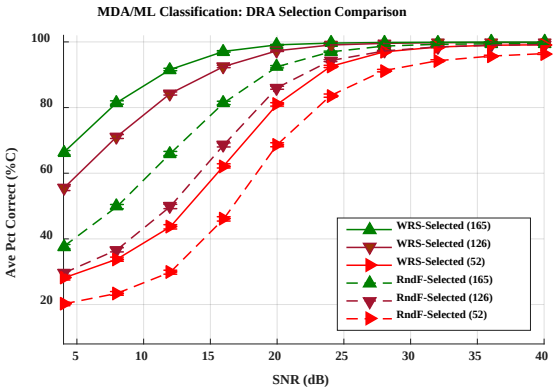



**Figure 7. MDA/ML classification performance showing a comparison of pre-classification WRS-selected features used for Fig. 6 results and corresponding post-classification RndF-selected DRA subsets based on Fig. 3a containing $N_{DRA}$ = 165, $N_{DRA}$ = 126, and $N_{DRA}$ = 52 features selected using the sorted RndF Gini ranking in Fig. 4a.**

reduced dimensional feature subset containing 157 of the 270 full-dimensional features (an approximate 42% reduction) that 1) produced statistically equivalent classification performance as the full-dimensional set while inherently reducing the computational complexity (processing time, memory, etc.) required for real-time security augmentation, and 2) yielded an SNR "gain" (reduction in required SNR to achieve a given average percent correct %C classification performance) of $4.0 < G_{\Delta dB} < 8.0$ dB at %C = 90% when compared with

equal dimension DRA subsets selected using a *post-classification* variable importance metric produced by the Random Forest (RndF) classifier.

The effectiveness of WRS-based DRA for CB-DNA features, and corresponding abysmal performance of the previously effective RndF-based DRA process used for RF-DNA features, is attributed to inherent CB-DNA feature "information" that appears to possess more exploitable characteristics than those occurring in corresponding RF-DNA features of the same collected bursts. This phenomena remains an area of interest for future studies, including the investigation of other pre-classification statistical tests and their potential benefit for DRA feature selection. Collectively considering all results, DNA-based fingerprint discrimination continues to be relevant as it pertains to exploitable physical layer information contained in IIoT communications.

## 6. Acknowledgment

The views in this paper are those of the authors and do not reflect the official policy or position of the Air Force Institute of Technology, the Department of the Air Force, the Department of Defense, or the US Government. This paper is approved for public release, Case#: 88ABW-2018-4444.

## 7. References

[1] Amtel Corporation, "AVR Low Power 2.4 GHz Transceiver for ZigBee." AT86FR230 Spec Sheet, 5131E-MCU Wireless-02/09." 2009.

[2] Atmel Corporation, "R2016: RZRAVEN Hardware User's Guide." Rev. #8117D-AVR-04/08, 2008.

[3] Arce, I., et al., "Avoiding the Top 10 Security Design Flaws." [Online]. Available: https:// cybersecurity. ieee.org/blog/2015/11/13/avoiding-the-top-10-security-flaws/, Nov 2015.

[4] Bihl, T.J., M.A. Temple, and K.W. Bauer, "An Optimization Framework for Generalized Relevance Learning Vector Quantization with Application to Z-Wave Device Fingerprinting." *Proc of the 50th Hawaii Int'l Conference on System Sciences (HICSS17)*, Waikoloa Village, HI, pp. 2379-2387, Jan. 2017.

[5] Bihl, T.J., M.A. Temple, K.W. Bauer, and B.W. Ramsey "Dimensional Reduction Analysis for Physical Layer Device Fingerprints with Application to ZigBee and Z-Wave Devices." *IEEE Military Communications Conf (MILCOM15)*, pp. 360-365. 2015.

[6] Breiman, L., "Random Forests." *Machine Learning*, Vol. 45, No. 1, pp. 5-32, 2001.

[7] Dubendorfer, C.K., B.W. Ramsey, and M.A. Temple, "ZigBee Device Verification for Securing Industrial Control and Building Automation Systems." *Critical Infrastructure Prot VII, IFIP Advances in Information and Communication Technology*, pp. 47-62, 2013.

[8] Erinle, B. "Cyber Security for National Defense." [Onine]. Available: http://themilitaryengineer.com/ index.php/item/259-cyber-security-for-national-defense. 2010.

[9] Field Communications Group, "Connecting the World of Process Automation." [Online]. Available: https:// fieldcommgroup.org/, Nov 2017.

[10] Gupta, U., "Application of Multi-Factor Authentication in Internet of Things Domain." *Int'l Jour of Computer Applications*, Vol. 123, No. 1, pp. 29–31, Aug 2015.

[11] Hollander, M., D.A. Wolfe and E. Chicken. "Nonparametric Statistical Methods." Hoboken: Wiley, 2014.

[12] James, G., D. Witten, T. Hastie, and R. Tibshirani, *An Introduction to Statistical Learning*, New York: Springer Science+Business Media, 2017.

[13] Lennvall, T., S. Svensson and F. Hekland, "A Comparison of Wirelesshart and ZigBee for Industrial Applications." *IEEE Int'l Workshop on Factory Communication Systems*, 2008.

[14] Lopez, J., Jr., N.C. Liefer, M.A. Temple, and C.R. Busho. "Enhancing Critical Infrastructure and Key Resources (CIKR) Level-0 Physical Process Security Using Field Device Distinct Native Attribute Features." *IEEE Trans. Inf. Forenics Secur.*, vol. 13, no. 5, pp. 1215–1229. DOI: 10.1109/TIFS.2017.2779447, 2017.

[15] Lopez, J., M.A. Temple, and B.E. Mullins, "Exploitation of HART Wired Signal Distinct Native Attribute (WS-DNA) Features to Verify Field Device Identity and Infer Operating State." *Springer LNCS*, Vol. 8985, pp. 24-30, Mar 2016.

[16] Mehta, A. "Could an Air Conditioner Take Down a Military Base? The Pentagon is Worried." [Online]. Available: https://www.fifthdomain.com/, 2017.

[17] Patel, H.J., M.A. Temple, and R.O. Baldwin, "Improving ZigBee Device Network Authentication Using Ensemble Decision Tree Classifiers with RF-DNA Fingerprinting." *IEEE Trans on Reliability*, Vol. 64, No. 1, pp. 221-233, 2015.

[18] Patel, H.J., M.A. Temple, R.O. Baldwin, and B.W. Ramsey. "Introduction of a Random Forrest Classifier to ZigBee Device Network Authentication Using RF-DNA Fingerprinting." *Jour of Information Warfare (JIW)*, Vol. 13, No. 3, pp. 33-45, 2014.

[19] PCI Industrial Computer Manufacturer's Group, "Industrial Internet of Things." [Online]. Available: https://www.pwc.com/gx/en/technology/pdf/industrial-internet-of-things.pdf, Nov 2017.

[20] PCI Security Standards Council, "Multi-Factor Authentication." Information Supplement, [Online]. Available: https://www.pcisecuritystandards.org/pdfs/

Multi-Factor-Authentication-Guidance-v1.pdf, Ver. 1, Feb 2017.

[21] PwCIL, “The Industrial Internet of Things.” [Online], Available: https://www.pwc.com/gx/en/technology/pdf/industrial-internet-of-things.pdf, Pricewaterhouse Coopers International Limited, UK, 2016.

[22] Ramsey, B.W., M.A. Temple, and B.E. Mullins, "A PHY Foundation for Multi-Factor ZigBee Node Authentication.” *IEEE Global Communications Conf (GLOBECOM12)*, pp. 795-800, 2012.

[23] Reising, D.R., M.A. Temple, and J.A. Jackson. “Authorized and Rogue Device Discrimination Using Dimensionally Reduced RF-DNA Fingerprints.” *IEEE Trans. Inf. Forensics Secur.*, vol. 10, no. 6, pp. 1180–1192, DOI: 10.1109/TIFS.2015.2400426, 2015.

[24] Research and Markets, "802.15.4 IoT Markets: A Market Dynamics Report.” Market Report ID: #4392927, [Online]. Available: https://www.research-andmarkets.com/research/bwqff5/802_15_4_iot, 2017.

[25] Rondeau, C.M., C.M. Betances, and M.A. Temple, “Securing ZigBee Commercial Communications Using Constellation-Based Distinct Native Attribute (CB-DNA) Fingerprinting.” *Jour of Secure Comm Nets*, vol. 2018, DOI: 10.1155/2018/1489347, 2018.

[26] Stefanidis, K., and A.G. Voyiatzis, "An HMM-Based Anomaly Detection Approach for SCADA Systems." *Int’l Conf on Information Security Theory and Practice (WISTP)*. Springer LNCS 9895, pp. 85–99, DOI: 10.1007/978-3-319-45931-8_6, 2016.

[27] Stone, S.J., M.A. Temple, and R.O. Baldwin, "Detecting Anomalous PLC Behavior Using RF-Based Hilbert Transform Features and a Correlation-Based Verification Process.” *Int’l Jour on Critical Infrastructure Protection 9*, pp. 41-51, Feb 2015.

[28] Stone, S.J., and M.A. Temple, "RF-Based Anomaly Detection for PLCs in Critical Infrastructure Apps.” *Int’l Jour on Critical Infrastructure Protection*, Vol. 5, No. 2, pp. 66-73, 2012.

[29] Talbot, C.M., M.A. Temple, T.J. Carbino, and A.J. Betances. “Detecting Rogue Attacks on Commercial Wireless Insteon Home Automation Systems.” *Jour of Computer Security*, vol. 74, no. C, pp. 296–307, DOI: 10.1016/j.cose.2017.10.001, May 2018.

[30] US Department of Homeland Security, “Critical Infrastructure: Homeland Security Starts with Home-town Security.” US Department of Homeland Security, [Online]. Available: https://www.hsdl.org/?view&did=760074. Aug 2010.

[31] US Department of Homeland Security, "Recommended Practice: Improving Industrial Control Systems Cyber-security with Defense-In-Depth Strategies.” [Online]. Available: https://ics-cert.us-cert.gov/. Sep 2016.

[32] Weiss, J., and J. Lopez. “The Gap in ICS Cyber Security and Safety–Level 0,1 Devices.” *ISA Power Industry Division (POWID) Conf*, Jun 2018.