

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

Dan Shan, Kai Zeng, Paul Richardson and Weidong Xiang
ECE and CIS departments
University of Michigan - Dearborn
Dearborn, Michigan, USA 48128
Email: {danshan, kzeng, richarpc, xwd}@umd.umich.edu

Abstract—In this work, we study a special kind of primary user emulation (PUE) attack, named wireless microphone user emulation (WMUE) attack in white space cognitive radio networks. In WMUE attacks, a malicious user emulates wireless microphone (WM) signals in order to block secondary users. Existing work on WMUE attack detection deals with single channel scenario. Although multi-channel WM (MCWM) systems are common, detecting WMUE attacks under a multi-channel setting in noisy environments has not been well studied and the existing solution for single channel case cannot be directly applied. In a practical multi-channel WM system, the audio signals on different channels mix with each other and are contaminated by noises, which introduce great challenges on WMUE attack detection. We propose a novel multi-channel WMUE attack detection scheme which is based on the cross-correlation between the demodulated FM signal and the acoustic signal. The audio interferences, audio noises, and RF noises are all resisted by the cross-correlation. To reduce computation complexity, we propose a 1.5-bit FM demodulator whose outputs are represented by only 0, 1 and -1. Moreover, we set up a MCWM system and developed a hardware based prototype to evaluate the performance of the proposed scheme. Experimental results show that, the proposed scheme can effectively detect multi-channel WMUE attacks within 0.25 second with detection rate larger than 0.9 and false alarm rate lower than 0.1 under low signal-to-noise ratios.

I. 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 [1]–[3] on white space to detect the presence of PUs (incumbent signals), mainly including TV signals and wireless microphone (WM) signals. When PUs emerge, SUs are required to evacuate from the spectrum in order to avoid interference to the PUs. Exploiting this policy adversely, an attacker may block all SUs within an area by emulating the signal of a certain type of PU. This kind of attack is named primary user emulation (PUE) attack [4].

Emulating the WM signal is much easier than emulating TV signal, since the latter one is usually transmitted from a tower with preknown location. By evaluating the received signal's coverage area [4], [5] or channel characteristics [6], [7], one can differentiate between the signal from a PUE attacker

and the real TV signal. However, these detection techniques cannot be applied to detect the attack that emulates WM signals, named WM user emulation (WMUE) attack, because both locations and channel characteristics of legitimate WM systems are hard to acquire and validate. Detecting a WMUE attack is challenging, while launching a WMUE attack is as simple as building a FM modulator which is mature and cheap.

Existing work detects WMUE attacks in a single-channel system by comparing the FM signals with the acoustic signals acquired simultaneously. Although multi-channel WM (MCWM) systems are common in practice, PUE attacks in such systems have not been studied and the existing solution proposed in [8] cannot be directly applied.

Detecting WMUE attacks in a practical MCWM system in noisy environments faces new challenges. Firstly, multiple WM users in the same MCWM system may speak simultaneously; 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 are mixed together. Secondly, the audio signal and FM signals are further decorrelated 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 channels in the same MCWM system).

Aiming at solving these challenges, we introduce a novel WMUE attack detection scheme, which is motivated by the idea in [8], but tackles the multi-channel case. The basic idea is to check the cross-correlation between demodulated FM signal and the mixed acoustic signal. Since the cross-correlator is very robust in noisy environments, both the audio signals in other audio channels and noises have limited effect to the detection performance.

The cross-correlator has good anti-noise ability, but also suffers from high computation complexity. To address this issue, we propose a 1.5-bit FM receiver, which directly maps the FM signal to a piece of acoustic signal whose amplitude is represented by 0, 1 or -1. The computation complexity of such technique is much lower than the method of demodulating the FM signal with conventional technique [9], [10] in digital domain and then reducing the data precision by a three-level quantizer. Moreover, the 1.5-bit FM receiver leads to a multiplierless cross-correlator, and the entire detection scheme enjoys low complexity.

We evaluate the performance of the whole detection scheme by real-world testing, which includes an off-the-shelf MCWM system and a WSD prototype realized by RF components and an oscilloscope. Experimental results show that, the proposed

This material is based upon work supported by the National Science Foundation (NSF) CAREER award under Grant Number (CNS-1149500), and co-supported by NSF under Grant Number 1002113.

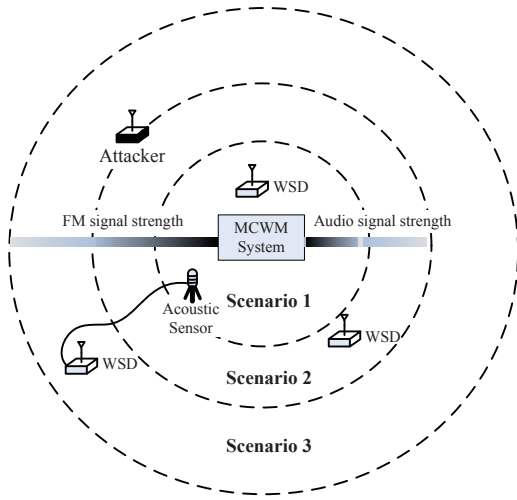


Fig. 1. The system model and three scenarios considered in this paper. Scenarios differ from each other in the quality of FM signals and acoustic signals.

scheme requires only -3 to 0 dB signal-to-noise ratio (SNR) when two audio channels are used, and requires about 5-6 dB SNR when four audio channels are used, with the performance of detection rate $\beta > 0.9$ and false alarm rate $\alpha < 0.1$. The detection time is as low as a quarter second.

Our contributions are summarized as follows:

- Propose a cross-correlation based WMUE attack detection scheme with the ability to resist noises and interferences in MCWM systems;
- Propose a 1.5-bit FM demodulator which enjoys low complexity and simplifies the cross-correlator;
- 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.

II. SYSTEM MODEL

A. 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), 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. On the other hand, all FM signals are separate since they locate at different channels. As a result, the WSD is able to acquire FM signals $s_m(t)$ for $m = 1, \dots, M$, 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 .

We consider that the quality of acoustic signal $a(t)$ drops much faster than the qualities of $s_m(t)$ when the propagation distance increases, and define three operating scenarios for the WSD who is sensing the spectrum and acoustic signal:

- *Scenario 1*: The WSD locates very close to the MCWM system, so both $s_m(t)$ and $a(t)$ have good quality;
- *Scenario 2*: The WSD locates a little far away from the MCWM system, so all $s_m(t)$ still have good quality, but $a(t)$ has poor quality;
- *Scenario 3*: The WSD locates very far away from the MCWM system, so all $s_m(t)$ have poor quality, but high-quality $a(t)$ is acquired by the sensor near the MCWM system and sent to the WSD through the infrastructure.

These three scenarios are illustrated in figure 1. We assume that the power of $s_m(t)$ is above the noise floor at each WSD in all scenarios, so that f_m for $m = 1, \dots, M$ can be estimated by the WSD [11]. Since f_m can only be a multiple of 25 kHz [12], 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 $m = 1, \dots, M$.

B. Attacker Model

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

It is undesirable for the attacker to convert the demodulated FM signal to audio signal and send it to the loudspeaker, since an unexpected or strange audio sound would be easier to be noticed by human, thus makes the attacker be detected easily. Therefore, we consider that the attacker does not generate any audio signal. 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 $s_m(t)$.

C. 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 $s_m(t)$ for $m = 1, \dots, M$ during the same time span. 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 and alternative hypothesis, respectively.

III. 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.

The basic procedures of the proposed scheme are shown in figure 2. The WSD first records the RF signal $s_m(t)$ on one channel and acoustic signal $a(t)$ simultaneously. Then it down-converts $s_m(t)$ to an IF signal $s_m^{(IF)}(t)$, and feeds the latter one into a low-complexity FM demodulator. Finally, the scheme computes the peak value X of the cross-correlation

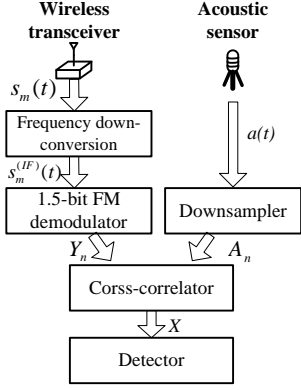


Fig. 2. Basic procedures of the proposed WMUE attack detection scheme.

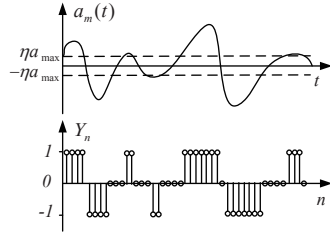


Fig. 3. The relationship between $a_m(t)$ and the desired output Y_n of a 1.5-bit FM demodulator.

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.

We will explain each component in the detection scheme in the following subsections.

A. Preliminaries

Before introducing the proposed FM demodulator, we will analyse some properties of the FM signal $s_m(t)$ first. We assume that $s_m(t)$ is down-converted to intermittent frequency (IF) centred at f_I which is kept constant for different wireless channels, and denote this IF signal as $s_m^{(IF)}(t)$. In other words, a superheterodyne receiver is considered here. Then we have [3]

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

For most superheterodyne receivers,

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

and

$$f_I > 2\Delta f a_{max} \quad (3)$$

where f_{max} and a_{max} denote the maximum frequency and maximum amplitude of $a_m(t)$, respectively. Then we define T as a number that satisfies

$$2f_{max} \leq 1/T < f_I \quad (4)$$

and

$$2\Delta f a_{max} \leq 1/T \quad (5)$$

and get the following lemma.

Lemma 1: The expression $|\int_{t_0}^{t_0+T} s_m^{(IF)}(t) e^{j2\pi g_k t} dt|$ is a monotonic decreasing function with respect to $|g_k - f_{t_0}|$, where

$$f_{t_0} := f_I + \Delta f a_m(t_0) \quad (6)$$

and

$$|g_k - f_{t_0}| \leq 1/T. \quad (7)$$

One can easily verify the correctness of *Lemma 1*, and derive the following lemma:

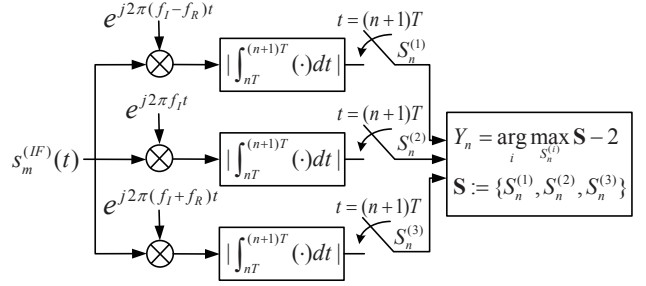


Fig. 4. The proposed 1.5-bit FM demodulator.

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

Since we focus on the m^{th} wireless channel here, we drop the index m if doing this would not cause misunderstanding.

B. 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 (1) if and only if

$$Y_{m,n} = \begin{cases} -1, & a_m(nT) < -\eta a_{max} \\ 0, & -\eta a_{max} \leq a_m(nT) < \eta a_{max} \\ 1, & \text{others} \end{cases} \quad (8)$$

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

Figure 3 shows the relationship between $a_m(t)$ and the desired output of a 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 information contained in $Y_{m,n}$ is maximized. For example, if the amplitude of $a_m(t)$ is evenly distributed over $[0, a_{max}]$, $\eta = 0.5$.

Proposition 1: The demodulator shown in figure 4 with output $Y_n = \arg \max_i S_n^{(i)} - 2$ is the 1.5-bit FM demodulator

defined in *Definition*, where $\mathbf{S}_n := \{S_n^{(1)}, S_n^{(2)}, S_n^{(3)}\}$, $g_1 = f_I - f_R$, $g_2 = f_I$, $g_3 = f_I + f_R$, and $f_R = 2\eta a_{max} \Delta f$.

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

$$\begin{aligned} & |g_1 - f_{nT}| \\ &= |f_R + \Delta f a_m(nT)| \\ &= |2\eta a_{max} \Delta f + \Delta f a_m(nT)| \\ &< |\Delta f a_m(nT)| \\ &= |g_2 - f_{nT}| \end{aligned} \quad (9)$$

and it is easily shown that $|g_1 - f_{nT}| < |g_3 - f_{nT}|$. Then according to *Lemma 2*, $S_n^{(1)} > S_n^{(2)}$ and $S_n^{(1)} > S_n^{(3)}$. As a result, $\arg \max_i S_n^{(i)} = 1$ and $Y_n = -1$.

In 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 [13]; however, their 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 exactly match any signal transmitted. Moreover, this 1.5-bit FM demodulator can also be interpreted as a sampler for the audio signal $a_m(t)$

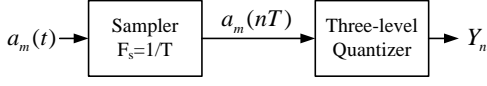


Fig. 5. The proposed 1.5-bit FM demodulator can be interpreted as a sampler for the audio signal $a_m(t)$ with sampling frequency $F_s = 1/T$ followed by a three-level quantizer.

with sampling frequency $F_s = 1/T$ followed by a three-level quantizer, as shown in figure 5.

This FM demodulator may also operate in digital domain, if the input $s_m^{IF}(t)$ is sampled. It is easily shown that, basic principle of this demodulator still holds in digital domain, and the complexity of this 1.5-bit FM demodulator is lower than conventional FM demodulators.

C. Audio Signal Processing

We model the acoustic signal $a(t)$ arriving at the WSD under H_0 (no attack) as

$$a(t)|H_0 = a^{(T)}(t) \otimes h(t) = \sum_{j=1}^J h_j a^{(T)}(t - t_j) + z(t) \quad (10)$$

where $a^{(T)}(t) := \sum_{m=1}^M a_m(t)$ denotes the mixed acoustic signal generated by the loudspeaker of the MCWM system, $h(t) := \sum_{j=1}^J h_j \delta(t - t_j)$ represents the impulse response of the acoustic channel between the loudspeaker and WSD, and $z(t)$ denotes the acoustic noises.

In practice, the acoustic signal travels slower than the RF signal. We address this issue in the acoustic channel model $h(t)$, and define the time delay t_0 of the first multipath as the acoustic signal's propagation delay D_m/v (instead of the conventional value 0), where D_m denotes the distance between the m^{th} WM and the WSD, and v denotes the speed of sound in the air. Other values of t_j when $j \neq 0$ also incorporate this propagation delay.

At the WSD side, $a(t)$ is sampled by the acoustic sensor at a high frequency, typically at 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 high-frequency noises.

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

$$A_n|H_0 = a^{(T)}(t) \otimes h(t) \otimes h_s(t) + z_L(t) \quad (11)$$

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 audio channel response, and define

$$d(nT) := h(t) \otimes h_s(t) := \sum_{l=1}^L d_l \delta(nT - t_l T) + Z_n \quad (12)$$

where t_l is a non-negative integer, and Z_n denotes the samples of noises which are independent to $a_m(t)$ for $m = 1, \dots, M$.

Combining (11) and (12), we get

$$\begin{aligned} A_n|H_0 &= \sum_{m=1}^M a_m(t) \otimes \sum_{l=1}^L d_l \delta(nT - t_l T) \\ &= \sum_{m=1}^M \sum_{l=1}^L d_l a_m(nT - t_l T) + Z_n. \end{aligned} \quad (13)$$

On the flip side, A_n under H_1 (attack case) is modelled by

$$A_n|H_1 = Z_n. \quad (14)$$

D. The Cross-correlator

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

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

According to (8) and (13), 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:

$$Y_{m,n} = a_m(nT) + Q_{m,n} \quad (15)$$

where $Q_{m,n}$ denotes the quantization error at $t = nT$. Then from (13) to (15), we get

$$\begin{aligned} \text{Corr}(A_n|H_0, Y_{m,n}, p) &:= \frac{1}{P_{m,p}} \sum_{n=0}^{W-1} (A_n|H_0) Y_{m,n-p} \\ &= \begin{cases} \frac{1}{P_{m,p}} (C_{m,p}^{(0)} + C_{m,p}^{(1)} + C_{m,p}^{(2)}), p = t_l \\ \frac{1}{P_{m,p}} (C_{m,p}^{(2)} + C_{m,p}^{(3)}), \text{others} \end{cases} \end{aligned} \quad (16)$$

where

$$P_{m,p} := \sqrt{\left(\sum_{n=0}^{W-1} (A_n)^2 \right) \left(\sum_{n=0}^{W-1} (Y_{m,n-p})^2 \right)} \quad (17)$$

$$C_{m,p}^{(0)} := d_l \sum_{n=0}^{W-1} |a_m(n - t_l)|^2 \quad (18)$$

$$C_{m,p}^{(1)} := \sum_{n=0}^{W-1} \sum_{\substack{m', l' \\ |m-m'|+|l-l'| \neq 0}} a_{m'}(n - t_{l'}) a_m(n - t_l) \quad (19)$$

$$C_{m,p}^{(2)} := \sum_{n=0}^{W-1} \sum_{m=1}^M \sum_{l=1}^L d_l a_m(n - t_l) Q_{m,n} + \sum_{n=0}^{W-1} Z_n Y_{n-p} \quad (20)$$

$$C_{m,p}^{(3)} := \sum_{n=0}^{W-1} \sum_{m'=1}^M \sum_{l'=1}^L d_{l'} a_{m'}(n - t_{l'}) a_m(n - p). \quad (21)$$

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

Similarly, $\text{Corr}(A_n|H_1, Y_{m,n}, p)$ is obtained by setting $a_m(t) = 0$ for $m = 1, \dots, M$ in (16):

$$\text{Corr}(A_n|H_1, Y_{m,n}, p) = \frac{1}{P_{m,p}} \sum_{n=0}^{W-1} Z_n Y_{n-p}. \quad (22)$$

The audio noises Z_n and quantification error Q_n are considered as uncorrelated to Y_n and $a_m(t)$, respectively. As a result, $\text{Corr}(A_n|H_1, Y_{m,n}, p)$ is close to 0. On

the flip side, due to the existence of $C_{m,p}^{(0)}$ given in (18), $\text{Corr}(A_n|H_0, Y_{m,n}, p)$ 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, $\text{Corr}(A_n|H_0, Y_{m,n}, p)$ is even larger because of $C_{m,p}^{(1)}$ given in (19). In any case, $\text{Corr}(A_n|H_0, Y_{m,n}, p)$ is expected to exceed $\text{Corr}(A_n|H_1, Y_{m,n}, p)$ when $p = t_l$.

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

$$X = \max_{p=0, \dots, \tau_{\max}} \{\text{Corr}(A_n, Y_{m,n}, p)\} \quad (23)$$

where τ_{\max} represents the maximum delay spread of the audio channel divided by T (and rounded to the nearest integer if necessary). (23) 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 .

2) *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*.

E. The Detector

According to the analysis in subsection III-D, $X|H_0$ 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.

The detection threshold X_0 falls in the range $(0, 1)$, because X is the output from the cross-correlator and $0 \leq X < 1$. The detection time equals to TW .

In order to get X , $\text{Corr}(A_n, Y_{m,n}, p)$ needs to be calculated for $(\tau_{\max} + 1)$ times with different values of p . In the definition of $\text{Corr}(A_n, Y_{m,n}, p)$ given in (16), 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 $\frac{1}{F_{m,p}}$ can be derived from an iterative way [14], 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.

IV. PERFORMANCE EVALUATION

A. The Real-World Testing Environment

To evaluate the performance of the proposed detection scheme, we set up a real-world testing environment as shown in figure 6. A MCWM system and a WSD prototype are set up in a $12m \times 7m$ room. The MCWM system contains a 8-channel WM receiver manufactured by Pyle Audio Inc. with model number PDWM8400, a 40W loudspeaker, and 8 WMs. The

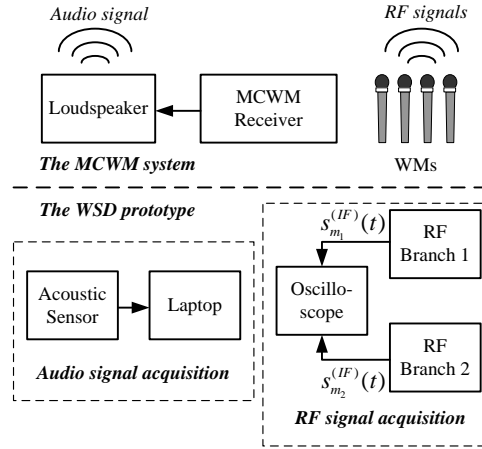


Fig. 6. Block diagram of the real-world testing environment.

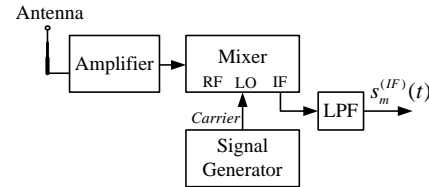


Fig. 7. Block diagram of one RF branch.

carrier frequencies of these 8 channels are within the range of 170-240 MHz, which falls into the very high frequency (VHF) band.

The two RF branches in figure 6 share the same design as shown in figure 7, which contains a frequency down-conversion circuit realized by a level-7 mixer. A signal generator serves as the local oscillator of the WSD and provides carrier for the mixer. Moreover, the wireless signal is amplified by an amplifier at RF and filtered by a LPF at intermittent frequency. The frequency response (i.e., $2\Delta f_{a_{\max}}$) of the WM system is 18 kHz according to the manual, but measured at only 4 kHz. Then we set $f_I = 10$ kHz and $T = 0.5$ ms in the experiments.

B. Testing methods

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. As a result, 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, 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 $T' = 20$ us and $n' = 0, 1, \dots$. 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 algorithms are

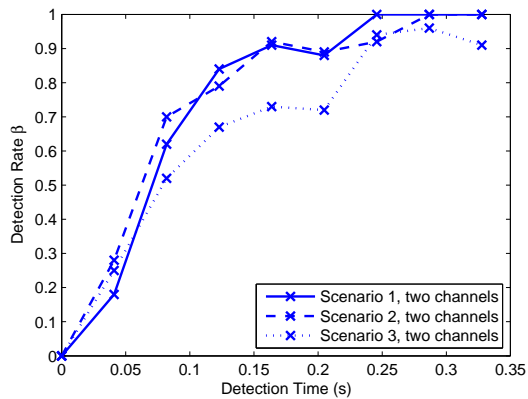


Fig. 8. Detection rate versus detection time in three scenarios in the case of two channels.

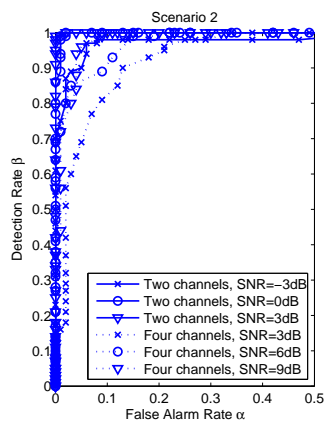


Fig. 9. ROC curves in *Scenario 2* under different SNR conditions in the cases of two channels and four channels, respectively.

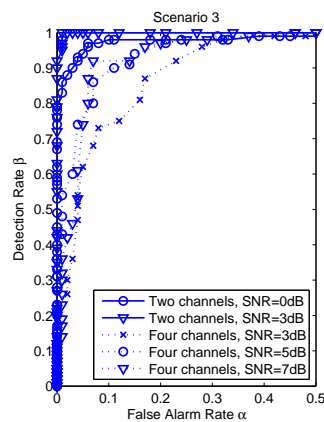


Fig. 10. ROC curves in *Scenario 3* under different SNR conditions in the cases of two channels and four channels, respectively.

executed twice and the results are averaged. We consider that the amplitude of audio signal follows uniform distribution for simplicity, and set $\eta = 0.5$.

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.

C. Testing Results

We first 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 8. 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 $W \geq 500$. We set $W = 500$ in the following experiments.

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 performs well when SNR is higher than -3 dB and 6 dB in the cases of two channels and four channels, respectively, as shown in figure 9. In *Scenario 3*, the lowest SNR required by the proposed scheme in the two cases are 0 dB and 5 dB, respectively, as shown in figure 10.

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

V. CONCLUSIONS

In this paper, we propose a novel and simple scheme to detect WMUE attacks imposed on MCWM systems in noisy environments. 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 and interferences. We set up a MCWM system and design a WSD prototype for performance evaluation. Hardware based experiments show that, the proposed scheme is able to detect WMUE attacks within 0.25 s in all scenarios with low SNR.

We conclude that, the proposed multi-channel WMUE attack detection scheme achieves good performances in noisy environments. Performance of the proposed scheme may be further enhanced by multiple antenna or collaborate sensing techniques, which are considered as our future works.

REFERENCES

- [1] A. Sahai, N. Hoven, and R. Tandra, "Some fundamental limits in cognitive radio," in *Proc Allerton Conf Commun Control Comput.*, 2004.
- [2] 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.
- [3] H.-S. Chen and W. Gao, "Spectrum sensing for tv white space in north america," *Selected Areas in Communications, IEEE Journal on*, vol. 29, no. 2, pp. 316–326, Feb. 2011.
- [4] R. Chen, J.-M. Park, and J. Reed, "Defense against primary user emulation attacks in cognitive radio networks," *Selected Areas in Communications, IEEE Journal on*, vol. 26, no. 1, pp. 25–37, Jan. 2008.
- [5] Z. Yuan, D. Niyato, H. Li, J. B. Song, and Z. Han, "Defeating primary user emulation attacks using belief propagation in cognitive radio networks," *Selected Areas in Communications, IEEE Journal on*, vol. 30, no. 10, pp. 1850–1860, Nov. 2012.
- [6] N. Nguyen, R. Zheng, and Z. Han, "On identifying primary user emulation attacks in cognitive radio systems using nonparametric bayesian classification," *Signal Processing, IEEE Transactions on*, vol. 60, no. 3, pp. 1432–1445, March 2012.
- [7] 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.
- [8] S. Chen, K. Zeng, and P. Mohapatra, "Hearing is believing: Detecting wireless microphone emulation attack in white space," *Mobile Computing, IEEE Transactions on*, 2011.
- [9] J. Garodnick, J. Greco, and D. Schilling, "Theory of operation and design of an all-digital fm discriminator," *Communications, IEEE Transactions on*, vol. 20, no. 6, pp. 1159–1165, 1972 Dec.
- [10] A. Saha and B. Mazumder, "A digital phase-locked loop for generating frequency discriminating codes and frequency multiplication," *Proceedings of the IEEE*, vol. 69, no. 4, pp. 472–473, 1981 April.
- [11] 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*, no. 69, p. 297306, Sep. 2011.
- [12] "FM Broadcast Translator Stations and FM Broadcast Booster Stations, 47 CFR Part 74." [Online]. Available: <http://transition.fcc.gov/mb/audio/bickel/part74rule.html>
- [13] L. Zadeh and J. Ragazzini, "Optimum filters for the detection of signals in noise," *Proceedings of the IRE*, vol. 40, no. 10, pp. 1223–1231, Oct. 1952.
- [14] D. Schmidl, T.M.; Cox, "Robust frequency and timing synchronization for ofdm," *Communications, IEEE Transactions on*, vol. 45, no. 12, pp. 1613–1621, Dec. 1997.