Coding-Scheme Classification with Applications to PHY-Layer Security

Garrett Vanhoy and Tamal Bose
Electrical and Computer Engineering
University of Arizona
Tucson, Arizona 85719
Email: {gvanhoy, tbose}@email.arizona.edu

***Abstract*—The exploitation of side-channel information (SCI) poses a threat to the security of even the most sophisticated systems. SCI generally refers to any information that is exposed from a system employing encryption other than the original or encrypted data. In wireless systems, this information can include signal attributes such as received signal strength, bandwidth, burst duration, modulation, and others. Although encryption prevents an eavesdropper from being able to completely understand traffic being generated between devices, SCI can be exploited to potentially circumvent encryption. Traffic patterns are an especially revealing form of SCI. For example, it has been shown that particular traffic patterns can be used to identify a web page that a user is currently browsing. Many existing techniques used to exploit or extract SCI require knowledge of the protocol being used between devices or being able to extract commonly unencrypted information from protocol headers. In this paper, we discuss how methods to hide physical layer parameters may still be overcome using classification techniques.**

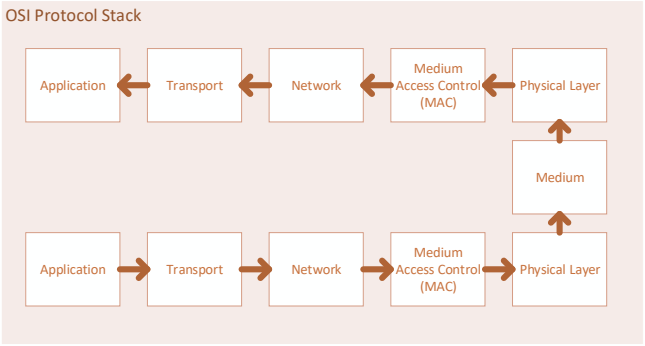

Fig. 1: OSI Protocol Stack

I. Introduction

The fulfillment of security requirements such as data authenticity, integrity, and confidentiality, has always been a challenge for wireless communication systems. This is because in wireless systems, both malicious and legitimate users have access to the same media. In wired systems and malicious user needs to connect physically to the communication system in order to carry out attacks. With the proliferation of wireless access technologies such as Bluetooth and Wi-Fi, the threat of wireless the threat of wireless attacks has become a increasingly important issue. Addtionally, with the increasing availability of general-purpose software defined radios and sophisticated tools, the threat of wireless attacks has become more prominent. As a result there has been a substantial increase in the academic, government, and commercial communities in wireless security.

Communication systems typically adopt the OSI protocol architecture for transporting information which includes several layers. Information from one application to another typically traverses from the upper the layers of the protocol stack through the lower layers and then back up the stack again like that seen in Figure 1.

The protocol stack contains the application, transport, network, medium access control (MAC), and physical layers in decreasing level on the stack. Information from upper layers are always passed down to lower layers as "packets" of generic data to be moved and could take many different forms. Hence, layers are usually only aware of parameters related to their intended purpose and thus they present different vulnerabilities and require corresponding security measures to be in place. However, the physical layer can present vulnerabilities to other parts of the stack. The physical layer, which defines how a medium is used to relay bits that come from all other layers, is naturally more vulnerable to attacks than in wired systems [1]. It has been shown that through observing frame sizes at the physical layer, the type of application being used can be discerned [2].

There are several attacks that can be carried out against wireless systems that mirror that of wired systems. Eavesdropping, denial of service (DoS), spoofing, man-in-the-middle, and message injection are commonly discussed attacks in the realm of security. In wireless systems, since eavesdropping is easier to carry out, other attacks are become even more threatening as well as a result. For example, in a wired network, a common type of DoS attack simply overwhelms a server with many requests for a service, rendering it unable to properly serve many users. A malicous user might carry this out without any knowledge of the state of the server or other clients. In a wireless scenario malicious node can attempt to prevent service for a only for a particular client instead of all clients by understanding only manipulating some traffic.

In this work, background of the existing work in the field of physical layer security in Section II is provided. It is then shown how these existing approaches may be compromised

using common classification techniques with cyclostationary features in Section III. An eavesdropping scenario is the simulated and the results are presented in Section V. Finally we conclude in Section VI.

II. Physical Layer Security

In the realm of wireless security, physical layer security has seen a substantial increase in interest as it has been shown that encryption alone is not enough to achieve security requirements. The physical layer of wireless systems is exposed to both legitimate and nefarious users and this presents two major threats to security: jamming and eavesdropping. In this work, we focus primarily on threats related to eavedropping.

Eavedropping presents a threat to information confidentiality and is generally combatted using encryption. The success of encryption lies entirely in the assumption that an eavesdropper does not have enough computational capacity to "break" the encryption scheme compared to the intended receiver that has additional information about the incoming message. There has been plenty of work with brute-force type attempts to breaking encryption methods. However, the more imminent threat to the efficacy of encryption is the leakage of side channel information (SCI). Side channel information refers to any information that is neither the encrypted payload data (bits) or the unencrypted payload data (also bits). For example, almost every protocol prepends a header to the data payload. This header has a known format and sometimes contains fields that are common or of predictable values. If such a field is encrypted and is known to the eavesdropper it may be possible to reverse-engineer the encryption key in a short time. This is called a known plain-text attack. Other types of SCI that have been exploited include signal strength, bandwidth, data rate, and modulation scheme. These parameters are generally relayed through PHY-layer headers in transmitted frames, but they can also be inferred without the use of these bits using various DSP or statistical techniques. Hence, the obfuscation of SCI or prevention of SCI leakage has become an important topic in wireless physical layer security.

A. Existing Countermeasures

Preventing SCI leakage has been carried out in a few ways. First, SCI leakage can be prevented through effectively reducing the signal to noise ratio (SNR) seen by the eavesdropper relative to the intended receiver. One effective technique for accomplishing this is the use of low probability of intercept (LPI) waveforms that are typically applied in the realm of radar. Since the intended receiver knows the precise method (sequence) by which the transmitted signal has been spread, it achieves a spreading gain over the eavesdropper. Friendly jamming or artificial noise generation also reduces the effective SNR at the eavesdropper by transmitting artificial noise along with its message. By transmitting noise in unused space, time, or frequency, the eavesdropper receives a disadvantage over the intended receiver.

A common and critical assumption of existing work addressing the threat of eavesdropping is that the eavesdropper has similar capabilities as the intended receiver. These results are relevant to common scenarios where commercially available receivers are used in malicious ways such as one cell phone eavesdropping on another. In scenarios where a receiver is specifically designed for surveillance purposes, it is unlikely to be so restricted. There are very few, if any, existing approaches to the issue of eavesdropping that can guarantee an SNR advantage for the intended receiver in the presence of an eavesdropper *with sufficiently capable hardware*.

One such technique that can be used to fulfill security requirements even in the presence of a more capable eavesdropper is full frame encryption [3]. header bits that can be used to determine important information about the frame such as its length, modulation, coding scheme, and the protocol are generally sent unencrypted. By encrypting these bits and obfuscating the modulation and coding scheme, it becomes much more difficult for an eavesdropper to effectively determine what may be happening at higher layers of the OSI stack.

III. Signal Classification

In this section, the method by which signals of the same modulation, but different coding scheme, can be classified using cyclostationary features and modern machine-learning techniques.

A. Cyclostationary Signals

Many signals and systems have been modelled as wide-sense stationary stochastic processes where second-order statistics of the signal remain constant with time, but whose autocorrelation is independent of time. However, many man-made signals exhibit a periodic or an almost-periodic autorrelation function because they contain various periodic structures in time. From a practical perspective, if the received signal is considered as a stochastic process that is a sum of both additive white Gaussian noise (AWGN) and the signals of interest, they each fall into the category of almost-cyclostationary processes. The distinction between cyclostationary (CS) and almost-cyclostationary (ACS) is an important distinction made in the literature and more details can be found in [4]. Such signals can be said to be cyclostationary or periodically correlated and the analysis of this class of processes has been ongoing for decades [5]. Over the years, cyclostationarity has been studied rigorously in continuous and discrete, real and complex, and stochastic and non-stochastic contexts [4].

As an example, let $x(t)$ be a continuous-time real-valued stochastic process. The process $x(t)$ can be called wide-sense cyclostationary if its autocorrelation function, $R_x(\tau) = \int_{-\infty}^{\infty} x(t-\frac{\tau}{2})x(t+\frac{\tau}{2})dt$, is periodic. Due to this periodicity, the process $x(t)$ can be expanded in a Fourier series such that

$$R_x(t,\tau) = \sum_{\alpha \in A} R_x^{\alpha}(\tau) e^{j2\pi\alpha t}, \tag{1}$$

and

$$R_x^{\alpha}(\tau) \triangleq \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} R_x(t,\tau) e^{-j2\pi\alpha t} dt \tag{2}$$

where τ is the lag parameter and A is the set of cycle frequencies α such that $R_x^\alpha(\tau) \neq 0$. Both the coefficients $R_x^\alpha(\tau)$, which are called the cyclic autocorrelation functions (CAF) and their Fourier transforms, which are called the cyclic spectra of the process $x(t)$, are useful in analysis and classification of signals and processes. The cyclic spectra is especially important for analysis as it represents the density of correlation between two spectral components of a process that are separated by α. This property is especially useful for detection because a process that produces additive white Gaussian noise (AWGN) contains no correlation between spectral components making it readily discernible from many signals.

The cyclic spectrum at cycle frequency α of the process $x(t)$ can be written

$$S_x^\alpha(f) = \int_{-\infty}^{\infty} R_x^\alpha(\tau) e^{-j2\pi f\tau} d\tau, \tag{3}$$

which can interpreted as the time-averaged statistical correlation of two spectral components seperated by cycle frequency α as the bandwidth of each spectral component approaches zero. For this reason, the cyclic spectrum can also be called the spectral correlation density (SCD). According to this definition $S^0(f)$ is actually the traditional power spectral density (PSD) of the process $x(t)$. The SCD has been used in a variety of signal processing and classification tasks in communications, radar, and others [6], [7], [8], [9]. This primarily stems from its ability to detect and characterize the presence of cyclic features such as cyclic prefix length, symbol period, or carrier frequency even in the presence of noise and other channel effects.

B. Estimation of the SCD

The difficulty with using this feature is that it is computationally complex to estimate for digital signals. A single estimate of the SCD of a reasonable resolution can require up to 65,536 of 32-point complex FFTs. The SCD is readily derivable in closed form for many continuous forms of communications signals and a substantial efforts has been made decades ago to estimate this quantity for a finite-duration digital signal. For a digital signal, estimating the SCD has two commonly-used methods which are optimized for computational efficiency. The FFT Accumulation method (FAM) [10] and Spectral Strip Correlation Algorithm (SSCA) [11] are the two variations to estimate the SCD. The FAM, due its data parallel computations and regular data access patterns, offers opportunities for exploiting parallelism [12], particularly on hardware architectures that allow fine-grained parallelism. The FAM is the most efficient computationally and is calculated as

$$X_{N'}(n,k) = \sum_{r=-N'/2}^{r=N'/2} a[r]x[n-r]e^{-j2\pi k(n-r)T_s}. \tag{4}$$

with

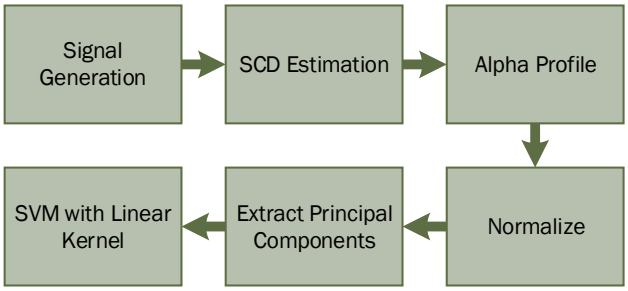

Fig. 2: Classification System

$$S_x^\alpha(n,k) = \frac{1}{N}\sum_{n=0}^{N-1} \frac{1}{N'} X_{N'}\left(n, k+\frac{\alpha}{2}\right) X_{N'}^*\left(n, k-\frac{\alpha}{2}\right). \tag{5}$$

where $N^{'}$ and N together determine a resolution in both the time and frequency domains and $a[n]$ is an arbitrary windowing function. Equation (4) is the sliding window discrete Fourier transform (DFT) with window $a[n]$.

C. Classification by Cyclostationarity

The proposed system to classify signals can be seen in Figure 2. The SCD is first estimated using the FAM method with parameters, $N^{'} = 32$, $N = 256$, and $a[n]$ is a hamming window of length $N^{'}$. Second, the maximum is taken along the α axis of estimate of $S_x^\alpha(f)$ to narrow down the number of features the classifier needs to train with. Next, these values are normalized with respect to the maximum value in each set so as to emphasize the relative values of all the features. Then, using principal component analysis, the 4096 features are narrowed down to 25 features that account for the majority of the variance between the classes. To classify among each class, a support vector machine with a linear kernel is trained on a subset of the provided training data.

IV. Simulations

Simulations were constructed using GNU Radio to generate signals for the classification system. The signals were generated with random data for every frame. Using different bits for every frame ensures that classification is not done based on the actual data in the frame. After this, several different channel codes were implemented including a convolutional code of rate 1/2, trellis code of rate 3/4, and no coding at all. These codes are commonly used to decrease bit error rates in wireless systems. Next, each bit was mapped onto a 16-QAM signal constellation and pulse-shaped with an interpolating root-raised cosine filter of transition with one-fourth the sample rate. To simulate the reception of a bandpass signal, the signal was then carrier modulated to a frequency of 1/4 the sample rate. The signal was then carried through an additive white Gaussian noise (AWGN) channel. The imaginary part of the signal was then discarded so that the SCD features that will be generated later are more rich.

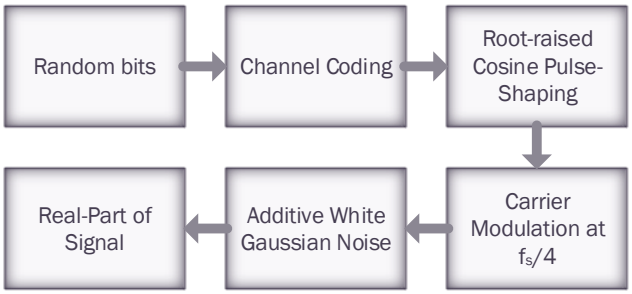

Fig. 3: Simulated Wireless System

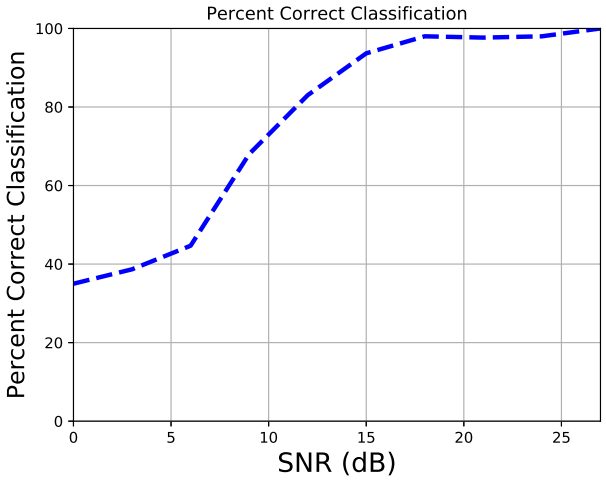

Fig. 4: Percent Correct Classification vs SNR

V. Results

To evaluate the ability of the proposed classification system, 100 signals of each type of channel coding were generated for each SNR tests. The classifier was trained on a random sample of 1/3 of the total samples generated. The results are shown in Figure 4. The results show that although the same modulation is sent, it may be possible to determine the channel coding with some degree of accuracy.

VI. Conclusions

In the field of wireless security, the exposition of side-channel information poses a threat to many systems. It has been shown in recent studies that even basic information about a waveform such as the number of bits in each frame can be used to compromise data confidentiality. Many techniques have been proposed to deter eavesdropping of malicious users that have similar hardware, but the scenario where eavesdroppers have more capable hardware has not been studied in detail. This work shows that even in the face of techniques such as full frame encryption, which hides information critical to determining the number of bits in a frame, it is possible to classify the channel coding being used. This suggests that if the dictionary of possible channel codes of obfuscating modes are known to an eavesdropper, that with enough time and samples the frame length could be determined with a high degree of certainty.

VII. Acknowledgements

This project was partially supported by the Broadband Wireless Access and Applications Center (BWAC); NSF Award No. 1265960.

References

[1] Y. Zou, X. Wang, and L. Hanzo, "A Survey on Wireless Security: Technical Challenges, Recent Advances and Future Trends," vol. 104, no. 9, pp. 1–31, 2015. [Online]. Available: http://arxiv.org/abs/1505.07919

[2] S. Chen, R. Wang, X. F. Wang, and K. Zhang, "Side-channel leaks in web applications: A reality today, a challenge tomorrow," *Proceedings - IEEE Symposium on Security and Privacy*, pp. 191–206, 2010.

[3] H. Rahbari and M. Krunz, "Full frame encryption and modulation obfuscation using channel-independent preamble identifier," *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 12, pp. 2732–2747, 2016.

[4] W. A. Gardner, "Cyclostationarity: Half a century of research," *Signal Processing*, vol. 86, no. 4, pp. 639–697, 2006.

[5] H. L. Hurd, *An investigation of periodically correlated stochastic processes*. University Microfilms, 1970.

[6] Q.-Q. Lu, M. Li, and X.-J. Wang, "Improved method of spectral correlation density and its applications in fault diagnosis," *Beijing Keji Daxue Xuebao/Journal of University of Science and Technology Beijing*, vol. 35, no. 5, pp. 674–681, 2013.

[7] D.-S. Yoo, J. Lim, and M.-H. Kang, "Atsc digital television signal detection with spectral correlation density," *Journal of Communications and Networks*, vol. 16, no. 6, pp. 600–612, 2014.

[8] Z.-B. Zhang, L.-P. Li, and X.-C. Xiao, "Detection and chip rate estimation of mpsk signals based on cyclic spectral density," *Xi Tong Gong Cheng Yu Dian Zi Ji Shu/Systems Engineering and Electronics*, vol. 27, no. 5, pp. 803–806, 2005.

[9] L. Zhu, H.-W. Cheng, and L.-N. Wu, "Identification of digital modulation signals based on cyclic spectral density and statistical parameters," *Journal of Applied Sciences / Yingyong Kexue Xuebao*, vol. 27, no. 2, pp. 137–143, 2009.

[10] S. R. Schnur, "Identification and classification of ofdm based signals using preamble correlation and cyclostationary feature extraction," DTIC Document, Tech. Rep., 2009.

[11] D. Simic and J. Simic, "The strip spectral correlation algorithm for spectral correlation estimation of digitally modulated signals," in *Telecommunications in Modern Satellite, Cable and Broadcasting Services, 1999. 4th International Conference on*, vol. 1, 1999, pp. 277–280 vol.1.

[12] A. R. Castro, L. C. Freitas, C. C. Cardoso, J. C. Costa, and A. B. Klautau, "Modulation classification in cognitive radio."