A Comparison of PHY-Based Fingerprinting Methods Used to Enhance Network Access Control

Timothy Carbino, Michael Temple, Juan Lopez Jr.

A Comparison of PHY-Based Fingerprinting Methods Used to Enhance Network Access Control

Timothy J. Carbino, Michael A. Temple, and Juan Lopez Jr.

Air Force Institute of Technology, Electrical and Computer Engineering,
WPAFB Ohio, 45433, USA
{timothy.carbino,michael.temple,juan.lopez.ctr}@afit.edu

Abstract. Network complexity continues to evolve and more robust measures are required to ensure network integrity and mitigate unauthorized access. A physical-layer (PHY) augmentation to Medium Access Control (MAC) authentication is considered using PHY-based Distinct Native Attribute (DNA) features to form device fingerprints. Specifically, a comparison of waveform-based Radio Frequency DNA (RF-DNA) and Constellation-Based DNA (CB-DNA) fingerprinting methods is provided using unintentional Ethernet cable emissions for 10BASE-T signaling. For the first time a direct comparison is achievable between the two methods given the evaluation uses the same experimentally collected emissions to generate RF-DNA and CB-DNA fingerprints. RF-DNA fingerprinting exploits device dependent features derived from instantaneous preamble responses within communication bursts. For these same bursts, the CB-DNA approach uses device dependent features derived from mapped symbol clusters within an adapted two-dimensional (2D) binary constellation. The evaluation uses 16 wired Ethernet devices from 4 different manufacturers and both Cross-Model (manufacturer) Discrimination (CMD) and Like-Model (serial number) Discrimination (LMD) is addressed. Discrimination is assessed using a Multiple Discriminant Analysis, Maximum Likelihood (MDA/ML) classifier. Results show that both RF-DNA and CB-DNA approaches perform well for CMD with average correct classification of $\%C$=90% achieved at Signal-to-Noise Ratios of $SNR \geq$ 12.0 dB. Consistent with prior related work, LMD discrimination is more challenging with CB-DNA achieving $\%C$=90.0% at SNR=22.0 dB and significantly outperforming RF-DNA which only achieved $\%C$=56.0% at this same SNR.

Keywords: Network Access Control, Physical-layer Distinct Native Attribute, RF-DNA, CB-DNA, Device Fingerprinting, MDA/ML

1 Introduction

Network Access Control (NAC) solutions implement strategies which allow devices and/or users access to a given network. There are many NAC solutions that can be employed by a network administrator to include mapping Medium Access

Control (MAC) addresses to specific ports, device credentials, and querying the hardware and software of a device. Each of these potential solutions suffer from weakness to include an attackers ability to spoof specific device information or steal device credentials. As each year passes technical capability expands and more devices are able to connect to a network. This expansion creates unique security challenges and increases the potential for unauthorized access. Physical-layer (PHY) augmentation of MAC based authentication processes provides one means to improve security and network authentication reliability. The envisioned PHY-augmented authentication process utilizes a device's digital ID (e.g., MAC address) and PHY features extracted from the device's communication signal. Ideally, the device's fingerprint consists of unique PHY features that enable reliable discrimination. The final authentication decision, to allow or deny network access, is based on 1) presentation of an authorized MAC address and 2) a statistical match between the current Distinct Native Attribute (DNA) features of the device presenting the MAC address and the stored DNA for the actual device assigned the MAC address.

The majority of PHY-based fingerprinting methods are based on features generated from transient, invariant or entire burst responses as discussed in the review presented in [1]. It is concluded in [1] that many of the PHY fingerprinting techniques discussed lack proper performance evaluation. It is the goal of this work to conduct performance evaluation between the two most prevalent approaches in [1]. The contributions of the research presented in this paper includes: 1) a direct comparison of performance in waveform-based Radio Frequency DNA (RF-DNA) and Constellation-Based DNA (CB-DNA) approaches by, 2) utilizing for the first time the same collected emissions for both approaches, 3) the CB-DNA approach is expanded for the first time to include *conditional* constellation point sub-clusters, and 4) expand CB-DNA classification to include Like-Model Discrimination (LMD).

The paper is organized as follows. Section 2 provides background information and related work on some of the most recent works in device fingerprinting. Section 3 discusses the experimental setup and outlines the PHY-based RF-DNA and CB-DNA device fingerprinting approaches. This is followed by device discrimination results in Sect. 4 and a summary and conclusions in Sect. 5.

2 Background

Device hardware fingerprinting is possible due to variations in manufacturing processes and device components. These variations inherently induce PHY feature differences that vary across devices [2]. Amplifiers, capacitors, inductors and oscillators also possess slight imperfections that influence device fingerprints [2–5]. The resultant variation can cause deviation in communication symbol rate, center frequency, and induce AM/FM/PM conversion [2]. Thus, it is possible to exploit device imperfections even when the intrinsic components used are supposedly identical [1, 6].

As noted previously, the review in [1] focused primarily on PHY based fingerprinting techniques, with non-PHY based approaches prior to 2009 only briefly addressed. Non-PHY based fingerprinting techniques as in [7–12] are relevant and can be used to fingerprint devices by actively probing or passively monitoring network packet traffic. Fingerprinting is accomplished by exploiting clock-skew via round trip time and inter-arrival time estimation in the collected network traces. These non-PHY based approaches are noted here for completeness and a comparison of PHY based and non-PHY based approaches is the subject of subsequent research.

PHY based device fingerprinting works in [6, 15–19, 21] generally rely on invariant *non-data modulated* Region of Interest (ROI) within the burst (turn-on transient, preamble, midamble, etc.) to extract fingerprint features. Additional works [1–5], utilize the *data modulated* burst response regions to extract their fingerprint features from device dependent modulation errors. Transient-based approaches are generally avoided given 1) the limited duration of the transient response, and 2) the transient response being influenced by environmental conditions that affect the communication channel and limit its usefulness [4]. As noted in Sect. 3.2, the CB approaches require a signal constellation for calculating error statistics and thus are only applicable for CB communication applications. This is not a constraint of the RF-DNA approach presented in Sect. 3.1 which has been successfully used for both communication applications [15,16,18,19,21], and non-communication applications such as discriminating between device components and operational states [6, 17, 22, 23].

A new approach to CB-DNA was first introduced in [20] which included development of a 2D binary signal constellation for unintentional wired Ethernet emissions using features from two binary *composite* constellation point clusters. Nearest Neighbor (NN) and Multiple Discriminant Analysis, Maximum Likelihood (MDA/ML) classifiers were used to assess device discrimination for Cross-Model (manufacturer) Discrimination (CMD) with the MDA/ML classifier out performing NN. Results here extend this earlier work by 1) exploiting discriminating feature information in multiple *conditional* constellation point sub-clusters that form the binary composite clusters, and 2) assessing Like-Model (serial number) Discrimination (LMD) capability as required for the envisioned network device ID authentication process.

3 Experimental Methodology

This work varies from traditional PHY fingerprinting approaches in that it is fingerprinting wired network devices via the unintentional RF emissions given off by the Ethernet cable. The experimental methodology here was adopted from [24] and is summarized briefly for completeness. The emission collection setup included interconnecting two computers using 10BASE-T Ethernet signaling over a category 6 Ethernet cable. A LeCroy WavePro 760Zi-A oscilloscope operating at a sample frequency is f_s=250M Samples/Sec (MSPS) and a high sensitivity Riscure 205HS near-field RF probe were used to collect the unintentional

RF emissions. An in-line baseband filter with bandwidth of W_{BB}=32Mhz was used to limit the collection bandwidth. The Ethernet cable and RF probe were placed in a test fixture to maintain relative cable-to-probe orientation while the Ethernet cards were swapped in and out for collection.

As shown in Table 1 [20], a total of 16 network cards were tested, with four cards each from D-Link (DL), Intel (IN), StarTech (ST), and TRENDnET (TN). The last four MAC address digits show that some devices vary only by a single digit and are likely from same production run. Four unique LAN transformer markings are provided and used to analyze results. The LAN transformer is the last part that the signal goes through prior to reaching the RJ45 output jack [20].

Table 1. Ethernet Cards Used for Emission Collection [20]

Manufacturer	Reference	MAC Address Last Four	LAN Transformer Markings		
D-Link	DL1	D966	Bi-Tek	IM-1178LLF	1247I
	DL2	DA06	Bi-Tek	IM-1178LLF	1247I
	DL3	DA07	Bi-Tek	IM-1178LLF	1247I
	DL4	60E0	Bi-Tek	IM-1178LLF	1247I
TRENDnET	TN1	9B55	Bi-Tek	IM-1178LLF	1247I
	TN2	9334	Bi-Tek	IM-1178LLF	1247I
	TN3	9B54	Bi-Tek	IM-1178LLF	1247I
	TN4	9B56	Bi-Tek	IM-1178LLF	1247I
Intel	IN1	1586	BI	HS00-06037LF	1247
	IN2	1A93	BI	HS00-06037LF	1247
	IN3	1A59	BI	HS00-06037LF	1247
	IN4	1A9E	BI	HS00-06037LF	1247
Star Tech	ST1	32CB	FPE	G24102MK	1250a1
	ST2	32B4	FPE	G24102MK	1250a1
	ST3	96F4	FPE	G24102MK	1320G1
	ST4	3048	FPE	G24102MK	1250a1

3.1 RF-DNA Fingerprinting

The RF-DNA fingerprinting approach has been most widely used for *intentional* signal responses of wireless devices [15, 16, 18, 19]. For this work, the RF-DNA approach adopts the technique introduced in [24] for collecting *unintentional* RF emissions from Ethernet cables and producing RF-DNA fingerprints on a burst-by-burst basis. Useful RF-DNA has been historically extracted from invariant signal amble regions [6, 15, 16] and thus the 10BASE-T preamble response was targeted here for initial assessment. RF-DNA features can be extracted from various ROI responses, a few of which include Time Domain (TD) [16], Spectral Domain (SD) [18], Fourier Transform (FT) [18], and Gabor Transform (GT) [15].

Instantaneous amplitude $\{a(k)\}$, phase $\{\phi(k)\}$, and frequency $\{f(k)\}$ are TD sequences used for RF-DNA fingerprint generation using the preamble as

the ROI; k denotes discrete time samples. Composite RF-DNA fingerprints are generated by 1) centering (mean removal) and normalizing $\{a(k)\}$, $\{\phi(k)\}$, and $\{f(k)\}$, 2) dividing each TD sequence into N_R equal length subregions as illustrated in Fig. 1, 3) calculating three statistical features of variance (σ^2), skewness (γ), and kurtosis (κ) for *each* TD sequence to form *Regional Fingerprint* $F_{R_i}^{a,\phi,f}$ as in (1) for i=1, 2, ..., N_R, and 4) concatenating $F_{R_i}^{a,\phi,f}$ to form the final 1×(3×N_R) *Composite RF-DNA Fingerprint* F_C^{RF} as in (2) [17]. Statistical features across entire ROI response are commonly included as well, hence the regional indexing in (2) to N_R+1 total elements.

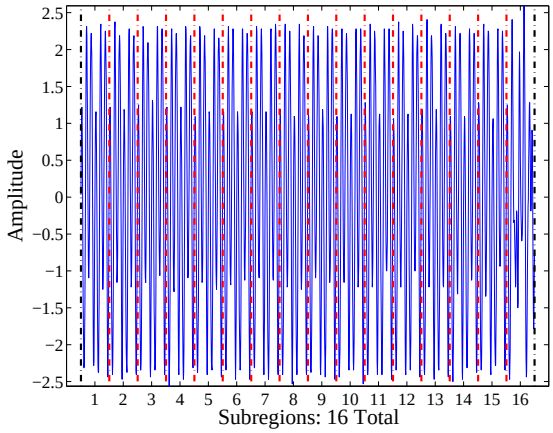

Fig. 1. Representative 10BASE-T Preamble Time Domain Amplitude Response Used for Fingerprint Generation. The 6.4 μs Preamble is Divided into N_R=16 Subregions.

The total number of RF-DNA features in (2) is a function of N_R, TD responses, and statistics. Varying N_R provides a means to investigate performance for various feature vector sizes. Fingerprints were generated over the preamble ROI using three TD responses ($\{a(k)\}$, $\{\phi(k)\}$, $\{f(k)\}$), three statistics (σ^2, γ, κ) per response, for N_R=16, 31, 80 with (2) and produced RF-DNA fingerprints having N_{Feat}=144, 279, and 720 total features, respectively.

$$F_{R_i}^{a,\phi,f} = [\sigma_{R_i}^2, \gamma_{R_i}, \kappa_{R_i}]_{1\times 3} \tag{1}$$

$$F_C^{RF} = [F_{R_1}^{RF} : F_{R_2}^{RF} : F_{R_3}^{RF} : \ldots : F_{R_{N_R+1}}^{RF}]_{1\times(3\times N_R)} \tag{2}$$

3.2 CB-DNA Fingerprinting

As with RF-DNA fingerprinting approach, the majority of CB fingerprinting works utilize intentional RF emissions from wireless devices with unique features derived from modulation errors in the constellation space, i.e., differences

(error) between received projected symbol points and ideal transmitted constellation points [1–3,5]. The CB-DNA approach adopted here differs from previous approaches by utilizing statistical features from *unconditional* and *conditional* projected symbol clusters (not modulation errors) in a 2D constellation space.

The CB-DNA fingerprinting process used here was adopted from [20] and is summarized here for completeness. CB-DNA fingerprints were generated from a single burst with example constellations being illustrated in Fig. 2 for the four card manufacturers with blue circles and black squares clusters representing Binary 0 and Binary 1, respectively. This research expands on [20] by utilizing for the first time *conditional* subclusters illustrated in Fig. 3 for card manufacturer StarTech. The *conditional* subclusters are based not only on the current demodulated bit but the proceeding and succeeding bit as well. The eight distinct *conditional* sub-clusters correspond to the eight possible bit combinations that can precede and succeed the bit being estimated i.e., bit combinations of [0 **X** 0], [0 **X** 1], [1 **X** 0], and [1 **X** 1], where **X** denotes the bit being estimated.

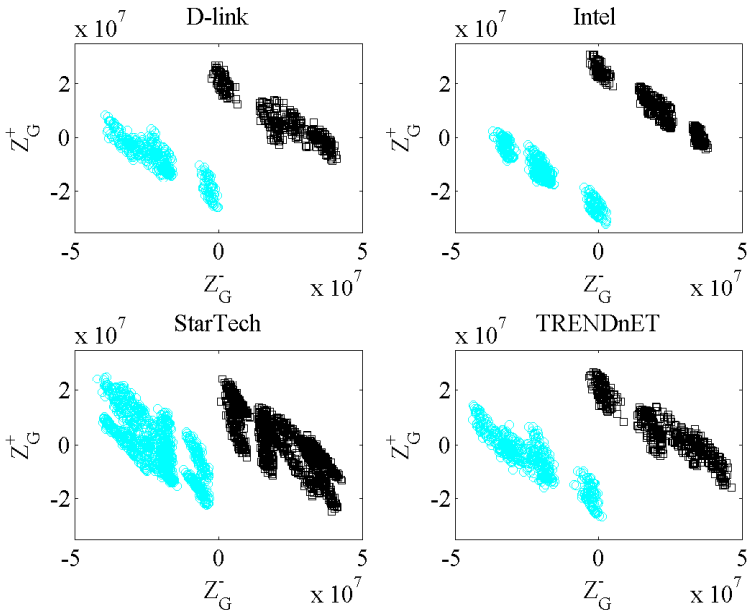

Fig. 2. Device Constellations Consisting of Approximately 1,400 Symbols (1 Burst) for Each of the Four Card Manufacturers. Each *Composite Cluster* Represents a Binary 0 (Blue Circle) or Binary 1 (Black Square).

CB-DNA fingerprint generation begins by dividing constellation points into their respective *unconditional* and *conditional* cluster regions for a total of N_{CR}=2+8=10. Statistical CB-DNA features are then calculated for each cluster region using the mean (μ), variance (σ^2), skewness (γ), and kurtosis (κ)

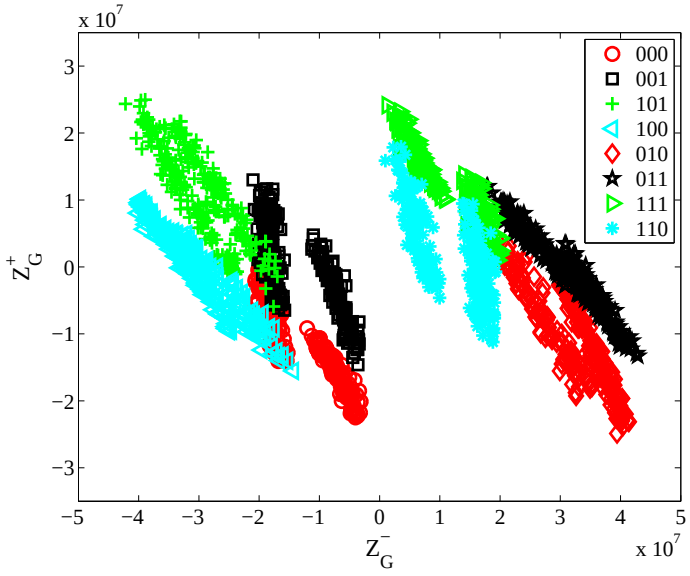

Fig. 3. 2D Binary Constellation for ST1 showing each *Composite Cluster* Comprised of Four Distinct *Sub-Clusters* Corresponding to the Four Possible Combinations of Bits Preceding and Succeeding the Bit Being Estimated.

along the Z_G^- and Z_G^+ dimensions shown in Fig. 3. Joint statistics in both the Z_G^- and Z_G^+ direction are also considered and include covariance (*cov*), coskewness ($\beta_{1\times2}$), and cokurtosis ($\delta_{1\times3}$) which provide an extra six features per region. The resultant statistics form a *Regional Cluster Fingerprint* $F_{R_i}^{CB}$ given by (3), where the superscripted $-/+$ sign denotes constellation dimension and i=1, 2, ..., N_{CR}. The final *Composite CB-DNA Fingerprint* F_C^{CB} is of dimension $1 \times (14 \times N_{CR})$ and constructed by concatenating $F_{R_i}^{CB}$ from (3) as shown in (4) [20]. The total number of CB-DNA features in (4) is a function of N_{CR}, statistics, and dimensions i.e., Z_G^- and Z_G^+. Varying N_{CR} provides a means to investigate performance for various feature vector sizes. Fingerprints were generated using N_{CR}=2, 8, and 10 (μ, σ^2, γ, κ, *cov*, $\beta_{1\times2}$, $\delta_{1\times3}$) with 4 statistics from each of the Z_G^- and Z_G^+ dimensions and 6 joint statistics producing CB-DNA fingerprints having N_{Feat}=28, 112, and 140 total features, respectively.

$$F_{R_i}^{CB} = \left[\mu_{R_i}^-, \mu_{R_i}^+, \sigma_{R_i}^{2-}, \sigma_{R_i}^{2+}, \gamma_{XR_i}^-, \gamma_{R_i}^+, \kappa_{R_i}^-, \kappa_{R_i}^+, cov, \beta_{1\times2}, \delta_{1\times3},\right]_{1\times14} \tag{3}$$

$$F_C^{CB} = \left[F_{R_1}^{CB} : F_{R_2}^{CB} : F_{R_3}^{CB} : \ldots : F_{R_{N_{CR}}}^{CB}\right]_{1\times(14\times N_{CR})} \tag{4}$$

3.3 Device Discrimination

The effect of varying SNR on discrimination performance was assessed to characterize the effect of varying channel conditions. This was done by adding in-

dependent like-filtered Additive White Gaussian Noise (AWGN) N_{Nz} realizations to each experimentally collected emission to achieve the desired SNR for Monte Carlo simulation. Given an average collected SNR=30.0 dB, device discriminability was assessed for simulated $SNR \in$ [12 32] dB in 2 dB steps. For Monte Carlo simulation results in Sect. 4, a total of N_{Nz}=6 independent AWGN realizations were generated, filtered, power-scaled and added to the collected signal responses to generated signals at the desired SNR. Given N_{Nz} =6 AWGN realizations and N_S=1000 collected signal responses per card, a total of N_F=$N_S \times N_{Nz}$=6000 independent fingerprints per card were available for discrimination assessment.

Consistent with prior related work [6, 15, 16], device discriminability was assessed using a MDA/ML classification process. MDA/ML processing was implemented for N_C=4 and 16 classes using an identical number of *Training* (N_{Tng}) and *Testing* (N_{Tst}) fingerprints for each class. A total of N_F=24,000 (CMD) and N_F=6,000 (LMD) fingerprints were generated for each N_C per Sect. 3.1 and Sect. 3.2 for RF-DNA and CB-DNA methods, respectively. MDA/ML *training* was completed for each N_C using N_{Tng}=$N_F/2$ fingerprints and K-fold cross-validation with K=5 to improve MDA/ML reliability. This involves: 1) dividing the training fingerprints into K equal size disjoint blocks of $N_{Tng}/5$ fingerprints, 2) holding out one block and training on K-1 blocks to produce projection matrix $\mathbf{W}$, and 3) using the holdout block and $\mathbf{W}$ for validation [25]. The $\mathbf{W}$ from the best training iteration is output and used for subsequent MDA/ML *testing* assessment. The process is repeated to generate an SNR-dependent $\mathbf{W}(SNR)$ for each analysis SNR.

4 Discrimination Results

The MDA/ML classification results are presented for CMD (manufacturer) and LMD (serial number) performance using the 16 devices in Table 1. Device fingerprint generation occurs using identical burst-by-burst emissions per methods in Sect. 3.1 and Sect. 3.2, with RF-DNA using only the burst preamble and CB-DNA using the entire burst to include preamble. A total of 1000 bursts are processed from each device with three AWGN realizations added to each burst to create 3000 fingerprints per device for classification. Discrimination results are based on two classification models created per Sect. 3.3. The CMD results are based on N_{Tst}=12,000 testing fingerprints and LMD results are based on N_{Tst}=3,000 fingerprints. An arbitrary performance benchmark of %C=90% correct classification is used for comparative assessment with summary analysis based on CI=95% binomial confidence intervals. Given the large number of independent trials for all results in Sect. 4, the resultant CI=95% confidence intervals are less than the vertical extent of data markers in Fig. 4 through Fig. 6 and therefore omitted for visual clarity.

Fig. 4 shows average RF-DNA results for CMD and LMD. The %C=90% benchmark is achieved for CMD with all three N_R values at $SNR \geq$21 dB, with N_R=80 performance starting out with %C=92% at SNR=12 dB and the

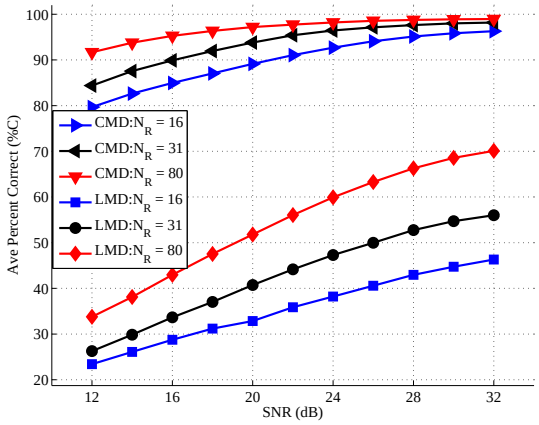

Fig. 4. RF-DNA Fingerprinting Averages for *Cross-Model Discrimination* (CMD) Using N_C=4 Classes and *Like-Model Discrimination* (LMD) Using N_C=16 Classes with N_R=16, 31 and 80 Sub-Regions.

other N_R=16 and N_R=31 cases requiring an additional 6.0 dB and 10.0 dB gain in SNR, respectively, to achieve $\%C$=92%. At SNR ≈26 dB the markers for N_R=80 and N_R=31 begin to overlap suggesting those two MDA/ML models yield statistically equivalent performance at SNR=26.0 dB and higher with N_R=16 %C being slightly lower. The LMD results for RF-DNA in Fig. 4 never reach the $\%C$=90% benchmark. However, the N_R=80 case outperforms the others by approximately 15% and 20% at SNR=30 dB. Thus, the RF-DNA model for N_R=80 was chosen for comparison with the CB-DNA model.

Fig. 5 shows CMD and LMD results for CB-DNA fingerprinting while varying the use of composite clusters and sub-clusters. The CMD and LMD results using N_{CR}=2 are about 5% and 25% respectively worse in correct classification with respect to the N_{CR}=10 cases. For CMD the N_{CR}=10 model achieves 96% correct classification on average at 12 dB where the N_{CR}=2 model peaks out at 94% at 32 db, which shows that the N_{CR}=10 model is superior. The results for CMD with N_{CR}=8 are similar to N_{CR}=10. LMD results for N_{CR}=8 are constantly a few percentage points lower than N_{CR}=10 and requires an additional 4 dB gain to achieve $\%C$=90 over N_{CR}=10 at 22 dB. LMD increases the complexity of the classification and reaches an average of %C=90% across all 16 device for N_{CR}=10 with average collected SNR=22.0 dB.

$\%C$ classification results for RF-DNA and CB-DNA Fingerprinting are provided in Fig. 6 for CMD and LMD. The CMD comparison shows that CB-DNA reaches $\%C$=96% at SNR=12 dB while RF-DNA reaches $\%C$=96% at $SNR \approx 16$ dB (approximately 6.0 dB higher). The LMD comparison shows that CB-DNA consistently out performs RF-DNA by at least 24% at all SNR levels.

The results in Fig. 6 enable direct comparison of RF-DNA and CB-DNA Fingerprinting however, average $\%C$ performance hides individual class interactions.

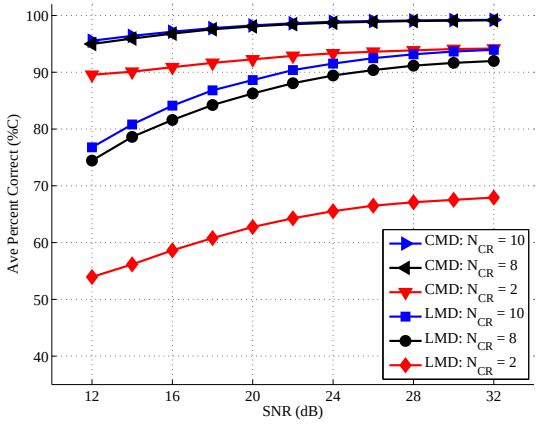

Fig. 5. CB-DNA Fingerprinting Averages for *Cross Model Discrimination* (CMD with N_C=4) and *Like Model Discrimination* (LMD with N_C=16) Using Composite (N_{CR}=2), Sub-Cluster (N_{CR}=8), and Combined (N_{CR}=10) Constellation Statistics.

Thus, MDA/ML confusion matrix results for SNR=24.0 dB are introduced to highlight cross-class misclassification for CMD (Table 2) and LMD (Table 3); matrix rows represent *input class* and matrix columns represent *called class*. The table entries are presented as $\%C$ CB-DNA / $\%C$ RF-DNA with bold entries denoting best or equivalent performance.

The CMD confusion matrix in Table 2 is nearly symmetric with *all* misclassification occurring between DL and TN devices. This is attributable to DL and TN devices using identical LAN transformers as indicated in Table 1. The diagonal entries show that CMD performance, for CB-DNA is better than or equivalent to RF-DNA. The resultant CMD averages for CB-DNA ($\%C$=98.9%) and RF-DNA ($\%C$=98.21%) are pursuant with Fig. 6.

Table 2. CMD confusion matrix for CB-DNA and RF-DNA Fingerprinting at SNR=24 dB and 12,000 trials per class. Entries presented as % CB-DNA / % RF-DNA with bold entries denoting best or equivalent performance.

	DL	**IN**	**ST**	**TN**
DL	**98.10** / 96.05	0.0 / 0.0	0.0 / 0.0	1.90 / 3.95
IN	0.0 / 0.0	**100.00** / **100.00**	0.0 / 0.0	0.0 / 0.0
ST	0.0 / 0.0	0.0 / 0.0	**100.00** / **100.00**	0.0 / 0.0
TN	2.49 / 3.20	0.0 / 0.0	0.0 / 0.0	**97.51** / **96.80**

The LMD confusion matrix results in Table 3 summarize misclassification of the complete 16-by-16 confusion matrix. Results are presented as individual manufacturer confusion matrices with "Other" entries representing all misclassi-

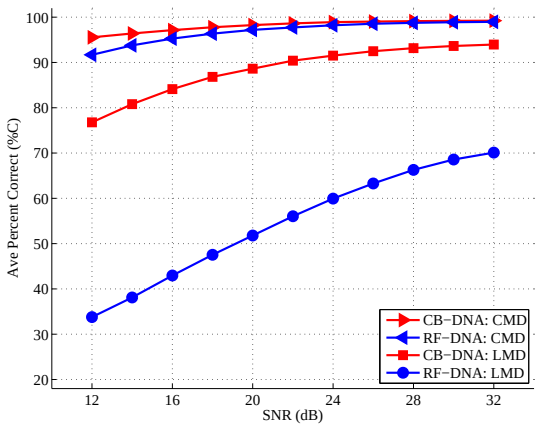

Fig. 6. RF-DNA vs. CB-DNA Fingerprinting Averages for *CMD* (Manufacturer Discrimination) Using N_C=4 Classes with (N_R=16) Sub-Regions and *LMD* (Serial Number Discrimination) Using N_C=16 Classes with (N_{CR}=10) Cluster Regions.

Table 3. LMD confusion matrix for CB-DNA and RF-DNA Fingerprinting at SNR=24.0 dB and 3,000 trials per class, highlighting errors within manufacturing groups. The "Other" column represents all other manufacturers. Entries presented as % CB-DNA / % RF-DNA with bold entries denoting best or equivalent performance.

	DL1	**DL2**	**DL3**	**DL4**	**Other**
DL1	**86.27** / 43.10	0.0 / 9.43	13.73 / 20.47	0.0 / 25.77	0.0 / 1.23
DL2	0.03 / 10.83	**97.57** / 72.80	0.80 / 6.50	0.63 / 3.50	0.97 / 6.37
DL3	15.77 / 22.47	0.03 /5.40	**83.93** / 57.77	0.27 / 12.40	0.0 / 1.96
DL4	0.0 / 29.30	0.37 / 7.13	0.10 / 17.07	**97.20** / 45.67	2.33 / 0.83
	TN1	**TN2**	**TN3**	**TN4**	**Other**
TN1	**97.90** / 43.37	1.10 / 16.44	0.03 / 8.13	0.67 / 30.13	0.30 / 1.93
TN2	1.36 / 23.90	**89.57** / 43.60	1.77 / 6.80	3.07 / 19.13	4.23 / 6.57
TN3	0.0 / 4.43	2.83 / 4.60	**84.87** / 85.07	11.90 / 5.50	0.40 / 0.40
TN4	1.60 / 32.60	5.20 / 16.60	6.76 / 8.13	**84.27** / 39.84	2.17 / 2.83
	IN1	**IN2**	**IN3**	**IN4**	**Other**
IN1	**91.67** / 78.43	5.63 / 9.00	2.40 / 11.40	0.30 / 1.17	0.0 / 0.0
IN2	4.50 / 9.73	**94.30** / 60.67	1.20 / 14.60	0.00 / 15.00	0.0 / 0.0
IN3	1.46 / 11.00	0.67 / 17.10	**97.67** / 61.77	0.20 / 10.13	0.0 / 0.0
IN4	0.0 / 4.07	0.0 / 19.13	0.0 / 12.07	**100.00** / 64.73	0.0 / 0.0
	ST1	**ST2**	**ST3**	**ST4**	**Other**
ST1	**90.53** / 67.63	0.70 / 1.20	4.40 / 13.27	4.37 / 17.90	0.0 / 0.0
ST2	0.43 / 0.83	**97.47** / 77.80	0.77 / 13.40	1.33 / 7.97	0.0 / 0.0
ST3	4.46 / 10.93	1.07 / 14.30	**83.57** / 60.60	10.90 / 14.17	0.0 / 0.0
ST4	3.63 / 19.33	1.30 / 6.93	7.50 / 17.50	**87.57** / 56.24	0.0 / 0.0

fications outside the manufacturing group. The results here are consistent with prior CMD results in Table 2, with 1) the IN and ST devices are never misclassified as another manufacturer, and 2) nearly 100% of the DL "Other" misclassifications being TN devices, and vice versa–this confusion is again attributed to DL and TN devices using identical LAN transformers as indicated in Table 1. Most notably in Table 3 are bold diagonal entries showing that CB-DNA outperformed RF-DNA performance for *all* devices.

5 Summary and Conclusions

A PHY augmentation to MAC-based authentication is addressed using PHY-based Distinct Native Attribute (DNA) features to form device fingerprints. Specifically, a previous Radio Frequency (RF-DNA) fingerprinting approach and new Constellation Based (CB-DNA) fingerprinting approach that exploits 2D constellation statistics are considered. The two methods are compared using fingerprints generated from the same set of unintentional 10BASE-T Ethernet cable emissions. Prior to this preliminary investigation it was hypothesized that CB-DNA would outperform RF-DNA. Considerable differences in the amount of burst information being exploited was the basis for this conjecture, i.e., RF-DNA fingerprinting only exploits a fraction of the Ethernet burst (64 preamble symbols) while CB-DNA exploits the entire Ethernets burst (average of 1,400 symbols here).

When comparing RF-DNA results here to previous related work [15–17], it is noted that Cross-Model Discrimination (CMD) results are consistent but Like-Model Discrimination (LMD) results are poorer. One reason for this is more stringent signaling characteristics of the Ethernet standards as well as the devices here sharing similar LAN transformer markings.

As measured by average percentage of correct classification (%C), the final RF-DNA vs. CB-DNA outcome shows that CB-DNA outperforms RF-DNA for the 16 devices considered. For CMD there was only a marginal difference at SNR=24 dB with CB-DNA at %C=98.9% and RF-DNA at %C=98.21%. Of particular note for CMD is that 100% of the misclassification error occurred between DL and TN devices which use the same LAN transformer. For LMD there was considerable improvement at SNR=24 dB, with CB-DNA at %C=91.5% and RF-DNA at %C=59.9%. LMD is generally more challenging than CMD and results show that both approaches suffer when classifying LMD. However, CB-DNA performance remained above the 90% threshold and only suffered a 6.2% degradation in %C while RF-DNA dropped by more than 30%.

From a device authentication and network security perspective, LMD performance is most important. Results here show that CB-DNA outperformed RF-DNA by a considerable margin. LMD results at the collected SNR=30.0 dB include like model %C=94% for CB-DNA and only %C=69% for RF-DNA.

These CB-DNA results are encouraging and work continues to improve performance. This includes investigating alternatives such as the Generalized Relevance Learning Vector Quantized-Improved (GRLVQI) classifier which provides

a direct indication of feature relevance on classifier decision [21, 26]. Feature relevance enables dimensional reduction analysis, which in-turn reduces processing complexity and enhances real-world applicability. Furthermore, the use of CB-DNA for device verification and rogue detection and rejection remains under investigation as well.

References

1. Danev, B., Zanetti, D., Capkun, S.: On Physical-Layer Identification of Wireless Devices. In: ACM Computing Surveys (CSUR), vol. 45, no. 1, p. 6 (2012)
2. Huang, Y., Zheng, H., Radio Frequency Fingerprinting Based on the Constellation Errors. In: Communications (APCC), 2012 18th Asia-Pacific Conf on. IEEE, pp. 900-905 (2012)
3. Brik, V., Banerjee, S., Gruteser, M., Oh, S.: Wireless Device Identification with Radiometric Signatures. In: Proc of the 14th ACM Intl Conf on Mobile computing and networking. ACM, pp. 116-127 (2008)
4. Danev, B., Luecken, H., Capkun, S., El Defrawy, K.: Attacks on Physical-Layer Identification. In: Proc of the third ACM Conf on Wireless network security. ACM, pp. 89-98 (2010)
5. Edman, M., Yener, B.: Active Attacks Against Modulation-Based Radiometric Identification. Technical report 0902, Rensselaer Institute of Technology (2009)
6. Cobb, W. E., Laspe, E. D., Baldwin, R. O., Temple, M. A., Kim, Y. C.: Intrinsic Physical-Layer Authentication of Integrated Circuits. Information Forensics and Security, IEEE Trans on, vol. 7, no. 1, pp. 14-24, (2012)
7. Desmond, LCC., Cho, CY., Tan, CP., Lee RS.: Identifying unique devices through wireless fingerprinting. Proceedings of the first ACM conference on Wireless network security. ACM, (2008)
8. Kohno, T., Broido, A., Claffy, K. C.: Remote physical device fingerprinting. IEEE Transactions on Dependable and Secure Computing 2(2), pp. 93–108 (2005)
9. Franklin, J., McCoy, D., Tabriz, P., Neagoe, V., Randwyk, J. V., Sicker, D.: Passive Data Link Layer 802.11 Wireless Device Driver Fingerprinting. In: Usenix Security (Vol. 6) (2006).
10. Gao, K., Corbett, C., Beyah, R. A.: A Passive Approach to Wireless Device Fingerprinting. In: Proc. of IEEE/IFIP DSN, pp. 383–392 (2010)
11. Uluagac, A., Radhakrishnan, S., Corbett, C., Baca, A., Beyah, R.: A Passive Technique for Fingerprinting Wireless Devices with Wired-side Observations. In: Proceedings of the IEEE Conference on Communications and Network Security (CNS,) pp. 305–313 (2013)
12. Francois, J., Abdelnurt, H., State, R., Festort, O.: Ptf: Passive Temporal Fingerprinting. In: Proc. of IFIP/IEEE International Symposium on Integrated Network Management, pp. 289–296 (2011)
13. Gubbi, J., Buyya, R., Marusic, S., Palaniswami, M.: Internet of Things (IoT): A Vision, Architectural Elements, and Future Directions. In: Future Generation Computer Systems, vol. 29, no. 7, pp. 1645-1660, (2013)
14. Zhou, L., Chao, H. C.: Multimedia Traffic Security Architecture for the Internet of Things. In: Network, IEEE, vol. 25, no. 3, pp. 35-40, (2011)
15. Reising, D. R., Temple, M. A., Oxley, M. E.: Gabor-Based RF-DNA Fingerprinting for Classifying 802.16e WiMAX Mobile Subscribers. In: Computing, Networking and Communications (ICNC), 2012 Intl Conf on. IEEE, pp. 7-13 (2012)

16. Ramsey, B. W., Temple, M. A., Mullins, B. E.: PHY Foundation for Multi-Factor ZigBee Node Authentication. In: Global Communications Conf (GLOBECOM), 2012 IEEE, pp. 795-800 (2012)
17. Cobb, W. E, Garcia, E. W., Temple, M. A., Baldwin, R. O., Kim, Y. C.: Physical Layer Identification of Embedded Devices using RFDNA Fingerprinting. In: MILITARY COMMUNICATIONS Conf, MILCOM 2010, pp. 2168-2173 (2010)
18. Williams, M. D., Munns, S., Temple, M. A., Mendenhall, M. J.: RF-DNA Fingerprinting for Airport WiMax Communications Security. In: Network and System Security (NSS), 2010 4th Intl Conf on, pp. 32-39 (2010)
19. Williams, M. D., Temple, M. A., Reising, D. R.: Augmenting Bit- Level Network Security Using Physical Layer RF-DNA Fingerprinting. In: Global Telecommunications Conf (GLOBECOM 2010), 2010 IEEE, pp. 1-6 (2010)
20. Carbino, T. J., Temple, M. A., Bihl, T. : Ethernet Card Discrimination Using Unintentional Cable Emissions and Constellation-Based Fingerprints. In: Computing, Networking and Communications (IWCNC), 2015 Intl Workshop on, Accepted, to appear Feb (2015)
21. Reising, D. R.: Exploitation of RF-DNA for Device Classification and Verification Using GRLVQI Processing. Technical report DTIC Doc (2012)
22. Stone, S. J., Temple, M. A., Baldwin, R. O.: RF-Based PLC IC Design Verification. In: 2012 DMSMS and Stand Conf. (DMSMS12) (2012)
23. Wright, B. C.: PLC Hardware Discrimination using RF-DNA Fingerprinting. Technical Report DTIC Document (2014)
24. Carbino, T. J., Baldwin, R. O.: Side Channel Analysis of Ethernet Network Cable Emissions. In: 9th Intl Conf on Cyber Warfare and Security, ICCWS (2014)
25. Duda, R. O., Hart, P. E., Stork, D. G.: Pattern Classification. John Wiley and Sons (2012)
26. Mendenhall, M. J., Merényi, E.: Relevance-Based Feature Extraction for Hyperspectral Images. Neural Networks, IEEE Trans on, vol. 19, no. 4, pp. 658-672 (2008)