PHY-Layer Authentication Using Hierarchical Modulation and Duobinary Signaling

Vireshwar Kumar*, Jung-Min “Jerry” Park*, T. Charles Clancy*, Kaigui Bian†
*Bradley Department of Electrical and Computer Engineering, Virginia Tech
†School of Electronics Engineering and Computer Science, Peking University

Abstract—In a cognitive radio network, the non-conforming behavior of rogue transmitters is a major threat to opportunistic spectrum access. One approach for facilitating spectrum enforcement and security is to require every transmitter to embed a uniquely-identifiable authentication signal in its waveform at the PHY-layer. In existing PHY-layer authentication schemes, known as blind signal superposition, the authentication/identification signal is added to the message signal as noise, which leads to a tradeoff between the message signal's signal-to-noise (SNR) and the authentication signal's SNR under the assumption of constant average transmitted power. This implies that one cannot improve the former without sacrificing the latter, and vice versa. In this paper, we propose a novel PHY-layer authentication scheme called hierarchically modulated duobinary signaling for authentication (HM-DSA). HM-DSA introduces some controlled amount of inter-symbol interference (ISI) into the message signal. The redundancy induced by the addition of the controlled ISI is utilized to embed the authentication signal. Our scheme, HM-DSA, relaxes the constraint on the aforementioned tradeoff and improves the error performance of the message signal as compared to the prior art.

I. INTRODUCTION

In the spectrum sharing paradigm, a heterogeneous mix of wireless devices need to coexist without causing harmful interference to each other [1]–[3]. To ensure the viability of spectrum sharing, effective and low-cost spectrum (rule) enforcement measures must be adopted. One of the critical challenges in spectrum enforcement is identifying non-conforming (i.e., “rogue”) transmitters that violate spectrum access rules prescribed by the spectrum regulatory authorities (e.g., FCC). In this paper, we propose a PHY-layer authentication scheme that can be used to identify rogue transmitters.

In this heterogeneous coexistence model, we can define two types of intended receivers—unaware and aware receivers [4]. An unaware receiver is able to correctly demodulate and decode the message signal, but cannot authenticate the received signals, either because it has no knowledge of the authentication scheme or it does not share the key with the transmitter. Also a receiver that does not intend to authenticate the received signals is classified as an unaware receiver. On the other hand, a receiver, interested in the message signal as well as the authentication signal embedded into the message signal in order to identify the transmitter and authenticate its signals, is called an aware receiver.

A PHY-layer authentication scheme should embed the authentication signal into the message signal such that it enables the aware receiver to extract the authentication signal, while at the same time, enables the unaware receiver to recover the message signal without requiring the unaware receiver to change its demodulation or decoding procedure. As a result, in most of the existing schemes [4]–[6], the authentication signal is added to the message signal as noise. In such an approach, the message signal is decoded in the presence of the authentication signal, thus resulting in decreased signal-to-noise ratio (SNR) for the message signal, assuming average transmission power has not been increased to embed the authentication signal. This means that the degradation in the message signal’s SNR is significant when the authentication signal’s SNR is increased to a level sufficient for authenticating the embedded signal at the receiver [7]. Hence, there is a fundamental tradeoff between the message signal’s SNR and the authentication signal’s SNR.

In this paper, we propose a novel PHY-layer authentication scheme that is based on duobinary signaling, a waveform shaping technique that has been traditionally used to increase bandwidth efficiency [8]. Our scheme, called hierarchically modulated duobinary signaling for authentication (HM-DSA), utilizes the redundancy induced in the message signal due to the addition of inter-symbol interference (ISI) to embed the authentication signal. A high-level description of HM-DSA was briefly discussed in [9]. In this paper, we elaborate on the idea.

The main contributions of this paper are summarized below.
• We propose a novel PHY-layer authentication scheme and show that our approach relaxes the constraint on the aforementioned tradeoff that affects the existing schemes.
• The proposed approach achieves a detection performance advantage (for message signals), compared to prior art.

The rest of the paper is organized as follows. We provide the related work in Section II. We describe our approach in Section III, and discuss our methodology for PHY-layer authentication in Section IV. We analyze our results in Section V, and compare our scheme with the prior art in Section VI. Section VII concludes the paper by highlighting the main contributions.

II. RELATED WORK

It has been shown in [10] that the identification of a device based on transmission imperfections exhibited by its radio transmitter, is prone to impersonation attacks. Therefore, in order to achieve transmitter authentication in a cognitive radio environment, it is imperative for a transmitter to embed an
authentication/identification signal into the message signal at PHY-layer in such a way that an aware receiver is able to authenticate the transmitter without hampering an unaware receiver’s ability to decode the message signal. One approach to achieve this, is to add the authentication signal to the message signal as noise [4]. The schemes following this approach can be categorized as blind signal superposition. To limit the detrimental effects of the authentication signal on the message signal, the principle of hierarchical modulation [5], [6] is often applied—i.e., the authentication signal (low priority signal) is carried on the low-power, high-resolution constellation while the message signal (high priority signal) is embodied by the high-power, low-resolution constellation. However, this leads to the aforementioned tradeoff between the error performance of the message signal and the authentication signal. We note that although both the schemes, HM-DSA and the one proposed in [5], employ amplitude based hierarchical modulation, each scheme uses a different approach for constructing the authentication signal’s constellation. HM-DSA uses duobinary signaling to create the constellation, whereas the scheme proposed in [5] embeds the authentication into the message signal as noise. While the schemes presented in [11]–[13] avoid the aforementioned tradeoff, they require every intended receiver to be aware of the authentication mechanism.

III. TECHNICAL BACKGROUND

In this section, we provide our approach for intentionally adding ISI to the message signal at the transmitter, and removing ISI to estimate the message signal at the receiver. We assume that the message signal is transmitted in blocks of binary sequences, each with length $N$, and represented by $\{d_n\}, \ n = 1, 2 \cdots, N$. Using non-return-to-zero (NRZ) encoding, a bipolar state sequence, $\{w_n\}$, is generated from $\{d_n\}$. Further, a duobinary sequence, $\{y_n\}$, is generated by adding the delayed and weighted states of $\{w_n\}$. It is achieved by using a digital filter represented by $y_n = w_n + \delta \cdot w_{n-1}$, where $0 < \delta < 1$. Hence, the ISI introduced to each $y_n$, corresponding to the state $w_n$, comes only from the preceding state, $w_{n-1}$. Moreover, the extent of ISI is controlled by $\delta$.

For $w_n = \pm 1$, we obtain a four-level hierarchically modulated output—i.e., $y_n$ has one of four possible values: $+1 + \delta$, $+1 - \delta$, $-1 + \delta$ or $-1 - \delta$ as shown in Figure 1a. We note that the four-level output of $y_n$ is used to express one of the two binary values of the message signal, and hence there is an inherent redundancy in this process. We also observe that the encoded signal, $y_1$, is given by $y_1 = w_1 + \delta \cdot w_0$, where $w_1$ and $w_0$ are the bipolar states for $d_1$ and $d_0$, respectively. Hence, we require an extra bipolar state, $w_0$ and a corresponding bit, $d_0$, to start the encoding of the message signal, $\{d_n\}$, $n = 1, 2 \cdots, N$. Bit $d_0$ is called an initialization bit, and the bipolar state $w_0$ is called an initialization state.

Having estimated the received sequence as $\{\hat{y}_n\}$, the following two decoding methods can be utilized to estimate the received binary message sequence, $\{\hat{d}_n\}$.

1) Symbol-by-Symbol Detection (SSD): With this method, $\{\hat{d}_n\}$ is directly obtained using the following decoding decision rule.

\[
\hat{d}_n = \begin{cases} 
0, & \text{if } \hat{y}_n = -1 + \delta \text{ or } -1 - \delta; \\
1, & \text{if } \hat{y}_n = +1 + \delta \text{ or } +1 - \delta.
\end{cases}
\]  

(1)

2) Maximum Likelihood Sequence Detection (MLSD): The four-level duobinary signal incurs an increase in the number of constellation points in Euclidean space compared to binary signaling. This implies that the error performance of the message decoding using SSD is inferior to that of binary signaling. However, the receiver can exploit the following inherent conditions, arising out of addition of controlled ISI, to significantly improve the message signal’s error performance.

- There can never be a direct transition from signal levels $+1 + \delta$ to $+1 - \delta$ and $-1 - \delta$ to $-1 + \delta$ (e.g., with $\delta = 0.3$, $+1.3$ has to be followed by a signal level $-0.7$ before transitioning to signal level $+0.7$).
- The change in polarity in consecutive signal levels (e.g., $+0.7$ to $-0.7$) is possible if and only if there is change in consecutive bits (e.g., $1$ to $0$).

Moreover, the encoded sequence is generated from a bipolar sequence and has memory of length 1—i.e., the current state is related only to the previous state. Hence, we can use MLSD (based on Viterbi trellis decoding) with two states (i.e., $+1$ and $-1$) to obtain an estimate of the bipolar state sequence, $\{\hat{w}_n\}$ as shown in Figure 1b. Finally, the estimated binary sequence, $\{\hat{d}_n\}$, is generated from $\{\hat{w}_n\}$ using NRZ decoding.

IV. HIERARCHICALLY MODULATED DUOBINARY SIGNALING FOR AUTHENTICATION (HM-DSA)

A. Network Model

We assume that Alice, Bob, Charlie, and Eve share the same wireless medium. Alice, the transmitter, and Bob, the aware receiver, have agreed on a keyed authentication scheme.
that allows Bob to verify the messages he receives from Alice. To authenticate, Alice embeds an authentication signal to the message signal. Charlie, the unaware receiver, does not know the authentication scheme and cannot authenticate Alice’s messages, but can demodulate and decode the message signal. Eve, the adversary, has knowledge of the authentication scheme, but does not know the key. In this paper, our focus is on the PHY-layer authentication technique that enables Alice to embed the authentication signal into the message signal transmitted to Bob so that the embedded signal is compatible with Charlie.

B. Embedding and Extraction of the Authentication Signal

The binary message signal to be transmitted by Alice, $MS_a$, is divided into blocks of binary message sequences each with length $N$. Each block of the binary message sequence is represented by $\{d_n\}$, $n = 1, 2, \cdots, N$. The bipolar state sequence, $\{w_n\}$ and the duobinary sequence, $\{y_n\}$ are obtained as described earlier. However, the initialization bit is replaced with an authentication bit, which is the core idea behind HM-DSA. Hence, the encoding of each block of the binary message sequence is initiated with a bit of the authentication signal to be transmitted by Alice, $AS_a$, represented by $\{a_k\}$, $k = 1, 2, \cdots, K$. This implies that the authentication bit, $a_k$ is embedded into the $k^{th}$ block of the message signal, and hence, for $k^{th}$ block of the embedded signal, the first encoded symbol is given by $y_1 = w_1 + \delta \cdot w_0$, where $w_1$ and $w_0$ are the bipolar states for $d_1$ and $a_k$, respectively.

Table I illustrates the HM-DSA scheme through an example. Assume that the message sequence, $\{011011\}$ is divided into two sequences $\{011\}$ and $\{011\}$. In HM-DSA, the encoding is initiated by $w_0$ which is obtained by NRZ encoding of the authentication bit $a_1 = 1$ for the first sequence, and $a_2 = 0$ for the second sequence. Therefore, the first symbol of the resulting encoded sequences for the same binary message sequences are different. In this way, we have embedded $a_1 = 1$ and $a_2 = 0$ into the first binary message sequence and the second binary message sequence, respectively.

Having estimated the received signal, Charlie, the unaware receiver decodes only the message signal using SSD—i.e., the estimated binary message sequence, $MS_c$, is obtained using the decision rule given by equation (1). Table II provides an example, illustrating the results of performing SSD of the message sequence encoded in Table I.

Bob, the aware receiver, estimates the message signal, $MS_b$ by utilizing MLSD to decode the received sequence. Since the initialization bit is an authentication bit (of the estimated authentication signal, $AS_b$), it also has to be estimated through MLSD. In effect, for each block, the aware receiver executes a minimum distance check on the first received signal level—i.e., it estimates the path out of the four possible paths between the initialization state and the first state of the message signal in the trellis shown in Figure 1b. Hence, the authentication bit and the first message bit are determined by the estimated initialization state and the first message state, respectively.

In other words, for each block, the aware receiver determines the closest signal level out of the four possible levels, and decides the first bit of the estimated message signal, $MS_b$, to be 1 or 0. Hence, although the aware receiver utilizes MLSD for $MS_b$, the authentication bit and first bit in each block of $MS_b$ are determined by SSD. From Figure 1b, we can readily observe that the signal level of $-1 + \delta$ or $-1 - \delta$ can be detected for the first zero-valued message bit only if the authentication bit’s state is $+1$ or $-1$, respectively. But if the first signal level is $+1 - \delta$ along with first message bit’s state $+1$, the authentication bit’s state has to be $-1$. Similarly, the authentication bit’s state can be decided to be $+1$ if the first signal level is $+1 + \delta$ along with first message bit’s state $+1$.

V. Analysis

A. Minimum Euclidean Distance

From Figure 1a, we observe that the message and authentication signals are embodied in two different constellations, i.e., the message signal is carried in the low-resolution constellation and the authentication signal is carried in the high-resolution constellation. The effect of this multi-resolution modulation can be observed when we compare the minimum Euclidean distance of the symbols in the message signal’s constellation with that of the authentication signal’s constellation. We follow the discussions of generalized quadrature amplitude modulation (QAM) in [14] to obtain the minimum Euclidean distance for the message signal as

$$d_{m, HM-DSA} = \sqrt{E_b \cdot \left( \frac{1 - \delta}{\sqrt{(1 + \delta^2)}} \right)}.$$  \hspace{1cm} (2)$$

where $E_b$ is the bit energy. The minimum Euclidean distance for the authentication signal is given by

$$d_{a, HM-DSA} = \sqrt{E_b \cdot \left( \frac{\delta}{\sqrt{(1 + \delta^2)}} \right)}.$$ \hspace{1cm} (3)$$

Hence, the error performance of $MS_c$ and $AS_b$ are dependent on $d_{m, HM-DSA}$ and $d_{a, HM-DSA}$, respectively. However, the error performance of message signal at the aware receiver, $MS_b$, is improved by using MLSD instead of SSD. Figure 2 shows the bit error rate (BER) vs. $E_b/N_0$ curves in HM-DSA with $\delta = 0.3$ and $N = 16$, where $N_0$ is the noise power spectral density. We use the error performance of the message signal without ISI, and using binary phase shift keying (BPSK) modulation, $MS_{b_{bpsk}}$, as the benchmark. We observe that while the error performance of $MS_c$ is significantly worse as compared to that of $MS_{b_{bpsk}}$, the error performance of $MS_b$ is very close to that of $MS_{b_{bpsk}}$.

Further, we can readily see that the error performance of $MS_c$, although worse than that of $MS_b$, is noticeably better than that of $AS_b$. This means that when the ISI is $\delta = 0.3$, the shift in the constellation points (from their conventional positions) is not significant enough to cause a significant drop in error performance of message signal detection. However, this relatively small shift in amplitude makes decoding of the authentication signal difficult.
We note that in HM-DSA, the error performance of $MS_b$ is affected by using the initialization bit as the authentication bit. Since the first bit in a message block is detected using SSD, the first few bits are more prone to error than the later bits in the message block. However, if the initialization bit is kept constant, i.e., no authentication bit is embedded, and the receiver has perfect knowledge of the initialization bit, the error performance of $MS_b$ using MLSD would be asymptotically equal to that of binary signaling.

**B. Effect of $N$**

In HM-DSA, as one authentication bit is inserted in each block of $N$ message bits, the authentication rate (the rate at which the authentication signal is embedded into the message signal) is given by $1/N$. It can readily be inferred that increasing $N$ leads to a lower authentication rate, but larger trellis length for MLSD resulting in higher error performance of $MS_b$. On the other hand, changing $N$ doesn’t affect the error performance of $MS_c$, since $MS_c$ is decoded using SSD. Also, changing $N$ in MLSD does not improve the error performance of $AS_b$ as the authentication signal is determined by each block’s first received symbol which, in turn, is estimated through SSD.

We note that the authentication rate as high as 1 can be achieved by using $N = 1$. However, in this case, MLSD actually performs SSD which means that the error performance of $MS_b$ as well as $MS_c$ are equal. However, even after adding the controlled ISI, $MS_b$ can be detected with nearly the same error performance as without ISI if MLSD, with sufficiently long trellis length ($N >> 1$), is used. Note that we need to authenticate the transmitter, and not the data contained in the message signal. Hence, even very low authentication rate is sufficient for PHY-layer authentication [6].

**C. Effect of $\delta$**

It is evident from the equations (2) and (3) that the minimum Euclidean distance of the message signal degrades and the minimum Euclidean distance of the authentication signal improves, when we increase the ISI by increasing $\delta$. Hence, as the presence of the authentication signal becomes more dominant (by increasing $\delta$) in HM-DSA, the error performance of $MS_c$ degrades. This phenomenon can be attributed to the fact that the message signal’s detection at the unaware receiver in HM-DSA is constrained by the tradeoff between the message signal’s SNR and the authentication signal’s SNR. However, in blind signal superposition approaches, this tradeoff is unavoidable for both, the unaware as well as the aware receiver, because the authentication signal is embedded in the message signal in such a way that the authentication signal appears as noise to the message signal. On the other hand, in HM-DSA, the aware receiver can overcome the loss in performance by utilizing MLSD with larger value of $N$—i.e., larger block length and lower authentication rate. This is an advantageous feature of HM-DSA compared to prior art.

In Figure 3, we observe that while the error performance of $AS_b$ enhances by increasing the ISI—i.e., $\delta$, the error performance of $MS_c$ degrades. Note that the effect of the increase in ISI on $MS_b$ is compensated by increasing $N$, and hence the error performance of $MS_b$ remains nearly unaffected.

**D. Security**

We assume that the binary authentication signal can be generated using the technique described in [9]. Hence, the authentication signal is immune to impersonation and replay attacks. However, Eve can successfully corrupt the authentication signal by transmitting specific noise so that Bob cannot verify the authenticity of the message transmitted by Alice. We refer to such an attack as an *obstruction of authentication* (OOA) jamming attack [9].

In HM-DSA, the authentication signal is embedded into the message signal to obtain a hierarchically modulated signal, where the message signal is embodied by a high-power constellation while the authentication signal is carried on a low-power constellation. An adversary can emit just enough interference to exploit the power difference between the two constellations to disable the decoding of the authentication signal without disabling the decoding of the message signal. This attack can be quite effective against all PHY-layer authentication schemes that are based on hierarchical modulation.
we again obtain the relationship as $\delta$ Euclidean distance of the authentication signal in HM-DSA, and black crosses represent the embedded signal).

Fig. 4: Constellation (red circles represent the message signal and black crosses represent the embedded signal).

VI. COMPARISON WITH PRIOR ART

In this section, we compare HM-DSA against a benchmark that is representative of the prior art: authentication tagging using modulation (ATM) [6]. We utilize HM-DSA and ATM, respectively, with a quadrature phase-shift keying (QPSK) modulated message signal to obtain the embedded signal. Therefore, by design, ATM and HM-DSA have same message throughput. Also, the overall average transmission power is kept unchanged from standard QPSK.

In HM-DSA, the controlled ISI added to the QPSK signal results in a constellation with 16 possible symbols as shown in Figure 4a. On the other hand, ATM utilizes the phase based hierarchical modulation to embed the authentication signal which leads to a constellation of 8 possible symbol positions as shown in Figure 4b. In ATM, an authentication bit of 1 is embedded by shifting the phase of a QPSK message constellation symbol towards the $Q$-axis by $\theta$. An authentication bit of 0 is embedded by shifting the phase towards the $I$-axis by $\theta$.

ATM as well as HM-DSA intentionally corrupts the message signal to insert the authentication signal. Hence, the error performance of $MS$ decreases as the authentication signal becomes more prominent. For ATM, this degradation in the error performance can be formulated by the minimum Euclidean distance given by $d_{m,ATM} = \sqrt{E_b} \cdot (\cos \theta - \sin \theta)$. Hence, in order to limit the degradation in the error performance of $MS$ to the same extent in ATM and HM-DSA, i.e., to achieve the same minimum Euclidean distance, we obtain the relationship between $\theta$ (phase shift in ATM) and $\delta$ (ISI in HM-DSA), to be $\delta = \tan \theta$. The minimum Euclidean distance of the authentication signal in ATM is given by $d_{a,ATM} = \sqrt{E_b} \cdot \sin \theta$. Comparing it with the minimum Euclidean distance of the authentication signal in HM-DSA, we again obtain the relationship as $\delta = \tan \theta$. As a result, using $\delta = \tan \theta$, we obtain the same error performance of $MS$ in both these schemes.

In ATM, there is no way by which the aware receiver can enhance its performance. However, in HM-DSA, the aware receiver can enhance its performance by utilizing MLSD. For a PHY-layer authentication scheme to be viable, the receiver must be able to decode both the message and the authentication signals with sufficiently good error performance. Considering this requirement, we can conclude that HM-DSA enjoys an advantage over ATM.

If we use $\theta = \arctan \delta$ for ATM, the error performance curve labeled “$AS_0$ in HM-DSA” in Figure 2, would also correspond to the error performance of $AS_0$ in ATM. Moreover, the error performance curve labeled “$MS$ in HM-DSA” in Figure 2, would correspond to the error performance of $MS$ as well as $MS_0$ in ATM. Hence, we infer that HM-DSA has a significant error performance advantage over ATM in terms of $MS_0$. Note that this performance can be further enhanced by trading off with the authentication rate.

VII. CONCLUSION

We proposed a novel PHY-layer authentication scheme referred to as hierarchically modulated duobinary signaling with authentication (HM-DSA). One of the biggest drawbacks of most of the existing schemes is that they are constrained by the fundamental tradeoff between the message signal’s SNR and the authentication signal’s SNR. HM-DSA relaxes this constraint. However, our scheme’s advantage over the prior art in terms of the message signal’s error performance is achieved at the cost of lower rate of authentication, and higher complexity of transmitter/receiver.

REFERENCES


