EAI Endorsed Transactions
on Cognitive Communications

Research Article ICST.ORG

Detecting Multi-Channel Wireless Microphone User Emulation Attacks in White Space with Noise

Dan Shan, Kai Zeng *, Weidong Xiang, Paul Richardson

4901 Evergreen Rd, Dearborn, MI, USA, 48092

Abstract

Cognitive radio networks (CRNs) are susceptible to primary user emulation (PUE) attacks. Conventional PUE attack detection approaches consider television broadcasting as the primary user. In this work, however, we study a special kind of PUE attack named wireless microphone user emulation (WMUE) attack. Existing work on WMUE attack detection deals with single channel senario. Although multi-channel WM (MCWM) systems are common, detecting WMUE attacks under a multi-channel setting in noisy environments has not been well studied. In this work, we propose a novel multi-channel WMUE attack detection scheme which operates in low signal-to-noise ratio (SNR) environments with low computational complexity, thanks to the first 1.5-bit FM demodulator whose outputs are represented by only 0, 1 and -1. Experimental results show that, the proposed scheme can effectively detect multi-channel WMUE attacks within 0.25 second when SNR is lower than 6 dB.

Received on 20 May 2013; accepted on 12 June 2013; published on 14 July 2014

Keywords: CRN, WMUE, MCWM, FM demodulator

doi:10.4108/cogcom.1.1.e4

1. Introduction

Cognitive radio (CR) enables secondary users (SUs) to share the spectrum temporarily unused by primary users (PUs). To open the door for this new technique and enhance the spectrum efficiency, regulators in many countries have issued permission for radio frequency (RF) transmissions for license-exempt users on part of television (TV) bands, known as white space. The wireless devices that are carried by SUs and operate on white space are called white space devices (WSDs).

WSDs perform spectrum sensing on white space to avoid collisions to the signals from PUs (incumbent signals), mainly including TV signals and wireless microphone (WM) signals. Many spectrum sensing techniques are proposed to detect these two types of incumbent signals [3, 8, 12, 16, 20]. When PUs emerge, SUs are required to evacuate from the spectrum in order to avoid interference to PUs. Exploiting this policy adversely, an attacker may block all SUs within an area by emulating the signal of a certain PU. This kind of attack is named primary user emulation (PUE) attack [4].

Over the years, tremendous efforts have been expended in the area of PUE attack detection. By evaluating the received signal's coverage area, one can differentiate between the signal from a PUE attacker and the real TV signal [4, 22]. However, these detection techniques do not apply to the attack that emulates WM signals (named WM user emulation attack, or WMUE attack), because WM signals may be transmitted from anywhere. Moreover, WMUE attacks may be launched on a frequency band where no WM system has ever worked on; as a result, one cannot detect these attacks by comparing their channel-specific features with the features contained in real WM signals [14]. In short, detecting a WMUE attack is not easy, while launching a WMUE attack is as simple as building a cheap FM modulator.

Existing work detects WMUE attacks in a single-channel system by comparing the FM signal with the audio signal acquired simultaneously [5]. Since a WMUE attacker wants to abuse the white space and meanwhile hide himself, he is not willing to generate any audio signals correlated with the FM signal(s) he transmits, and this fact leads to low similarity between the FM signal and audio signal around the WM system.

*Corresponding author. Email: kzeng@umich.edu

Although multi-channel WM (MCWM) systems are are common, PUE attacks in these systems are rarely studied, leaving several open challenges. Firstly, multiple WM users in the same MCWM system may speak simultaneously. This situation frequently happens; for examples, multiple performers sing a song at the same time on a stage, or several invited speakers on a conference are having a heated discussion with many overlapped talks. Then the audio signals on different channels interfere each other, and the relationship between the mixed audio signal and the FM signals on multiple audio channels become more complicate. Secondly, the audio signal and FM signals are further contaminated by both acoustic noises and RF noises (we use the term "noise" to represent both thermal noise and interferences coming from other systems, but not including interferences coming from other audio channels in the same MCWM system). Thirdly, some WSDs have only one receiver branch and may monitor the FM signal only on one audio channel. As a result, RF signals on different audio channels may not be observed simultaneously.

An intuitive idea to solve these challenges is to check the cross-correlation between a demodulated FM signal and the audio signal acquired simultaneously. Since a WM user's speech is uncorrelated with noises and other users' speeches, interferences from other channels and noises can be resisted by a cross-correlator effectively. However, two issues remain: (1) this solution requires a FM demodulator which only works in high signal-to-noise ratio (SNR) conditions; (2) a cross-correlator conducts massive multiplications and has very high computation complexity. These issues are tackled by a major contribution in this work: a 1.5-bit FM receiver, which maps the FM signal to a piece of acoustic signal whose amplitude is represented by 0, 1 or -1. This is not only the first 1.5-bit FM receiver, but also the first FM receiver that works effectively when SNR is as low as -3 dB. This novelty not only lowers the complexity and SNR requirement of a FM demodulation, but also significantly reduces the complexity of a cross-correlator, since massive multiplications are eliminated by the simple coefficients 0 and ±1. The 1.5-bit FM receiver results in a cross-correlator with three-level quantization, which is the optimal quantization that processes the least information with the given quantization error [7].

We evaluate the performance of the proposed 1.5-bit FM demodulator by simulations, and evaluate the performance of the whole detection scheme in a real-world testing environment, which includes an off-the-shelf MCWM system and a WSD prototype. Based on the waveforms acquired in this real-world testing environment, we derive the detection rate β and false alarm rate α of the proposed detection scheme. Experiment results show that, the proposed scheme requires only -3 to 0 dB SNR when two audio channels are used, and requires about 5-6 dB SNR when four audio channels are used, with the performance that $\beta > 0.9$ and $\alpha < 0.1$. The detection time is as low as a quarter second.

Our contributions are summarized as follows:

- We propose a cross-correlation based WMUE attack detection scheme with the ability to resist noises and interferences in MCWM systems;
- We propose the first 1.5-bit FM demodulator which enjoys low complexity and simplifies the cross-correlator, and evaluate its performance by both theoretical analysis and computer based simulations;
- We design a hardware based prototype and validate the performance of the proposed detection scheme in a real-world environment.

Throughout the paper, "acoustical signal" and "audio signal" are synonymous. We use the terms "wireless channel" and "acoustic channel" to represent the channels experienced by RF signal and sound, respectively. All SNR's in this work are measured over the effective bandwidth of a FM signal which is at the level of 50 KHz, while those in some other works are measured over the entire 6MHz TV band [3, 8, 20]. The -3 dB SNR in this work is equivalent to -23.4 dB in those works, and is close to the limitation of those FM signal detection schemes.

2. Related Works

Various methods are proposed to detect PUE attacks. Among them, localization based methods draw much attention, with the basic principle that the location of some incumbent signal transmitters, for example, the TV towers, are preknown and hard to be emulated. By localizing the transmitter using received signal strength (RSS), one can differentiate between legitimate users and PUE attackers [4, 22]. Alternatively, PUE attacks may be detected through the fact that, the channel characteristics at different users are different and hard to be altered [6, 14]. Although this method is able to differentiate between different users, it cannot tell which user is the attacker. In other words, additional information about the legitimate user, like location or channel state information (CSI), are also required. All these methods cannot detect WMUE attacks, since both the locations and CSIs of MCWM users are hard to acquire.

The algorithm proposed in [5] detects the WMUE attacks by correlating the acoustic signal with the RF signal acquired simultaneously, and this principle is also adopted in this work. However, the work in [5] only considers the single-channel WM system, while this work covers both single-channel and multi-channel cases.

Authors in [2] propose a cooperative spectrum sensing scheme that maximizes the detection rate when PUE attacks exist. Moreover, a frequency hopping strategy is proposed in [13] to combat with PUE attacks under a game-theoretic model. These works are devoted to alleviating PUE attacks, rather than detecting PUE attacks.

Several all-digital FM receivers are proposed in [11, 15, 19], and all of them ignore noises. A FM receiver that works when SNR is as low as -3 dB is not found in the literature. Moreover, no 1.5-bit FM demodulator is found in the existing literature to detect WMUE attacks. In this work, a low-precision FM demodulator can significantly reduce the computation complexity, and is studied for the first time.

The design of multiplierless cross-correlators is discussed in [21], while design considerations and performance evaluation for the complex cross-correlator with three-level quantization are presented in [7]. These works guide us to the idea of 1.5-bit data precision; however, the main focus of this work is to detect PUE attacks, while the multiplierless cross-correlators is only part of the whole scheme.

Some preliminary results of this work are presented in [18]. In this paper, we add more technical details and evaluate the performance of the proposed FM demodulator in noisy environments through both theoretical analysis and simulations. Moreover, detecting threshold of the proposed WMUE attack detector is also discussed.

3. System Model

3.1. System Setup

A MCWM system is surrounded by a set of WSDs, as shown in figure 1. This MCWM system is composed of M audio transmitters (WMs) where $M \geq 1$, one MCWM receiver and one loudspeaker. The audio signals acquired by different WMs are modulated on different wireless channels, and are all received by the MCWM receiver and mixed together. We denote the audio signal and FM signal at the m^{th} WM as $a_m(t)$ and $s_m(t)$, respectively. Then the audio signal output $a^T(t)$ at the MCWM receiver equals to $\sum_{m=1}^{M} a_m(t)$, which is further amplified by the loudspeaker and overcast all acoustic signals generated by WM users. The WSD is able to acquire (some of) the FM signals $s_m(t)$, as well as acoustic signal $a(t)$ which contains $a^T(t)$, its reverberations and acoustic noises. The central frequency of $s_m(t)$ is denoted as f_m.

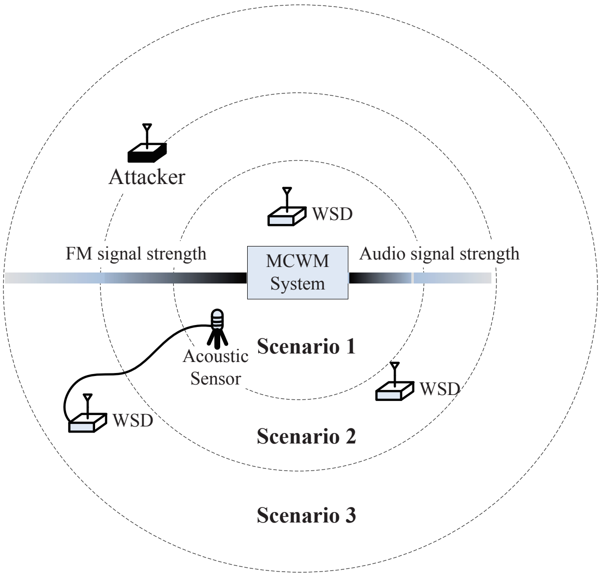

Figure 1. The system model and three scenarios considered in this paper. Scenarios differ from each other in the qualities of FM signals and acoustic signals.

According to [3], the FM signal can be modelled by

$$s_m(t) = A_C \cos\left[2\pi f_m t + 2\pi\Delta f \int_0^t a_m(t)dt + \theta\right] \quad (1)$$

where A_C and Δf control the amplitude and bandwidth of this FM signal, respectively, and θ represents a random phase with uniform distribution over $[0, 2\pi]$.

We consider that the quality of acoustic signal $a(t)$ drops much faster than the qualities of $s_m(t)$, when the propagation distance d increases. The reasons are twofold. Firstly, acoustic signals are more easily being blocked by obstacles like buildings, compared with FM signals operating on very high frequency (VHF) and ultra high frequency (UHF) bands. According to the measurement results in [9], FM signals may have more than 30 dB SNR when $d = 500$m, while effective ranges of the acoustic signals from most MCWM systems are less than 100m. Secondly, the sources of RF interference are much less than the sources of acoustic interferences, since different wireless systems operate on different frequency bands, while many types of acoustic interferences collide with human speeches in both time-domain and frequency-domain.

According to the propagation models above, we define three operating scenarios:

- *Scenario 1*: $d < 20m$, so both $s_m(t)$ and $a(t)$ are noise-free;
- *Scenario 2*: $20m < d < 200m$, so $s_m(t)$ is noise-free, but $a(t)$ is noisy;
- *Scenario 3*: $d > 200m$, so $s_m(t)$ is noisy, but high-quality $a(t)$ is acquired by the sensor close to the

MCWM system and sent to the WSD through infrastructure.

These three scenarios are also illustrated in figure 1. Our proposed WMUE attack detection algorithm covers all these three scenarios. For each scenario, we will focus on one WSD in the following analysis.

We assume that the power of $s_m(t)$ is above the noise floor at each WSD in all scenarios, so that f_m can be estimated by the WSD [8]. Since f_m can only be a multiple of 25 kHz [10], the WSD is able to adjust its estimates on f_m according to this rule. As a result, the WSD knows exact values of f_m for $m = 1, ..., M$.

3.2. Attacker Model

An attacker emulates the MCWM system by transmitting FM signals on one or multiple channels used by the legitimate MCWM system. These emulated FM signals and the FM signals transmitted by WMs are indistinguishable in terms of the modulation scheme and transmission power.

The attacker is not willing to convert the demodulated FM signal to audio signal and send it to the loudspeaker, since such audio signal would be very strange and expose the attacker directly, unless the original data transmitted by the attacker is just a piece of analogue audio signal (the exceptional case). Therefore, we consider that the attacker does not generate any audio signal, or generates audio signal that is not correlated to the FM signal. We consider that the "attacker" in the exceptional case is actually a legitimate WM system which may use the spectrum legally. We assume that the attacker has the ability to sense the spectrum and avoids collisions with existing MCWM systems. Therefore, there is one and only one source of $s_m(t)$.

3.3. The Detection Problem

The detection problem we study here is defined as the task to identify the source (either the MCWM system or the attacker) of $s_m(t)$, given a set of $a(t)$ and $s_m(t)$. It can be modelled as a hypothesis test:

- H_0: $s_m(t)$ is generated by the MCWM system;
- H_1: $s_m(t)$ is generated by the WMUE attacker.

H_0 and H_1 are called null hypothesis and alternative hypothesis, respectively.

4. The WMUE Attack Detection Scheme

The proposed WMUE attack detection scheme is based on the principle that, the acoustic signal and FM signals coming from the MCWM system correlate to each other, while those coming from the WMUE attacker do not. Then by evaluating the cross-correlation between the demodulated FM signal on a specific wireless channel and the acoustic signal, one can distinguish between a MCWM user and a WMUE attacker.

Basic procedures of the proposed scheme are shown in figure 2. The WSD first searches any FM-like signal on the frequency band interested. Once detecting a signal, it records the RF signal $s_m(t)$ and acoustic signal $a(t)$ simultaneously. Then it down-converts $s_m(t)$ to an intermediate frequency (IF) signal $s_m^{(IF)}(t)$, and feeds the latter one into a low-complexity FM demodulator. In other words, a superheterodyne receiver is considered here. Finally, the scheme computes the peak value X of the cross-correlation between the demodulated signal Y_n and the down-sampled acoustic signal A_n. X is close to 1 if $s_m(t)$ is transmitted from the MCWM system, and close to 0 if not. The same operations are repeated for other channels interested.

The cross-correlator suffers from very high computation complexity. To solve this problem, we first notice that reducing the data precision reduces the complexity of cross-correlator dramatically, but only degrades the performance slightly [7]. Consider the operation $\sum_n Y_n A_n$ required in the cross-correlator shown in figure 2; if Y_n equals to either 1 or -1, all the multiplications are unnecessary. Moreover, if $Y_n = 0$ at some points, the number of additions is also reduced. Motivated by these facts, we represent Y_n by only 0, 1 and -1. In other words, we propose a 1.5-bit FM demodulator with input $s_m^{(IF)}(t)$ and output Y_n, and show the relationship between the original audio signal $a_m(t)$ and the desired output Y_n in figure 3. This simplified FM demodulator in turn simplifies the cross-correlator significantly.

We introduce the 1.5-bit FM demodulator in subsection 4.2, and discuss its performance in noisy environments in subsection 4.3. Audio signal processing and the cross-correlator are introduced in subsections 4.4 and 4.5, respectively. Finally, the WMUE attack detector is given in subsection 4.6.

4.1. Preliminaries

We first analyse the properties of IF signal $s_m^{(IF)}(t)$ with central frequency f_I:

$$s_m^{(IF)}(t) = A_C \cos\left[2\pi f_I t + 2\pi\Delta f \int_0^t a_m(t)dt + \theta\right]. \quad (2)$$

For most superheterodyne receivers,

$$f_I > 2f_{max} \quad (3)$$

and

$$f_I > 2\Delta f\, a_{max} \quad (4)$$

where f_{max} and a_{max} denote the maximum frequency and maximum amplitude of $a_m(t)$, respectively. Then

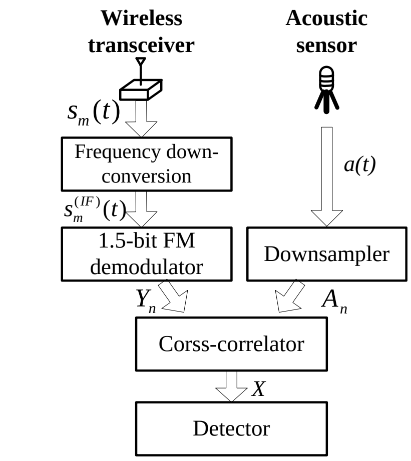

Figure 2. Basic procedures of the proposed WMUE attack detection scheme.

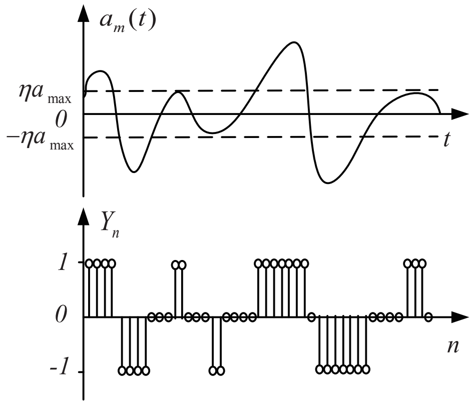

Figure 3. The relationship between $a_m(t)$ and the desired output Y_n of a 1.5-bit FM demodulator, where $-\eta$ and η are two decision thresholds.

we define T as a number that satisfies

$$2f_{max} \leq 1/T < f_I \tag{5}$$

and get the following observations.

Observation 1:

$$|\int_{nT}^{(n+1)T} s_m^{(IF)}(t)e^{j2\pi g_k t}dt\ | \approx A_C T\text{sinc}((g_k - f_{t_0})T) \tag{6}$$

where

$$f_{t_0} := f_I + \Delta f a_m(t_0) \tag{7}$$

$$|g_k - f_{t_0}| \leq 1/(2T) \tag{8}$$

and $\text{sinc}(x) := \sin(\pi x)/(\pi x)$.

Proof. Define $\theta_1 := 2\pi\Delta f \int_0^{t_0} a_m(t)dt$ and $\theta_2 := 2\pi\Delta f a_m(t_0)t_0$. From (2) we have

$$\begin{aligned} s_m^{(IF)}(t) &= A_C \cos\left[2\pi f_I t + 2\pi\Delta f \int_{t_0}^{t} a_m(t)dt + \theta_1 + \theta\right] \\ &\approx A_C \cos\left[2\pi f_{t_0} t + \theta'\right] \end{aligned} \tag{9}$$

where $\theta' := \theta_1 - \theta_2 + \theta$.

The approximation in (9) is due to the reason that, $1/T \geq 2f_{max}$ according to (5), $a_m(t)$ shows limited change during $[t_0, t]$, and $\int_{t_0}^{t} a_m(t)dt \approx a_m(t_0)(t - t_0)$. Then we get

$$\begin{aligned} &\int_{t_0}^{t_0+T} s_m^{(IF)}(t)e^{j2\pi g t}dt \\ &\approx \int_{t_0}^{t_0+T} A_C \cos\left[2\pi f_{t_0} t + \theta'\right] e^{j2\pi g t}dt \\ &= \frac{A_C}{2}[\underbrace{\int_{t_0}^{t_0+T} e^{j2\pi(g+f_{t_0})t+j\theta'}dt}_{d_1} + \underbrace{\int_{t_0}^{t_0+T} e^{j2\pi(g-f_{t_0})t-j\theta'}dt}_{d_2}] \end{aligned} \tag{10}$$

The integrands in d_1 and d_2 are two periodical functions with frequencies $g + f_{t_0}$ and $g - f_{t_0}$. According to (5) (7) and (8), $g + f_{t_0} \approx 2f_I > 2/T \geq 4|g - f_{t_0}|$. As a result, d_1 is much smaller than d_2. By ignoring d_1, (10) becomes

$$\begin{aligned} &\int_{t_0}^{t_0+T} s_m^{(IF)}(t)e^{j2\pi g t}dt \\ &\approx \frac{A_C}{2} \int_{t_0}^{t_0+T} e^{j2\pi(g-f_{t_0})t-j\theta'}dt \\ &= \frac{A_C}{2} e^{-j\theta''}(e^{j2\pi(g-f_{t_0})T} - 1)/(j2\pi(g - f_{t_0})) \end{aligned} \tag{11}$$

where $\theta'' := \theta' - 2\pi(g - f_{t_0})t_0$.

Then from (11) we have

$$\begin{aligned} &|\int_{nT}^{(n+1)T} s_m^{(IF)}(t)e^{j2\pi g_k t}dt| \\ &\approx |\frac{A_C}{2} e^{-j\theta''}(e^{j2\pi(g-f_{t_0})T} - 1)/j2\pi(g - f_{t_0})| \\ &= A_C(1 - \cos(2\pi T(g - f_{t_0})))/(4\pi^2(g - f_{t_0})^2) \\ &= A_C T\text{sinc}((g_k - f_{t_0})T) \end{aligned} \tag{12}$$

□

Observation 2: If $|g_1 - f_{nT}| \leq |g_2 - f_{nT}| \leq 1/(2T)$, $S_{m,n}^{(1)} \geq S_{m,n}^{(2)}$ where $S_{m,n}^{(k)} := |\int_{nT}^{(n+1)T} s_m^{(IF)}(t)e^{j2\pi g_k t}dt|$ and $f_{nT} := f_I + \Delta f a_m(nT)$.

Proof. It is easily shown that $S_{m,n}^k$ equals to the left part of (6) when $t_0 = nT$. According to *Observation* 1, $S_{m,n}^{(k)}$ is a monotonically decreasing function with respective to $|g - f_{nT}|$ during $[0, 1/(2T)]$. Therefore, *Observation* 2 holds. □

Since we focus on the m^{th} audio channel here, we drop the index m in $S_{m,n}^{(k)}$ if doing this would not cause misunderstanding.

4.2. The 1.5-bit FM Demodulator

Definition: A demodulator with output $Y_{m,n}$ is the 1.5-bit FM demodulator of the IF signal $s_m^{(IF)}(t)$ defined in (2) if and only if

$$Y_{m,n} = \begin{cases} -1, & a_m(nT) < -\eta a_{\max} \\ 0, & -\eta a_{\max} \leq a_m(nT) < \eta a_{\max} \\ 1, & others \end{cases} \tag{13}$$

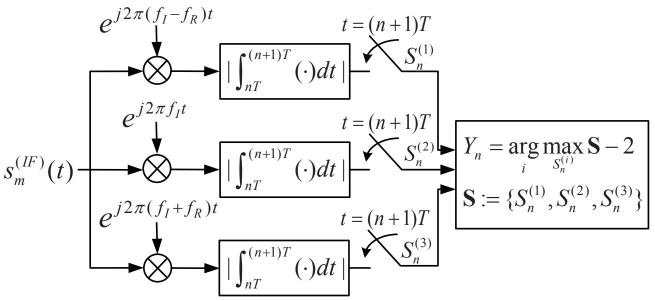
$e^{j2\pi(f_I-f_R)t}$
$e^{j2\pi f_I t}$
$e^{j2\pi(f_I+f_R)t}$
$s_m^{(IF)}(t)$
$|\int_{nT}^{(n+1)T}(\cdot)dt|$
$t=(n+1)T$
$S_n^{(1)}$
$S_n^{(2)}$
$S_n^{(3)}$
$Y_n = \underset{i \quad S_n^{(i)}}{\arg\max}\, \mathbf{S} - 2$
$\mathbf{S} := \{S_n^{(1)}, S_n^{(2)}, S_n^{(3)}\}$

Figure 4. The proposed 1.5-bit FM demodulator.

where $n = 0, 1, ...,$ while $-\eta$ and η are two decision thresholds with $0 < \eta < 1$.

Figure 3 shows the relationship between $a_m(t)$ and the desired output of this 1.5-bit FM demodulator. The thresholds $-\eta$ and η should guarantee that $Y_{m,n}$ equals to 0, 1 or -1 with equal probabilities, so that the amount of information contained in $Y_{m,n}$ is maximized. For example, if the amplitude of $a_m(t)$ is evenly distributed over $[0, a_{max}]$, $\eta = 0.5$.

Proposision 1: The demodulator shown in figure 4 with output $Y_n = \underset{i \quad S_n^{(i)}}{\arg\max}\, \mathbf{S_n} - 2$ is the 1.5-bit FM demodulator defined in *Definition*, where $\mathbf{S_n} := \{S_n^{(1)}, S_n^{(2)}, S_n^{(3)}\}$, $g_1 = f_I - f_R$, $g_2 = f_I$, $g_3 = f_I + f_R$, $f_R = 2\eta a_{\max}\Delta f$, and

$$|g_k - f_{nT}| \leq 1/(2T) \tag{14}$$

where $k = 1, 2, 3$.

Proof. When $a_m(nT) < -\eta a_{\max}$,

$$\begin{aligned} & |g_1 - f_{nT}| \\ = & |f_R + \Delta f a_m(nT)| \\ = & |2\eta a_{\max}\Delta f + \Delta f a_m(nT)| \\ < & |\Delta f a_m(nT)| \\ = & |g_2 - f_{nT}| \\ \leq & 1/(2T) \end{aligned} \tag{15}$$

and it is easily shown that $|g_1 - f_{nT}| < |g_3 - f_{nT}| \leq 1/(2T)$. Then according to *Observation* 2, $S_n^{(1)} > S_n^{(2)}$ and $S_n^{(1)} > S_n^{(3)}$. As a result, $\underset{i \quad S_n^{(i)}}{\arg\max}\, \mathbf{S_n} = 1$ and $Y_n = -1$.

By the same way, one can verify that $Y_n = 0$ when $-\eta a_{\max} \leq a_m(nT) < \eta a_{\max}$, and $Y_n = 1$ when $a_m(nT) \geq \eta a_{\max}$. □

The 1.5-bit FM demodulator proposed in figure 4 borrows the design of matched-filter [23]; however, their basic principles are different. In our system, the local signals fed into the multipliers, $e^{j2\pi f_I t}$ and $e^{j2\pi(f_I \pm f_R)t}$, do not necessarily match any pieces of the FM signal transmitted. Instead, our system is designed such that the integrators generate larger outputs when these local signals match the present signal better. Then by searching the largest outputs from three integrators, the demodulator determines the best value (0, 1 or -1) for Y_n.

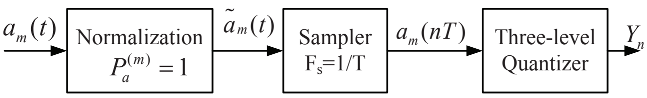

Figure 5. The proposed 1.5-bit FM demodulator can be interpreted as a sampler for the audio signal $a_m(t)$ with sampling frequency $F_s = 1/T$ followed by a three-level quantizer.

This 1.5-bit FM demodulator can also be interpreted as a sampler for the normalized audio signal $\widetilde{a}_m(t)$ with sampling frequency $F_s = 1/T$ followed by a three-level quantizer, as shown in figure 5. After normalization, we assume that the average power of $\tilde{a}_{\max}(t)$ equals to 1. This FM demodulator may also operate in digital domain, if the input $s_m^{IF}(t)$ is sampled with sampling rate G_s. It is easily shown that, basic principle of this demodulator still holds in digital domain, while the only change is to replace the integrators in figure 4 by adders.

By adopting integrators, the proposed 1.5-bit FM demodulator is able to work in low-SNR environments with low computation complexity. On the other hand, conventional digital FM demodulators [11, 15, 19] either suffer from high complexity or require high SNR.

4.3. Performance of the FM demodulator in Noisy Environments

We define the noisy IF signal fed into the FM demodulator as $\widehat{s}_m^{(IF)}(t)$, which is modelled as

$$\widehat{s}_m^{(IF)}(t) = s_m^{(IF)}(t) + w_m(t) \tag{16}$$

where $w_m(t)$ is the additive white Gaussian noise (AWGN) with power σ_m^2 at the m^{th} wireless channel. Then according to figure 4, output of the i^{th} integrator is given by

$$\begin{aligned} S_n^{(i)} &= |\int_{nT}^{(n+1)T} (s_m^{(IF)}(t) + w_m(t))e^{j2\pi g_i t}dt| \\ &= |\rho_n^{(i)}e^{j\phi_i} + \int_{nT}^{(n+1)T} w_m(t)e^{j2\pi g_i t}dt| \end{aligned} \tag{17}$$

where

$$\rho_n^{(i)} := A_C T\text{sinc}((g_k - f_{t_{nT}})T) \tag{18}$$

according to *Observation* 1, and ϕ_i denotes the angle of the complex value in (11) when $g = g_i$.

Let $N_m := \int_{t_0}^{t_0+T} w_m(t)e^{j2\pi g_i t}dt$ represent the random part in (17). One can easily verify that N_m is a random variable which follows complex normal distribution with mean value 0 and variance $\sigma_m^2 T$. As a result, $S_n^{(i)}$ in (17) follows Rician distribution whose probability

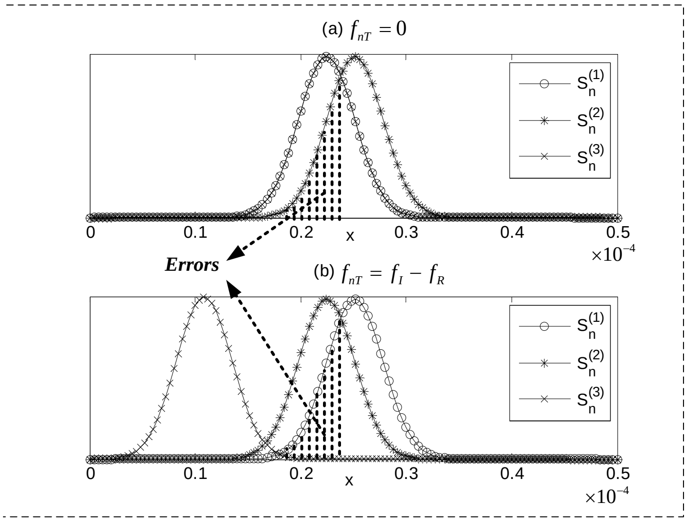

Figure 6. The PDF curves for $S_n^{(1)}$, $S_n^{(2)}$ and $S_n^{(3)}$ with (a) $f_{nT} = f_I$ and (b) $f_{nT} = f_I - f_R$, respectively, and $\gamma = 0$ dB.

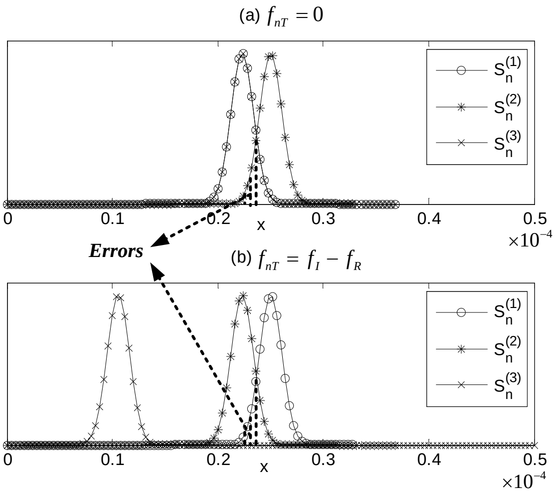

Figure 7. The PDF curves for $S_n^{(1)}$, $S_n^{(2)}$ and $S_n^{(3)}$ with (a) $f_{nT} = f_I$ and (b) $f_{nT} = f_I - f_R$, respectively, and $\gamma = 8$ dB.

density function (PDF) is given by [1]

$$p_i(x) = \begin{cases} \frac{x}{\sigma_m^2 T} e^{-\frac{x^2+(\rho_n^{(i)})^2}{2\sigma_m^2 T}} I_0(\frac{x\rho_n^{(i)}}{\sigma_m^2 T}),\ x \geq 0 \\ 0,\ x < 0 \end{cases} \tag{19}$$

where $\rho_n^{(i)}$ follows the same definition as in (18), and $I_0(x)$ is the modified Bessel function of the first kind with order zero.

Similar to the bit-error-rate (BER) performance in digital communication systems [1], performance of the proposed 1.5-bit FM demodulator is closely related to the ratio γ defined by $\gamma := {\rho_n}^2/(\sigma_m^2 T)$, which is proportional to ${A_C}^2 T/\sigma_m^2$ according to (18). Note that ${A_C}^2/\sigma_m^2$ just equals to (twice of) the SNR of $s_m(t)$. As a result, larger SNR leads to larger γ and better anti-noise ability.

Larger T also leads to larger γ. However, T is also restricted by (5) and (14), while larger T makes both *Observation* 1 and *Observation* 2 less accurate and may increase demodulation error (we call it modelling error). In practice, the optimal value of T may not strictly satisfy both (5) and (14) due to the trade-off between γ and modelling error. We will derive the optimal value of T by simulations in subsection 5.1.

PDF's of $S_n^{(1)}$, $S_n^{(2)}$ and $S_n^{(3)}$ are plotted in figure 6, with $A_C = 1$, $f_I = 50$ kHz, $f_R = 5$ kHz, $T = 50\mu s$ and $\gamma = 0$ dB. When $f_{nT} = f_I$ as shown in figure 6.a, $S_n^{(2)}$ has the best chance to be the largest one among $\{S_n^{(1)}, S_n^{(2)}, S_n^{(3)}\}$; as a result, Y_n tends to be 0, which is correct. When $f_{nT} = f_I - f_R$ as shown in figure 6.b, $S_n^{(2)}$ is most likely the largest one, and Y_n tends to be -1. The shaded area denotes the chance of decoding error. When γ is changed to 8 dB and other conditions keep unchanged, as shown in figure 7, the shaded areas become smaller compared with those in figure 6, and the demodulation performance is better.

Although a closed-form expression on the "BER" of this demodulator can be derived from PDF's, we note that this expression is valid only when this demodulator operates in analog domain. Performance of this FM demodulator operating in digital domain is affected by the sampling rate F_s, and we will show the normalized mean square error (NMSE) performance of this demodulator under the sampling rate adopted by the real-world experiments in subsection 5.1.

4.4. Audio Signal Processing

We model the acoustic signal $a(t)$ arriving at the WSD under H_0 as

$$a(t)|H_0 = a^{(T)}(t) \otimes h(t) = \sum_{j=1}^{J} h_j a^{(T)}(t - t_j) + z(t) \tag{20}$$

where $h(t) := \sum_{j=1}^{J} h_j \delta(t - t_j)$ represents the impulse response of the acoustic channel between loudspeaker and WSD, and $z(t)$ denotes audio noises.

In practice, acoustic signal travels slower than RF signal. To address this issue, we define the time $t = 0$ as the time when FM signal is detected by the WSD. Accordingly, all t_j's in the acoustic channel model $h(t)$ incorporate propagation delay of the acoustic signal. As an example, if there is line-of-sight with distance D between audio amplifier and WSD, $t_0 = D/v$ where v denotes the speed of sound in the air, while propagation delay of FM signal is much smaller than t_0 and has been ignored.

At the WSD side, $a(t)$ is sampled by the acoustic sensor at a high sampling rate (for example, 44.1 kHz). In order to match this acoustic signal with the FM demodulator output, we resample this acoustic signal at the rate $1/T$, which equals to the sampling rate of Y_n. Since $1/T$ =10 kHz is good enough to capture human voices, we consider this operation as a downsampler as shown in figure 2. Moreover, this downsampler features a lowpass filter with stop-band $1/T$ to resist out-of-band noises.

Denote the downsampled acoustic signal as A_n, which is obtained by

$$A_n|H_0 = a^{(T)}(t) \otimes h(t) \otimes h_s(t) + z_L(t) \tag{21}$$

where $h_s(t) := \sum \delta(t - nT)$ serves as the sampling function, and $z_L(t)$ denotes the lowpass-filtered noises. We combine the sampling operation with the acoustic channel response, and define

$$d(nT) := h(t) \otimes h_s(t) := \sum_{l=1}^{L} d_l \delta(nT - t_l T) + Z_n \tag{22}$$

where t_l is a non-negative integer, and Z_n denotes the samples of noises. Combining (21) and (22), we get

$$\begin{aligned} A_n|H_0 &= \sum_{m=1}^{M} a_m(t) \otimes \sum_{l=1}^{L} d_l \delta(nT - t_l T) \\ &= \sum_{m=1}^{M} \sum_{l=1}^{L} d_l a_m(nT - t_l T) + Z_n. \end{aligned} \tag{23}$$

On the flip side, A_n under H_1 is modelled by

$$A_n|H_1 = Z_n^{'} \tag{24}$$

which incorporates both audio noise and possible audio signal generated by the attacker. We denote the average powers of Z_n and $Z_n^{'}$ as P_Z and $P_Z^{'}$, respectively. When the attacker emulates both audio signal and FM signals, $P_Z^{'} > P_Z$, whereas $P_Z^{'} = P_Z$ if the attacker only emulates FM signals.

4.5. The Cross-correlator

In this subsection, we will set up the connection between the audio samples A_n and the FM demodulator outputs Y_n under three scenarios defined in subsection 3.1.

Scenario 1. We first look at the simplest scenario (scenario 1) in which both audio noises and RF noises are ignored.

According to (13) and (23), both A_n and Y_n are functions of $a_m(t)$ under H_0. Moreover, the relationship between Y_n and $a_m(t)$ can be simplified by the interpretation given in figure 5:

$$Y_{m,n} = \frac{a_m(nT)}{\sqrt{P_a^{(m)}}} + Q_{m,n} \tag{25}$$

where $P_a^{(m)}$ denotes the average power of $a_m(t)$, and $Q_{m,n}$ denotes the quantization error at $t = nT$. Then from (23) and (25), we get

$$\begin{aligned} &Corr(A_n|H_0, Y_{m,n}, p) \\ &:= \tfrac{1}{P_{m,p}} \sum_{n=0}^{W-1} (A_n|H_0) Y_{m,n-p} \\ &= \begin{cases} \tfrac{1}{P_{m,p}} (C_{m,p}^{(0)} + C_{m,p}^{(1)} + C_{m,p}^{(2)}), p = t_l \\ \quad \tfrac{1}{P_{m,p}} (C_{m,p}^{(2)} + C_{m,p}^{(3)}), others \end{cases} \end{aligned} \tag{26}$$

where

$$P_{m,p} := \sqrt{(\sum_{n=0}^{W-1} (A_n)^2)(\sum_{n=0}^{W-1} (Y_{m,n-p})^2)} \tag{27}$$

$$C_{m,p}^{(0)} := \frac{d_l}{\sqrt{P_a^{(m)}}} \sum_{n=0}^{W-1} |a_m(n - t_l)|^2 \tag{28}$$

$$C_{m,p}^{(1)} := \frac{1}{\sqrt{P_a^{(m)}}} \sum_{n=0}^{W-1} \sum_{\substack{m',l', \\ |m-m'|+|l-l'|\neq 0}} a_{m'}(n - t_{l'}) a_m(n - t_l) \tag{29}$$

$$C_{m,p}^{(2)} := \sum_{n=0}^{W-1} \sum_{m=1}^{M} \sum_{l=1}^{L} d_l a_m(n - t_l) Q_{m,n} + \sum_{n=0}^{W-1} Z_n Y_{n-p} \tag{30}$$

$$C_{m,p}^{(3)} := \frac{1}{\sqrt{P_a^{(m)}}} \sum_{n=0}^{W-1} \sum_{m'=1}^{M} \sum_{l'=1}^{L} d_l a_{m'}(n - t_{l'}) a_m(n - p). \tag{31}$$

and W determines the window size of this cross-correlator.

Similarly, $Corr(A_n|H_1, Y_{m,n}, p)$ is obtained by setting $a_m(t) = 0$ for $m = 1, ..., M$ in (26):

$$Corr(A_n|H_1, Y_{m,n}, p) = \frac{1}{P_{m,p}} \sum_{n=0}^{W-1} Z_n^{'} Y_{m,n-p}. \tag{32}$$

The audio noises Z_n and $Z_n^{'}$ and quantification error Q_n are considered as uncorrelated to Y_n and $a_m(t)$, respectively. As a result, $Corr(A_n|H_1, Y_{m,n}, p)$ is close to 0. On the flip side, due to the existence of $C_{m,p}^{(0)}$ given in (28), $Corr(A_n|H_0, Y_{m,n}, p)$ always contains some values that are much larger than 0 (but smaller than 1). If audio signals $a_m(t)$ on different channels are correlated with each other, $Corr(A_n|H_0, Y_{m,n}, p)$ is even larger because of $C_{m,p}^{(1)}$ given in (29). In any case, $Corr(A_n|H_0, Y_{m,n}, p)$ is expected to exceed $Corr(A_n|H_1, Y_{m,n}, p)$ when $p = t_l$.

Finally, we design the output X of the cross-correlator as

$$X = \max_{p=0,...,\tau_{\max}} \{Corr(A_n, Y_{m,n}, p)\} \tag{33}$$

where $\tau_{\max}$ represents the maximum delay spread of the audio channel divided by T (and rounded to the nearest integer if necessary). Equation (33) searches the peak value X of the cross-correlation between demodulated FM signal and down-sampled audio signal within the time window $[0, \tau_{\max}]$, and $X|H_0$ is expected to exceed $X|H_1$. This searching process synchronizes the demodulated FM signal $Y_{m,n}$ with the strongest (sampled) path in A_n.

Scenario 2 and Scenario 3. *Scenario 2* differs from *Scenario 1* only in that, the audio signal $a(t)$ has poor quality, or in other words, Z_n has larger amplitude. As a result, all the analysis in *Scenario 1* directly applies to *Scenario 2*.

Scenario 3 differs from *Scenario 1* only in that, $s_m(t)$ has poor quality. As a result, $Y_{m,n}$ is contaminated by both quantification error and noises. For simplicity, we merge the the quantification error into noises, and let $Q_{m,n}$ represent both. As a result, all the analysis in *Scenario 1* still applies to *Scenario 3*. The value of $Corr(A_n|H_1, Y_{m,n}, p)$ increases when the amplitude of Z_n or $Q_{m,n}$ increases. As a result, the window size W of the cross-correlator in *Scenario 2* and *Scenario 3* should be larger than the window size adopted in *Scenario 1*. However, larger window size also leads to longer detection time which equals to TW, and there is a trade-off between the detection performance and computation complexity. We will discuss this issue in section 5.

4.6. The Detector

According to the analysis in subsection 4.5, $X|H_0$ is expected to be greater than $X|H_1$ under all three scenarios. Then the proposed WMUE attack detector is given as follows:

The Detector: a WMUE attack is detected if and only if $X < X_0$, where X_0 is the detection threshold.

To determine the detection threshold X_0, we first ignore the quantification error $Q_{m,n}$, and assume that audio signals at different audio channels are uncorrelated and have same average power P_R measured at the WSD. Then the first term in (23) (the power of the mixed audio signal at the WSD) has the average power MP_R, $C^{(1)}_{m,p} = 0$ and $C^{(2)}_{m,p} = \sum_{n=0}^{W-1} Z_n Y_{n-p}$. It is easily shown that $P_{m,p} \propto W\sqrt{MP_R + P_Z}$, $C^{(0)}_{m,p} \propto W\sqrt{P_R}$, and $C^{(2)}_{m,p} \propto \sqrt{WP_Z}$, where P_Z denotes the average power of Z_n. Then we get

$$X|H_0 \propto \frac{W\sqrt{P_R} + \sqrt{WP_Z}}{W\sqrt{MP_R + P_Z}} \tag{34}$$

and

$$X|H_1 \propto 1/\sqrt{W}. \tag{35}$$

When W is large, $X|H_1$ approaches 0 and $X|H_0$ approaches $\sqrt{P_R}/\sqrt{MP_R + P_Z}$. The detection threshold X_0 equals to $\frac{1}{2}\sqrt{P_R}/\sqrt{MP_R + P_Z}$ accordingly, which only needs the second-order statistics of audio signal and background noise. X_0 may be further simplified as $1/2\sqrt{M}$ when P_Z is small compared with P_R.

When audio signals on different channels are correlated, $X|H_0$ increases while $X|H_1$ keeps unchanged. With the same detection threshold derived above, detection rate will be enhanced, while the false alarm rate is not affected.

4.7. Discussions

With the models (23) and (24), it seems that WMUE attacks can be detected simply by energy detection. However, such detection method is vulnerable if the attacker emulates audio signal in order to increase the audio noise floor. On the flip side, the proposed scheme always works as long as the emulated audio signal is uncorrelated to the FM signal.

In order to get X, $Corr(A_n, Y_{m,n}, p)$ needs to be calculated for $\tau_{max} + 1$ times with different values of p. In the definition of $Corr(A_n, Y_{m,n}, p)$ given in (26), the calculation of $\sum_{n=0}^{W-1} (A_n|H_0)Y_{m,n-p}$ requires only additions, because $Y_{m,n}$ only takes the values of 0 and ±1. Moreover, the normalization factor $\frac{1}{P_{m,p}}$ can be derived iteratively [17], and takes only one multiplication and one square root operation per update, only except for the first update (when $p = 0$). The 1.5-bit FM demodulator requires only three analogue multipliers and three integrators if operating at analogue domain, and takes three multiplications and three additions per sample if operating at digital domain. As a result, the whole detection scheme enjoys low computation complexity.

5. Experiments

Performance of the proposed WMUE attack detection scheme is determined by the performances of the 1.5-bit FM demodulator and the cross-correlator. We first conduct computer-based simulations to evaluate the NMSE performance of the FM demodulator, then prototype the whole scheme and conduct real-world experiments to evaluate detection rate and detection time.

5.1. Performance of the FM demodulator

Performance of the FM demodulator is quantified by NMSE, which is defined as

$$NMSE := \frac{\sum_{n=0}^{N_{All}-1} \left(Y^{(1)}_{m,n} - Y^{(2)}_{m,n}\right)^2}{\sum_{n=0}^{N_{All}-1} \left(Y^{(1)}_{m,n}\right)^2} \tag{36}$$

where $Y_{m,n}^{(1)}$ denotes the outputs of the ideal FM-demodulator defined in (13), $Y_{m,n}^{(2)}$ denotes the outputs from the proposed FM-demodulator shown in figure 4, and N_{All} represents the length of both outputs. In this definition, we have excluded the quantization error, since the task of this demodulator is not to recover the original audio signal $a_m(t)$, but to provide valid coefficients for the cross-correlator.

The NMSE performances are derived from computer based simulations, with $\Delta f = 5$kHz, f_I=10 kHz, $\eta = 0.5$ and $N_{All} = 500$. The audio signal $a_m(t)$ is read from an audio file which records a piece of human voice, with $a_{max} \approx 1$ (a little smaller than 1). Moreover, the IF signal $s_m^{IF}(t)$ is sampled at $G_s = 50$ kHz, which is the same sampling rate that will be used in our real-world testing, and the FM demodulator operates at digital domain.

The sampling interval T is considered as an important design parameter, since larger T leads to larger γ as discussed in subsection 4.3, but also reduces the sampling rate of $Y_{m,n}$ and may cause aliasing. Moreover, SNR of the FM signal, which equals to $A^2/(2\sigma_m^2)$, is also an important factor of NMSE.

We plot NMSEs of the FM-demodulator as a function of SNR in figure 8, when T equals to 0.1, 0.25, 0.5 and 1 ms, respectively. It is shown that the setting $T = 0.5$ms leads to the best performance in most cases, and is adopted in the following testing. The corresponding optimal sampling rate F_s^* is 2 kHz, which just satisfies (5) with the fact that the most energy in human voice concentrates in the frequency band below 1 kHz. On the flip side, F_s^* is smaller than the value required by (14) which equals to 10 kHz, since smaller sampling rate leads to longer integration window and better anti-noise ability, which compensates for the increased modelling error.

5.2. The Prototype

We prototype the proposed WMUE attack detection scheme by a commercial MCWM system and a self-designed WSD in a 12m × 7m room, as shown in figure 9. A commercial MCWM system and a WSD prototype are set up in a 12m × 7m room. The MCWM system contains an 8-channel WM receiver manufactured by Pyle Audio Inc. with model number PDWM8400, a 40W loudspeaker, and eight WMs. The carrier frequencies of these 8 channels are within the range of 170-240 MHz, which falls into VHF band. The WSD prototype is composed of two function blocks: (1) audio signal acquisition and (2) RF signal acquisition. An acoustic sensor connected to a laptop with sampling rate $F_s = 44.1$ kHz takes response of audio signal acquisition, while RF signal acquisition is realized by a multi-channel oscilloscope and two RF branches, which is capable to capture RF signals on two wireless channels simultaneously.

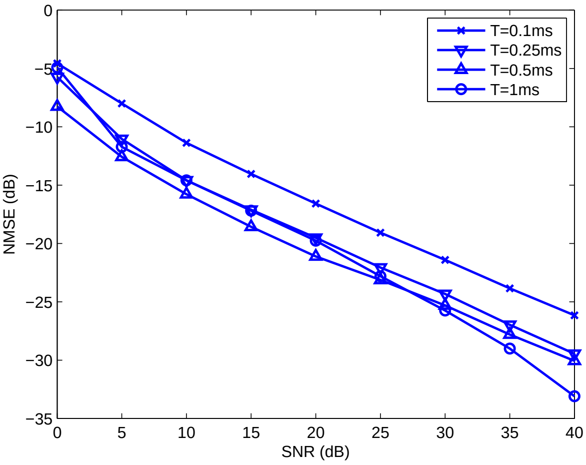

Figure 8. NMSE of the proposed 1.5-bit FM demodulator when T equals to 0.1, 0.25, 0.5 and 1 ms, respectively.

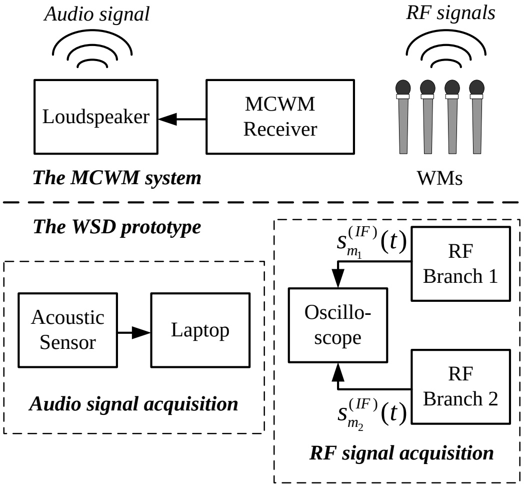

Figure 9. Block diagram of the real-world testing environment.

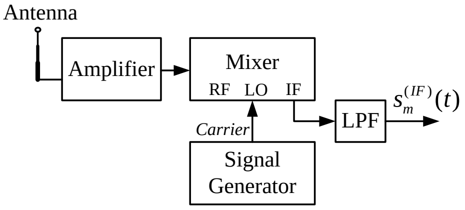

Figure 10. Block diagram of one RF branch.

The two RF branches in figure 9 share the same design as shown in figure 10, which is mainly a frequency down-conversion circuit realized by a level-7 mixer. A signal generator serves as the local oscillator. Moreover, the wireless signal is amplified by an amplifier at RF and filtered by a LPF at IF. Both the

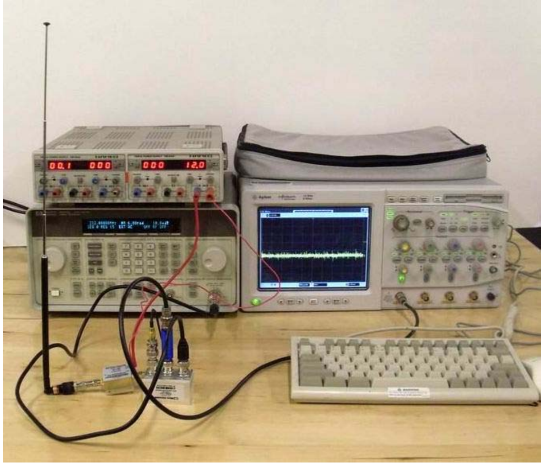

Figure 11. Photo of the WSD prototype.

image in the mixer's IF output and the out-of-band interferences are also rejected by the LPF; therefore, we do not apply any RF filter here. The IF signals $s_{m_1}^{(IF)}(t)$ and $s_{m_2}^{(IF)}(t)$ coming from two RF branches are recorded by the multi-channel oscilloscope with sampling rate $F_a = 50$ kHz and $f_I = 10$ kHz. Figure 11 shows the picture of the WSD prototype.

5.3. Testing Method

We consider such a WMUE attacker who replaces the speaker of a commercial WM system with an earphone and uses the modified system as his personal wireless phone. This attacker is very similar to a legitimate WM user and hard to be detected. As a result, the WM user is emulated by the MCWM system with loudspeaker turned on, while the WMUE attacker is emulated by the same MCWM system with loudspeaker turned off. Meanwhile, we define detection rate β as the rate that the WMUE attack is detected when the loudspeaker is turned off, and define false alarm rate α as the rate that the WMUE attack is detected when the loudspeaker is turned on.

For each scenario described in section 3.1, we test two cases that (1) two wireless channels or (2) four wireless channels are used simultaneously; we will use the terms "two channels" and "four channels" to represent these two test cases, respectively. The FM demodulator operates in digital domain with $t = n'T'$ where $T' = 20$ μs and $n' = 0, 1, \ldots$. We set $\tau_{max} = 200$, since the maximum delay spread of the acoustic channel experienced in our experiments does not exceed 0.1 s. Two RF branches are designed to emulate some WSDs with multiple antennae; the waveforms acquired by two RF branches are considered as two independent samples, upon which our detection scheme are executed twice and the results are averaged. We set $\eta = 0.5$ in both cases and all scenarios.

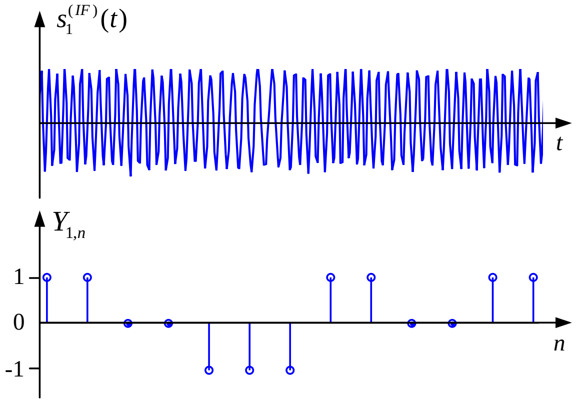

Figure 12. Waveform pieces of the IF signal and demodulator output acquired in the testing.

Performance of the proposed WMUE attack detection scheme in *Scenario 1* is evaluated by the original waveforms acquired in the experiments, with about 30 samples for each test case. For the other two scenarios, we add random noises to either the acoustic signal (in *Scenario 2*) or IF signals (in *Scenario 3*) with certain SNR.

5.4. Testing Results

A snapshot of the waveform pieces of $s_1^{(IF)}(t)$ and $Y_{1,n}$ derived in the experiments is plot in figure 12. The IF signal $s_1^{(IF)}(t)$ is close to a sine wave but with varying frequency; its amplitude is not constant due to the limited over-sampling rate (which equals to 5 according to the settings in subsection 5.2). The demodulator output $Y_{1,n}$ equals to 1 when the instant frequency of $s_1^{(IF)}(t)$ is high, while equals to -1 when the instant frequency of $s_1^{(IF)}(t)$ is low.

Next, we evaluate the relationship between detection rate β and detection time TW in three scenarios with the simpler case of two channels, as shown in figure 13. Both the SNR of the audio signal in *Scenario 2* and the SNR of the IF signals in *Scenario 3* are set to 3 dB, and the false alarm rate α in all curves are kept below 0.1. The proposed scheme achieves good performance in all scenarios when the detection time is no less than 0.25 s, or $W >= 500$. We focus on the case of $W = 500$ in the following experiments.

Finally, the performances in *Scenario 2* and *Scenario 3* under different SNR conditions are further evaluated by receiver operating characteristic (ROC), which represents detection rate β versus false alarm rate α. In *Scenario 2*, the proposed detection scheme achieves the performance $\alpha < 0.1$ and $\beta > 0.9$ (named good performance) when SNR is higher than -3 dB and 6 dB in the cases of two channels and four channels, respectively, as shown in figure 14. In *Scenario 3*, the SNRs required to achieve good performance in the two

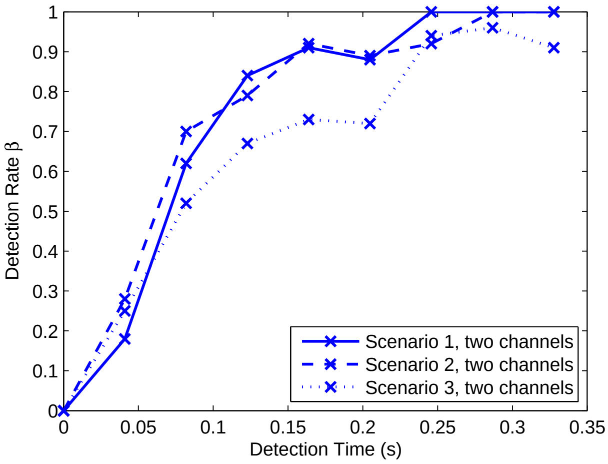

Figure 13. Detection rate versus detection time in three scenarios in the cases of two channels and four channels, respectively.

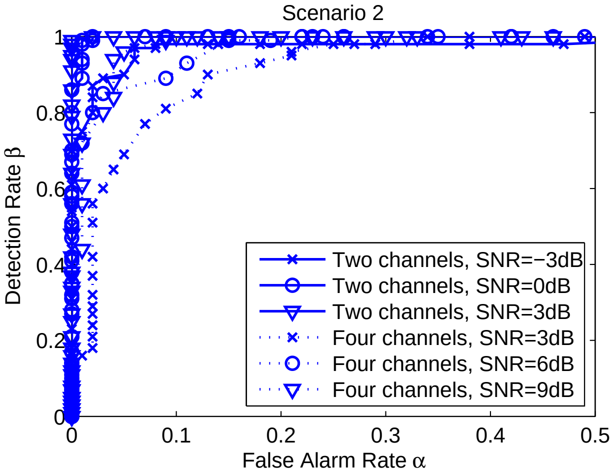

Figure 14. ROC curves in *Scenario 2* under different SNR conditions in the cases of two channels and four channels, respectively.

cases are 0 dB and 5 dB, respectively, as shown in figure 15.

These testing results validate that, the proposed scheme perform well in both noiseless environments and noisy environments.

6. Conclusions

In this paper, we propose a novel and simple algorithm to detect WMUE attacks imposed on MCWM systems in noisy environments. To the best of our knowledge, this is the first work that considers the MCWM systems. The cross-correlation between demodulated FM signal and the acoustic signal acquired simultaneously provides an effective way to detect WMUE attacks, and show good ability to resist noises/interferences. Moreover, computation complexity of the cross-correlator can be significantly reduced by the proposed 1.5-bit FM demodulator. The optimal sampling rate of the FM demodulator is 2 kHz according to the simulation results.

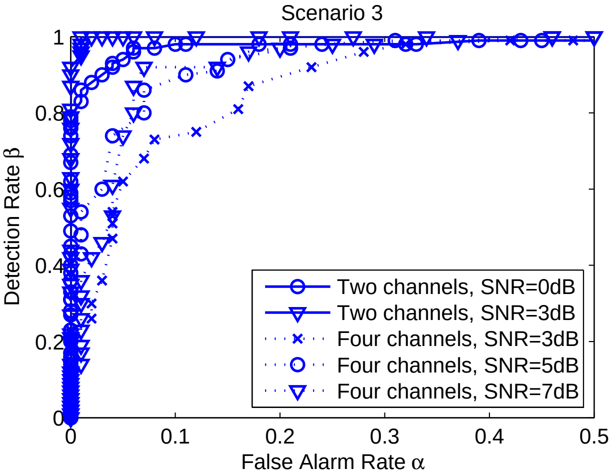

Figure 15. ROC curves in *Scenario 3* under different SNR conditions in the cases of two channels and four channels, respectively.

We set up a MCWM system and design a WSD prototype for performance evaluation. Hardware based experiments show that, the proposed algorithm is able to detect WMUE attacks within 0.25 s in all scenarios when two or four wireless channels are used simultaneously, with detection rate $\beta > 0.9$ and false alarm rate $\alpha < 0.1$. The minimum and maximum SNRs required to achieve such performance in various conditions equal to -3 dB and 6 dB, respectively.

We conclude that, both the 1.5-bit FM demodulator and the WMUE attack detection algorithm achieve good performances in noisy environments. Performance of the proposed scheme may be further enhanced by multiple antenna or collaborative sensing techniques, which are considered as our future works.

Acknowledgement. This material is based upon work partially supported by the US National Science Foundation CAREER award under Grant Number (CNS-1149500).

References

[1] *Communication Systems and Techniques*. New York: McGraw-Hill, 1966.

[2] C. Chen, H. Cheng, and Y.-D. Yao. Cooperative spectrum sensing in cognitive radio networks in the presence of the primary user emulation attack. *IEEE Transactions on Wireless Communications*, 10(7):2135–2141, 2007 July.

[3] H.-S. Chen and W. Gao. Spectrum sensing for tv white space in north america. *IEEE Journal on Selected Areas in Communications*, 29(2):316–326, Feb. 2011.

[4] R. Chen, J.-M. Park, and J. Reed. Defense against primary user emulation attacks in cognitive radio networks. *IEEE Journal on Selected Areas in Communications*, 26(1):25–37, Jan. 2008.

[5] S. Chen, K. Zeng, and P. Mohapatra. Hearing is believing: Detecting wireless microphone emulation attack in white space. *IEEE Transactions on Mobile Computing*, 12(3):401–411, 2013.

[6] Z. Chen, T. Cooklev, C. Chen, and C. P. Raez. Modeling primary user emulation attacks and defenses in cognitive radio networks. In *Proc. Performance Comput. Commun. Conf. (IPCCC)*, 2009.

[7] L. R. D'Addario, A. R. Thompson, F. R. Schwab, and J. Granlund. Complex cross correlators with three-level quantization design tolerances. *Radio Science*, 19:931–945, May-June 1984.

[8] H. S. Dhillon, J.-O. Jeong, D. Datla, M. Benonis, R. M. Buehrer, and J. H. Reed. A sub-space method to detect multiple wireless microphone signals in tv band white space. *Analog Integr Circ Sig Process*, (69):297âĂŞ306, Sep. 2011.

[9] T. Erpek, M. McHenry, and A. Stirling. Dynamic spectrum access operational parameters with wireless microphones. *IEEE Communications Magazine*, 49(3):38–45, Mar. 2011.

[10] FCC. Fm broadcast translator stations and fm broadcast booster stations, 47 cfr part 74.

[11] J. Garodnick, J. Greco, and D. Schilling. Theory of operation and design of an all-digital fm discriminator. *IEEE Transactions on Communications*, 20(6):1159– 1165, 1972 Dec.

[12] S. Kim, J. Lee, H. Wang, and D. Hong. Sensing performance of energy detector with correlated multiple antennas. *Signal Processing Letters, IEEE*, 18(8):671–674, Aug. 2009.

[13] H. Li and Z. Han. Dogfight in spectrum: Combating primary user emulation attacks in cognitive radio systems, part i: Known channel statistics. *IEEE Transactions on Wireless Communications*, 9(11):3566–3577, 2010 Nov.

[14] N. Nguyen, R. Zheng, and Z. Han. On identifying primary user emulation attacks in cognitive radio systems using nonparametric bayesian classification. *IEEE Transactions on Signal Processing*, 60(3):1432–1445, March 2012.

[15] A. Saha and B. Mazumder. A digital phase-locked loop for generating frequency discriminating codes and frequency multiplication. *Proceedings of the IEEE*, 69(4):472– 473, 1981 April.

[16] A. Sahai, N. Hoven, and R. Tandra. Some fundamental limits in cognitive radio. In *Proc Allerton Conf Commun Control Comput*, 2004.

[17] D. Schmidl, T.M.; Cox. Robust frequency and timing synchronization for ofdm. *IEEE Transactions on Communications*, 45(12):1613–1621, Dec. 1997.

[18] D. Shan, K. Zeng, P. Richardson, and W. Xiang. Detecting multi-channel wireless microphone user emulation attacks in white space with noise. In *The 8th International Conference on Cognitive Radio Oriented Wireless Networks (CROWNCOM 2013)*, 2013.

[19] B.-S. Song and I. S. Lee. A digital fm demodulator for fm, tv, and wireless. *Circuits and Systems II: Analog and Digital Signal Processing, IEEE Transactions on*, 42(12):821–825, 1995 Dec.

[20] S. Xu, S. Xu, and H. Wang. Svd based sensing of a wireless microphone signal in cognitive radio networks. In *Int. Conf. on Computational Science (ICCS)*, 2008.

[21] K.-W. Yip, Y.-C. Wu, and T.-S. Ng. Design of multiplierless correlators for timing synchronization in ieee 802.11a wireless lans. *IEEE Transactions on Consumer Electronics*, 49(1):107– 114, 2003 Feb.

[22] Z. Yuan, D. Niyato, H. Li, J. B. Song, and Z. Han. Defeating primary user emulation attacks using belief propagation in cognitive radio networks. *IEEE Journal on Selected Areas in Communications*, 30(10):1850–1860, Nov. 2012.

[23] L. Zadeh and J. Ragazzini. Optimum filters for the detection of signals in noise. *Proceedings of the IRE*, 40(10):1223–1231, Oct. 1952.