This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE Globecom 2010 proceedings.

Augmenting Bit-Level Network Security Using Physical Layer RF-DNA Fingerprinting

McKay D. Williams, Michael A. Temple, and Donald R. Reising

Department of Electrical and Computer Engineering
Air Force Institute of Technology
Wright-Patterson AFB, OH 45433 USA
Email: michael.temple@afit.edu

***Abstract*—Successful "cracking" of bit-level security compromises network integrity and physical layer augmentation is being investigated to improve overall security. Intra-cellular security is addressed here using device-specific RF "Distinct Native Attribute" (RF-DNA) fingerprints in a localized regional air monitor, with targeted applications including cellular networks such as the Global System for Mobile (GSM) Communications and last mile Worldwide Interoperability for Microwave Access (WiMAX) systems. Previous work demonstrated GSM inter-manufacturer classification (manufacturer discrimination) using RF-DNA fingerprinting and achieved accuracies of 92% at $SNR = 6$ dB. These results are extended here for intra-manufacturer classification (serial number discrimination). Historically, intra-manufacturer discrimination has posed the greatest challenge and RF-DNA fingerprinting has been effective with both Orthogonal Frequency Division Multiplexed (OFDM) and Direct Sequence Spread Spectrum (DSSS) network signals. Intra-manufacturer GSM results are provided here based on identical signal collection, fingerprint generation, and MDA/ML classification processes used for previous inter-manufacturer assessment. When comparing performance, the trend for GSM intra-manufacturer classification is consistent with previous work for other network-based signals and device classification is much more challenging. For classification accuracies of 80% or better, intra-manufacturer fingerprinting requires an increase of $20 - 25$ dB in SNR to achieve inter-manufacturer performance.**

I. Introduction

The vulnerability of wireless RF networks will remain as long as physics continues to support "free-space" propagation, i.e., everyone within range is assured "free" access. Furthermore, the longer a system remains employed the greater the opportunity for unauthorized activity given that system operational details become more available, and perhaps more significantly, the hardware and software used by unauthorized opportunists becomes increasingly capable. This is clearly evident with Global System for Mobile Communications (GSM) which employs cold war era 64-bit encryption that has been successfully "cracked" [1] and which remains under attack to this day [2], [3]. Thus, the exploitability of bit-level network security mechanisms remains a concern and it is not unreasonable to assume that emerging systems, e.g., last mile Worldwide Interoperability for Microwave Access (WiMAX), will be the next target of opportunistic "hackers."

The need for improved authentication and security measures in cellular-based network architectures (GSM, WiMAX, etc.) remains and Physical (PHY) layer security through regional air monitoring of RF signal characteristics is one viable alternative given that these characteristics are inherently difficult to mimic. The concept for using an "air monitor" to exploit PHY layer attributes and provide added security is not new [4], [5]. The work here builds upon these previous concepts and addresses intra-cellular air monitoring based on what are called RF "Distinct Native Attribute" (RF-DNA) fingerprints. While having air monitoring capability within each wireless device may be advantageous, physical size constraints preclude this option and regional intra-cellular monitoring at base stations is envisioned as the near-term employment alternative.

The *turn-on transient* region of burst-like communication signals possess distinct RF features that have been successfully exploited for RF fingerprinting [6]–[11]. In addition, reliable performance has been achieved with *near-transient* signal regions as well [5], [12]–[14] and the motivation for using near-transient features justified–the near-transient region includes the typical turn-on transient response plus a subsequent portion of the signal. Cellular-based signals generally contain pre-defined bit-level sequences and/or signal regions that contain essential information for maintaining communications. Of specific interest to this work are near-transient preamble and/or non-transient midamble responses when present. When based on pre-defined bit sequences or standards, and when identical for all users within a given cell, the signal responses will contain distinct "coloration" that is attributable to device RF components such as mixers, filters, amplifiers, etc., that are unique to a given device–the device's RF-DNA.

The methodology used here is consistent with [11]–[13], [15], [16] and a brief description of key processes is provided in Sect. II. Device classification results are presented in Sect. III based on experimentally collected GSM signals. Results for *inter-manufacturer* classification performance is first assessed and found to be consistent with results presented in [16]. These results are presented in Sect. III-A and based on using devices from four different manufacturers: Motorola, Nokia, Samsung, and Sony Ericsson. Serial number discrimination is then assessed in Sect. III-B for each of the manufacturers using *intra-manufacturer* classification with four like-model devices. While unintentional modulation

effects exist and are exploitable in both cases, the observed cross-device variation in the intra-manufacturer case is clearly reduced relative to the inter-manufacturer case. Thus, the RF-DNA for devices manufactured by a given manufacturer and having near-consecutive serial numbers is most "identical" and presents the greater classification challenge.

This challenge must be addressed for operational implementation as cells under the control of a given base station will generally contain like-model devices from various manufacturers. As such, work continues with intra-cellular GSM signals and has recently expanded to include intra-cellular OFDM-based WiMAX signals [17]. The favorable results of work presented in this paper and prior success with OFDM-based WiFi signals [5] suggest that RF-DNA fingerprinting should be well-suited for WiMAX applications.

II. Background

A. Statistical Fingerprint Generation

The statistical RF-DNA fingerprint ($\mathbf{F}$) for a signal is based on its amplitude (a), phase (ϕ) and/or frequency (f) characteristics. More specifically, the sequences $\{\bar{a}_c(n)\}$, $\{\bar{\phi}_c(n)\}$ and/or $\{\bar{f}_c(n)\}$ are generated from complex samples where the subscripted c denotes centering (mean removal) and the over bar denotes normalization (division by maximum value) [12], [13], [15], [16]. Using the selected characteristic sequence(s), the statistical fingerprint features are generated as standard deviation (σ), variance (σ^2), skewness (γ), and kurtosis (k) within specific signal regions. The *regional fingerprint markers* are generated by: 1) dividing each characteristic sequence into N_R contiguous, equal length subsequences, 2) calculating the four metrics for each subsequence and for all sequence elements ($N_R + 1$ total regions), and 3) arranging the metrics in a vector as

$$F_{R_i} = [\sigma_{R_i} \ \sigma^2_{R_i} \ \gamma_{R_i} \ k_{R_i}]_{1\times 4} \ , \tag{1}$$

where $i = 1, 2, ..., N_R + 1$. The marker vectors from (1) are concatenated to form the *composite characteristic vector* for each characteristic and is given by

$$\mathbf{F}^C = \left[F_{R_1} \vdots F_{R_2} \vdots F_{R_3} \ \ldots \ F_{R_{N_R+1}} \right]_{1\times 4(N_R+1)} . \tag{2}$$

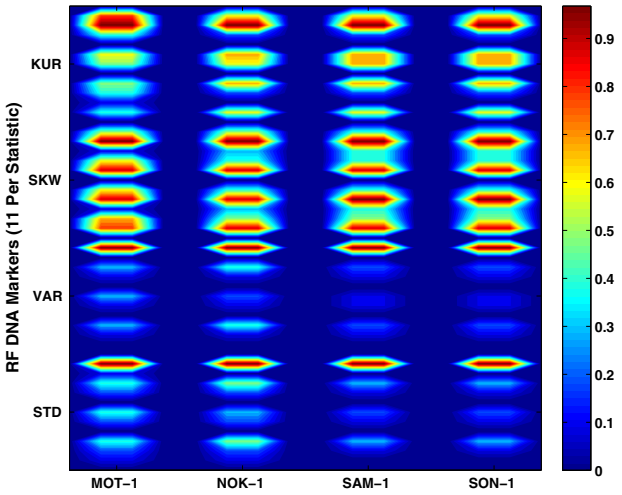
If only one signal characteristic is used, the expression in (2) represents the final fingerprint used for classification. When all signal characteristics are used, the final *RF-DNA fingerprint* is generated by concatenating vectors from (2) according to

$$\mathbf{F} = \left[\mathbf{F}^a \vdots \mathbf{F}^\phi \vdots \mathbf{F}^f \right]_{1\times 4(N_R+1)\times 3} . \tag{3}$$

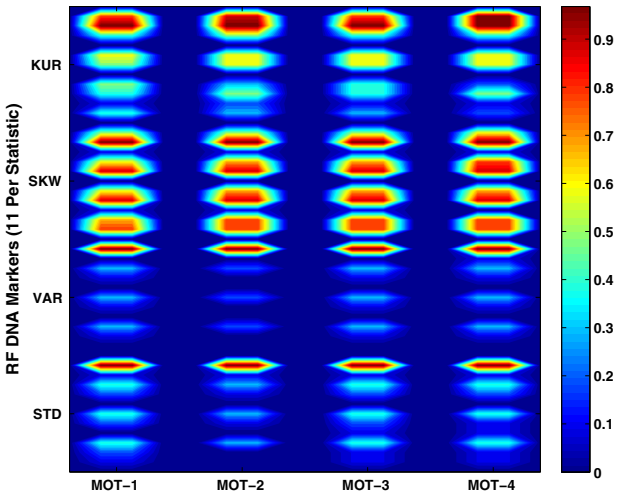
Previous GSM work in [15] showed that $N_R = 10$ subregions provides best case, or near best case, performance when using *phase-only* fingerprints–marginal gains in classification performance occur if amplitude and frequency characteristics are included. Thus, all results in this paper are for *phase-only* fingerprinting using $N_R = 10$ subregions. Accounting for $N_R + 1 = 11$ total signal regions, the resultant $\mathbf{F}$ in (3) contains a total of (4 Statistics)×(11 Regions) = 44 elements.

The RF-DNA fingerprints in Fig. 1 are provided to help the reader visualize statistical fingerprint behavior. These were generated by averaging 400 *phase-only* fingerprints extracted from midamble signal responses at $SNR = 20$ dB. While not actually used by the MDA/ML process as depicted here, the response behavior (features look similar, different, etc.) is generally consistent with resultant MDA/ML performance. The *inter-manufacturer* fingerprint behavior in Fig. 1(a) is consistent with results in [16] and illustrates variation across manufacturer, while the *intra-manufacturer* responses in Fig. 1(b) illustrate variation across serial number of Motorola devices.

The fingerprint representations in Fig. 1 were generated by grouping statistical feature elements from (3) to form the DNA



(a) *Inter-Manufacturer*



(b) *Intra-Manufacturer*

Fig. 1. Average RF-DNA fingerprints for four GSM cell phones using *Midamble* signal responses at $SNR = 20$ dB.

"markers." Within each marker, the statistics were scaled, compressed and/or expanded to span values between [0, 1], and subsequently quantized to a desired number of discrete levels. The quantized markers were vertically stacked to create an "RF-DNA Fingerprint" and presented in an electrophoresis-like plot. These particular reference fingerprints highlight the greater challenge posed by intra-manufacturer classification, i.e., cross-device variability in fingerprint features in Fig. 1(b) is visually lower and thus poses a greater challenge for MDA/ML intra-manufacturer classification.

B. MDA/ML Device Classification

The collection of statistical RF fingerprints, generated using (3) for each collected signal from each device, were input into a Multiple Discriminant Analysis/Maximum Likelihood (MDA/ML) process for device classification. MDA is an extension to Fisher's Linear Discriminant (FLD) when more than two classes (devices) are being considered. MDA reduces the dimensionality of higher-dimensional input data by projecting it into a lower-dimensional space with the goal of maximizing inter-class separation while reducing intra-class spread [18].

For the three class problem, MDA projects the multi-dimensional input data (44-dimensional *phase-only* fingerprints) into a 2-dimensional space. Device classification is performed using a ML classifier as derived from Bayesian Decision Theory, with the multi-dimensional input data classified as being affiliated with one of three possible classes. A Bayesian decision is based on known prior probabilities, probability densities, and relevant costs associated with making a decision. The decision process relies on an accurate representation of the class distribution and its parameters in order to define the likelihood. A sample is evaluated in each class likelihood and the sample is assigned the class label of the class likelihood yielding the maximum response. For ML classification, the prior probabilities are assumed to be equal and the costs uniform. The MDA/ML process was implemented using *K-fold cross-validation* with $K = 5$ to improve the reliability of classification results. While the required value of *K* can be data dependent, empirical testing with the collected data used here confirmed that $K = 5$ was sufficient to ensure reliability. This finding was consistent with common practice that suggests values of $K = 5$ and $K = 10$ are appropriate [19].

III. Device Classification Results

Inter-manufacturer and intra-manufacturer classification results are presented in Sec. III-A and Sec. III-B, respectively. All results were generated using identical collection and post-collection processing parameters, including: f_s = 23.75 Msps sample frequency, W_{BB} = 100 KHz baseband filter bandwidth, $N_R = 10$ fingerprint subregions ($N_R + 1 = 11$ total regions), and $K = 5$ MDA/ML cross-fold validation. Unlike previous results in [15], [16] which used midamble correlation for burst detection and alignment, results presented here are based on near-transient amplitude detection using a -9 dB detection threshold. Near-transient results are based on a signal region spanning 400 samples ($\approx$ 17 μSec) and midamble results are based on a signal region spanning 2200 samples ($\approx$ 92 μSec).

The signals were collected at a fixed location within a typical office environment (desks, chairs, metal filing cabinets, metal book cases, etc.) during a time frame with very heavy intra-cellular GSM activity. The impact of these conditions was 1) clearly observed in the dynamic resource allocation (number of channels, hopping sequence, etc.), and 2) manifest in the collected data which occupied 4 to 8 frequency channels and included the effects of multipath fading and power control. The observed Training Sequence Code (TSC) in the GSM midamble was *identical* for all collected signals.

For each SNR of a given classification experiment, 10000 classification decisions were made using approximately 1000 bursts per device. A sufficient number of independent Monte Carlo noise realizations were considered to achieve statistical significance and enable Confidence Interval (CI) assessment using CI = 95%. Resultant CI values are provided in all *tabular data* presented in the results section. For all *plotted data* in the results section, the size of *all* data markers has been adjusted such that their vertical extent is equal to, or greater than, the corresponding CI. The conventional error bar format for presenting CIs was avoided to enhance visual clarity.

The specific devices used for each inter-manufacturer and intra-manufacturer classification *case* are presented in Table I. As indicated, there were four like-model devices per manufacturer, with each having a unique serial number. For each case, classification results were generated for all four possible device combinations (four devices taken three at a time) and the three class MDA-ML process in Sect. II-B applied. Results from all device combinations for a given case are averaged to produce the overall average classification results. The device model numbers included: Motorola V191, Nokia 6125, Samsung SGH-a226, and Sony Ericsson W300i.

TABLE I
DEVICE USED FOR VARIOUS CLASSIFICATION CASES: MOTOROLA (MOT), NOKIA (NOK), SAMSUNG (SAM), AND SONY ERICSSON (SON) DEVICES AS INDICATED.

	Inter-Manufacturer				***Intra***-Manufacturer			
Case #	1	2	3	4	1	2	3	4
MOT-1	×				×			
MOT-2		×			×			
MOT-3			×		×			
MOT-4				×	×			
NOK-1	×					×		
NOK-2		×				×		
NOK-3			×			×		
NOK-4				×		×		
SAM-1	×						×	
SAM-2		×					×	
SAM-3			×				×	
SAM-4				×			×	
SON-1	×							×
SON-2		×						×
SON-3			×					×
SON-4				×				×

A. Inter-Manufacturer Classification

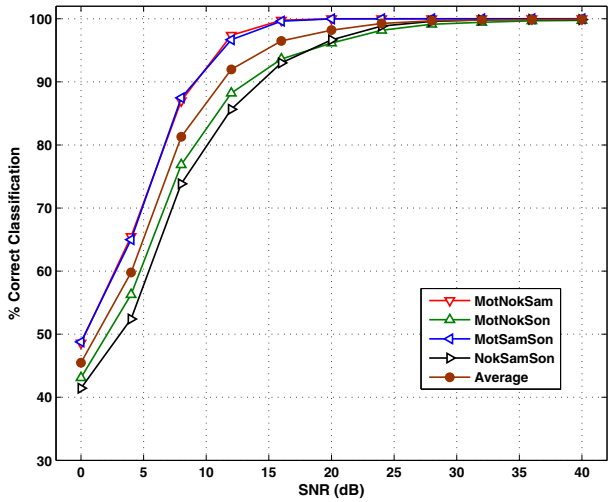
Inter-manufacturer classification performance was analyzed for all cases in Table I using using both midamble and near-transient signal responses at SNR $\in$ [0 40]. Figure 2 shows results for all four Case #1 device combinations and the average of all combinations. The trend observed in these results is that midamble fingerprinting is marginally superior on average for $SNR < 20$ dB. However, it is important to note that this benefit comes at the expense of greatly increased computation time, with midamble fingerprinting (2200 regional samples) taking nearly 250% longer to complete than near-transient fingerprinting (400 regional samples).

The trends in Fig. 2 are consistent with what was observed for all inter-manufacturer cases in Table I. This is reflected in Fig. 3 results which shows the midamble and near-transient combination averages for all four cases considered. As indicated, average classification accuracy of 90% or better is achieved for $SNR \geq 12$ dB using midamble fingerprints and $SNR \geq 16$ dB using near-transient fingerprints.

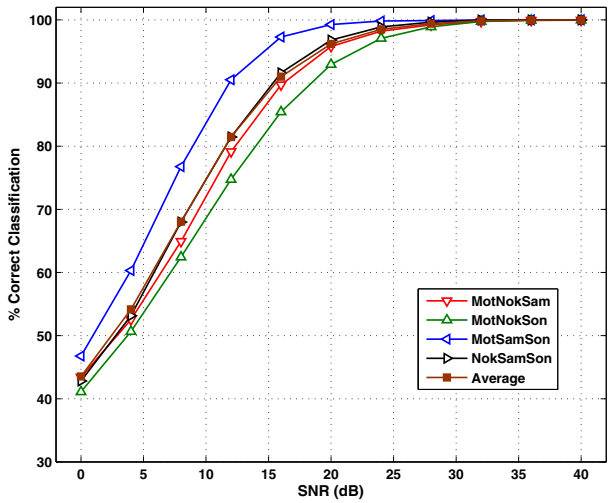
For the $SNRs$ considered, Fig. 3(a) shows that best case and worst case classification performance was achieved with the MotNokSam and NokSamSon combinations, respectively. It is insightful to consider these extremes by looking at representative MDA/ML classification confusion matrices. These are presented in Table II for midamble fingerprinting at $SNR = 20$ dB. Note that the CI values are for the correct class estimate percentages (diagonal entries).

B. Intra-Manufacturer Classification

Intra-manufacturer classification results were generated for all device combinations and cases shown in Table I using the midamble and near-transient signal regions. Figure 4 shows results for all four Case #1 Motorola device combinations and

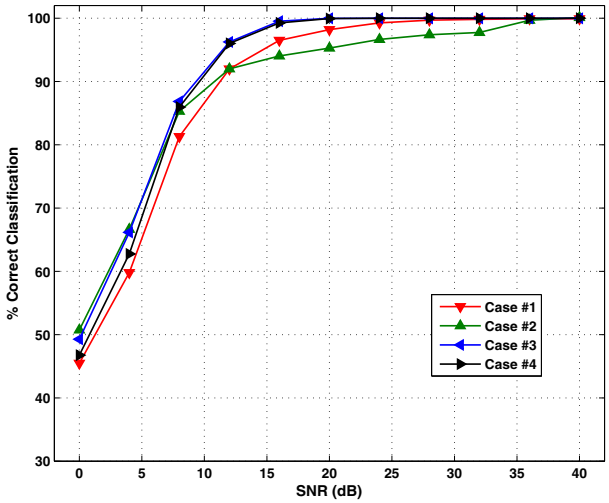


(a) Midamble Signal Region

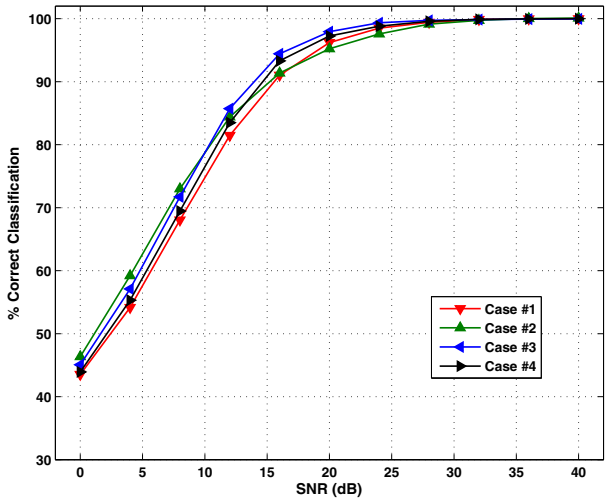


(b) Near-Transient Signal Region

Fig. 2. *Inter-Manufacturer* MDA/ML classification performance using all combinations of devices for Case #1 in Table I. The average across all permutations is shown with filled markers.



(a) Midamble Signal Region



(b) Near-Transient Signal Region

Fig. 3. Average *Inter-Manufacturer* MDA/ML classification performance for all cases in Table I. Averages calculated using results from all four device combinations within each case.

TABLE II
INTER-MANUFACTURER CASE #1 ***Midamble*** FINGERPRINTING: CONFUSION MATRICES FOR $SNR = 20$ DB.

MotNokSam : Overall Ave = 99.99%

	Class Estimate (%)			CI (%)
Actual Class	MOT-1	NOK-1	SAM-1	
MOT-1	**99.99**	0.01	0	0.02
NOK-1	0	**100**	0	0
SAM-1	0	0	**100**	0

NokSamSon : Overall Ave = 96.66%

	Class Estimate (%)			CI (%)
Actual Class	NOK-1	SAM-1	SON-1	
NOK-1	**93.87**	0	6.13	0.47
SAM-1	0	**100**	0	0
SON-1	3.88	0.02	**96.10**	0.38

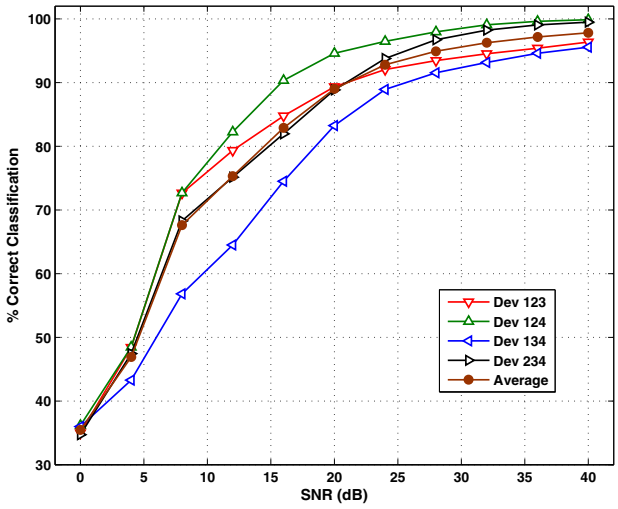
TABLE III
INTRA-MANUFACTURER CLASSIFICATION: CASE #1 ***Midamble*** FINGERPRINTING CONFUSION MATRICES FOR $SNR = 20$ DB.

Dev124 : Overall Ave = 94.60%

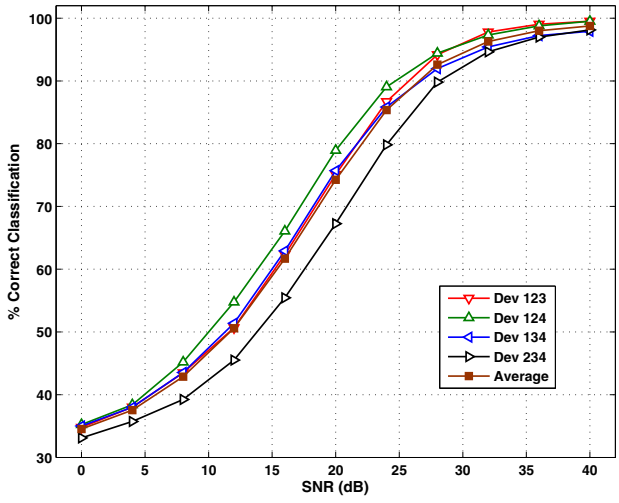
	Class Estimate (%)			CI (%)
Actual Class	MOT-1	MOT-2	MOT-4	
MOT-1	**99.85**	0.01	0.15	0.08
MOT-2	0.01	**93.85**	6.14	0.47
MOT-4	0.95	8.95	**90.10**	0.59

Dev134 : Overall Ave = 83.25%

	Class Estimate (%)			CI (%)
Actual Class	MOT-1	MOT-3	MOT-4	
MOT-1	**86.13**	13.87	0.01	0.68
MOT-3	20.61	**73.71**	5.68	0.86
MOT-4	0.01	10.09	**89.91**	0.59

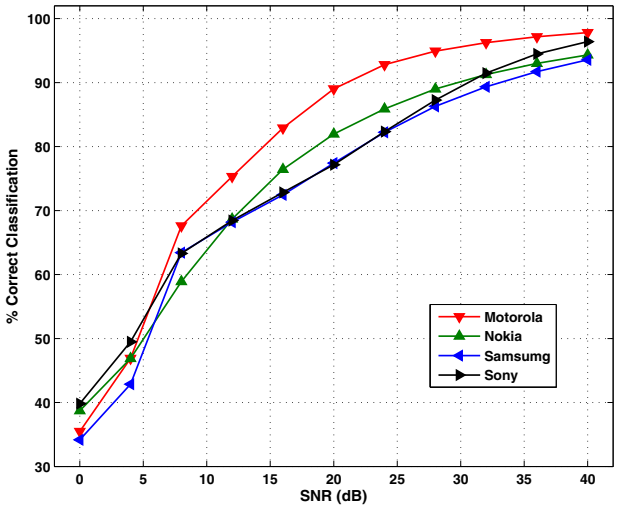


(a) Midamble Signal Region

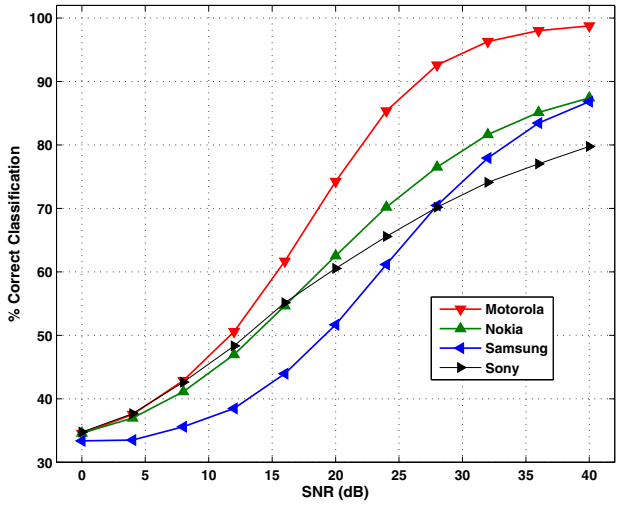


(b) Near-Transient Signal Region

Fig. 4. *Intra-Manufacturer* MDA/ML classification performance using all combination of Motorola devices from Case #1 in Table I. The average across all permutations is shown with filled markers.



(a) Midamble Signal Region



(b) Near-Transient Signal Region

Fig. 5. Average *Intra-Manufacturer* MDA/ML classification performance for all cases in Table I. Averages calculated using results from all four device combinations within each case.

the average of all combinations. The average results from Fig. 4 are shown in Fig. 5 along with average results for the other four inter-manufacturer cases. As indicated, average classification accuracy of 90% or better is achieved for $SNR \geq$ 20 dB using midamble fingerprints and $SNR \geq 26$ dB near-transient fingerprints (8 to 10 dB higher relative to what was demonstrated for inter-manufacturer classification).

Results in Fig. 5 show that Case #1 produced the best overall results and was the only case to achieve accuracy of 90% or better. For the range of $SNRs$ considered, Fig. 4(a) shows that best case classification performance was achieved with Motorola combination Dev 124 while worst case performance was achieved with combination Dev 134. As before, representative MDA/ML classification confusion matrices are provided for these combinations in Table III for $SNR = 20$ dB.

The Table II results are consistent with RF-DNA fingerprint responses in Fig. 1(a), with best case including the MOT-1 device which has the most unique fingerprint and worst case including the devices having the most similar fingerprints (NOK-1, SAM-1 and SON-1).

By comparison with corresponding inter-manufacturer results in Fig. 2 and Fig. 3, the intra-manufacturer results in Fig. 4 and Fig. 5 are notably poorer. This finding is consistent with previous RF fingerprinting work using 802.11 OFDM-based signals [5] which demonstrated that intra-manufacturer classification generally presents the greatest challenge. As before, midamble fingerprinting generally outperforms near-transient fingerprinting in the lower SNR region.

IV. Conclusion

Success with "cracking" bit-level encryption and pledges to continue hacking attacks remains a concern for existing cellular systems such as GSM and emerging systems such as WiMAX. Air monitoring with RF-DNA fingerprints is proposed here as one option for augmenting bit-level protection. Proof-of-concept demonstration is provided here using experimentally collected GSM signals. Previous *inter-manufacturer* device classification results in [15], [16] are extended to include *intra-manufacturer* device classification, i.e., serial number discrimination of like-model GSM devices. Consistent with work using 802.11 OFDM-based signals [5], intra-manufacturer classification presents the greatest challenge.

Fisher-based MDA/ML classification results are presented for *phase-only* RF-DNA fingerprinting, using identical collected data and post-collection processing for all cases, and show that average classification accuracy of 90% or better is achieved for 1) inter-manufacturer cases with $SNR \geq 12$ dB (midamble) and $SNR \geq 16$ dB (near-transient), and 2) intra-manufacturer cases with $SNR \geq 20$ dB (midamble) and $SNR \geq 26$ dB (near-transient).

Acknowledgment

This work sponsored by the Sensors Directorate, Air Force Research Laboratory, and the Tactical SIGINT Technology (TST) Program Office.

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

References

[1] N. Karsten and C. Paget, "GSM-SRSLY?, Presentation Slides, H4RDW4RE, Berlin Laboratory," 2009.

[2] J. Blau, "Open-Source Effort to Hack GSM, spectrum.ieee.org/Telecom/Wireless/Open-Source-Effort-to-Hack-GSM," 2009.

[3] M. Kassner, "Cracking GSM Encryption Just Got Easier, blogs.techrepublic.com.com/wireless, p=206." 2009.

[4] Y. Sheng, G. Chen, D. Kotz, and A. Campbell, "Detecting 802.11 mac layer spoofing using received signal strength." IEEE 27th Conference on Computer Communications (INFOCOM08), Apr 2008.

[5] R. Klein, M. Temple, and M. Mendenhall, "Application of wavelet-based RF fingerprinting to enhance wireless network security," *Jour of Communications and Networks*, vol. 11, no. 6, pp. 544–555, Dec 2009.

[6] O. Ureten and N. Serinken, "Wireless security through RF fingerprinting," *Canadian Journal of Electrical and Computer Engineering*, vol. 32, no. 1, pp. 27–33, Winter 2007.

[7] J. Hall, M. Barbeau, and E. Kranakis, "Radio frequency fingerprinting for intrusion detection in wireless networks," Jul 2005, DRAFT.

[8] ——, "Detection of transient in radio frequency fingerprinting using signal phase." IASTED Int'l Conf on Wireless and Optical Communications, May 2003.

[9] B. Danev and S. Kapkun, "Transient-based identification of wireless sensor nodes." ACM/IEEE Int'l Conf on Information Processing in Sensor Networks (IPSN09), Apr 2009.

[10] O. Tekbas, O. Ureten, and N. Serinken, "Improvement of transmitter identification system for low SNR transients," *IEE Electronics Letters*, vol. 40, no. 3, pp. 182–183, Jul 2004.

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

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

[13] ——, "Using spectral fingerprints to improve wireless network security." 2008 IEEE Global Communications Conference (GLOBECOM08), Mar 2008.

[14] R. Klein, M. Temple, M. Mendenhall, and D. Reising, "Sensitivity analysis of burst detection and RF fingerprinting classification performance." IEEE Int'l Conference on Communications (ICC09), Jun 2009.

[15] D. Reising, M. Temple, and M. Mendenhall, "Improved wireless security for gmsk-based devices using RF fingerprinting," *Int. J. Electronic Security and Digital Forensics*, vol. 3, no. 1, pp. 41–59, Mar 2010.

[16] ——, "Improving intra-cellular security using air monitoring with RF fingerprints." IEEE Wireless Communications and Networking Conference (WCNC10), Apr 2010.

[17] M. Williams, S. Munns, M. Temple, and M. Mendenhall, "Rf-dna fingerprinting for airport wimax communications security." 4th Int'l Conference on Network and Systems Security (NSS10), Sep 2010.

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

[19] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning; Data Mining, Inference, and Prediction*. Springer-Verlag, New York, New York, USA, 2001.