2023 IEEE Symposium on Electromagnetic Compatibility & Signal/Power Integrity (EMC+SIPI) | 979-8-3503-0976-8/23/$31.00 ©2023 IEEE | DOI: 10.1109/EMCSIPI50001.2023.10241678

Reconstruction of Sound Information Leakage Signals Obtained from Multiple Demodulation Methods

Taiki Kitazawa
Graduate School of Information Science
Nara Institute of Science and Technology
Nara, Japan
kitazawa.taiki.kq8@is.naist.jp

Seiya Takano
Graduate School of Information Science
Nara Institute of Science and Technology
Nara, Japan
takano.seiya.tp2@is.naist.jp

Yuichi Hayashi
Graduate School of Information Science
Nara Institute of Science and Technology
Nara, Japan
yu-ichi@is.naist.jp

***Abstract*— Speakerphones used in remote work environments have been reported to leak audio information through electromagnetic (EM) emanations. A method to improve information reconstruction quality has been proposed by simultaneous measurement of multiple leakage channels. However, acquiring sufficient reconstruction quality is difficult when the number of observable channels is limited due to limitations of measurement resources. In this study, we propose a method to increase the number of leakage channels virtually in order to reconstruct audio information with sufficient quality, even if observable channels are limited. We assume that leaked EM waves are modulated by multiple schemes and simultaneously extract audio information from a single leakage channel by amplitude and frequency demodulation methods. Furthermore, we synthesize audio signals obtained from different demodulation methods by a phase-aligned-spectrum-averaging method and indicate that the improvement in the reconstruction quality is comparable to measuring multiple leakage channels.**

Keywords—electromagnetic information leakage, speakerphone, TEMPEST, compromising electromagnetic emanations

I. Introduction

Information on human perception handled within input/output (I/O) devices is challenging to encrypt; therefore, it leaks instantly once retrieved. This I/O information propagates within devices as electrical signals; hence, the electromagnetic (EM) radiation caused by their time fluctuations can be used to acquire information [1]–[11].

TEMPEST is the codename for the threat of information leakage from input/output (I/O) devices through EM waves. The targets of such breaches have primarily been video display units (VDUs) [1]–[3], printers [4], keyboards [5], and other I/O devices. In addition, it has recently been discovered that speakerphones also pose a risk of leaking audio information [6], [7]. This indicates that it is possible to capture audio information from a switchboard for electrical power distribution at a distance of around 30 meters through the conducted emission in a speakerphone placed in a specific EM environment. This phenomenon of audio information leakage through EM is referred to as audio TEMPEST.

In display TEMPEST, multiple almost identical screen information can be observed because a screen refreshes several frames per second. However, in audio TEMPEST, applying averaging processing to reduce noise is challenging because little repetition of the same information is generated.

The same method used in display TEMPEST has been considered inapplicable to improve reconstruction quality in audio TEMPEST.

In contrast, a method to improve the reconstruction quality has been proposed that focuses on multiple frequency bands where audio information leaks from devices (leakage channels) and acquires the same multiple audio information by measuring those channels simultaneously [8]. However, sufficient reconstruction quality may not be achieved when measurement resources are limited.

In this paper, we propose a novel method to virtually increase the number of leakage channels in order to reconstruct audio information with the same quality as measuring multiple leakage channels, even if observable channels are limited.

We assume that EM waves leaked from a single channel are modulated by amplitude modulation (AM) and frequency modulation (FM) and acquire 2*N* signals, including audio information, through *N* measurable leakage channels by demodulating them individually. In the experiment, we applied a phase-aligned-spectrum-averaging method to signals measured from a single channel and demodulated them using AM and FM. Consequently, this reconstructed audio information has the same audio quality as the conventional method, which requires the measurement of multiple leakage channels.

II. Proposed Virtual Leakage Channels Considering Multiple Modulation Method

In this section, we explain our proposed method to increase leakage channels virtually by applying multiple demodulation methods to measurable EM waves, assuming that they are mixed multiple modulation signals on a single leakage channel. In this study, we focus on AM and FM and reconstruct audio using each demodulation. Reconstructed audio information is synthesized by spectrum averaging and phase alignment based on short-time Fourier transform (STFT) to improve the reconstruction qualities (Fig. 1(a)).

A. Acquirement of Audio Information Using Multiple Demodulation Schemes

In conventional audio TEMPEST, audio information is reconstructed using AM demodulation to measure leaked EM waves through conducted emission from power lines [6], [7].

Meanwhile, in other TEMPEST research, the clock frequency to control a keyboard has been demonstrated to fluctuate according to keyboard input because the clock depends on the supply voltage for the IC, which also fluctuates [9]. Therefore, the key input information can be reconstructed using FM demodulation to measure leaked EM waves at the fundamental and harmonic frequencies. Moreover, the same reconstruction has been reported in display TEMPEST [10], [11]; in TEMPEST against speakerphones, there is a possibility to reconstruct audio reconstruction using not only AM but also FM demodulation.

Given this assumption, a signal observed from a single leaked channel could be modulated by AM and FM. Therefore, we could virtually increase the leakage channels by demodulating them individually and acquiring two audio signals.

Here, we measure the leaked EM waves as the in-phase/quadrature-phase (I/Q) to apply AM and FM demodulation. We can write the measured signal $x(t)$ at time t as I/Q data using the following equation:

$$x(t) = I + jQ = Ae^{j(2\pi ft+\theta)}$$

$$\because I(t) = Acos(2\pi ft+\theta), Q(t) = Asin(2\pi ft+\theta) \quad (1)$$

where A is the amplitude, f is the instantaneous frequency, and θ is the phase. The actual receiver sample signals at a constant sampling time T_s:

$$x[i] = I + jQ = Ae^{j(2\pi fT_S i+\theta)}$$

$$\because I[i] = Acos(2\pi fT_s i+\theta), Q[i] = Asin(2\pi fT_s i+\theta) \quad (2)$$

We then apply AM and FM demodulation, assuming that the measured signals fluctuate independently in amplitude and frequency.

$$A[i] = \sqrt{(I[i])^2 + (Q[i])^2} \quad (3)$$

$$f[i] = \frac{1}{2\pi T_s} \arg\left(x[i] * \overline{x[i-1]}\right) \quad (4)$$

where we can reconstruct independently because (3) does not depend on the frequency, and (4) likewise does not depend on the amplitude.

Next, we explain the signal processing for the measured signals to be recognized as sound, as shown in Fig.1 (b). The sampling rate is higher than the typical one in audio because the measurement is conducted with several MHz bandwidths; they are then downsampled to 44.1 kHz. Finally, we reconstruct the audio by filtering to remove noise outside the audible range and rescaling it.

B. Synthesis Method for Multiple Sound Signals

In display TEMPEST, reconstruction qualities can be improved by simple averaging in the time domain using multiple almost identical image information by refresh screens [1], [2], [10]. However, the intensities and frequency characteristics of acquired signals differ for each frequency band and demodulation method, despite the information (characters or audio) being the same [2], [9]. Therefore, we consider that the simple averaging method in the time domain cannot improve the reconstruction quality of sound signals acquired by AM and FM demodulation.

In this section, we explain the phase-aligned-spectrum-averaging method to synthesize sound signals acquired by AM and FM demodulation.

Fig.1 (c) presents an overview of the phase-aligned-spectrum-averaging method. First, as mentioned in the previous subsection, we perform AM and FM demodulation and processing for audio recognition of I/Q data measured through leaked EM waves. We then perform an STFT on the sound signals using the window function to derive the spectrum at each time index. Here, in amplitudes of the spectrum, we add only spectral components exceeding a certain threshold P_{th} to suppress noise; furthermore, we reduce components below threshold P_{th} of the spectrum until -120 dB. Meanwnhile, we align the phases of the spectrum to that of an arbitrary input signal because they differ in each frequency bin according to the inputs. Finally, we perform an inverse STFT (ISTFT) and filter the time-frequency response after the addition and phase alignment processes in every time index.

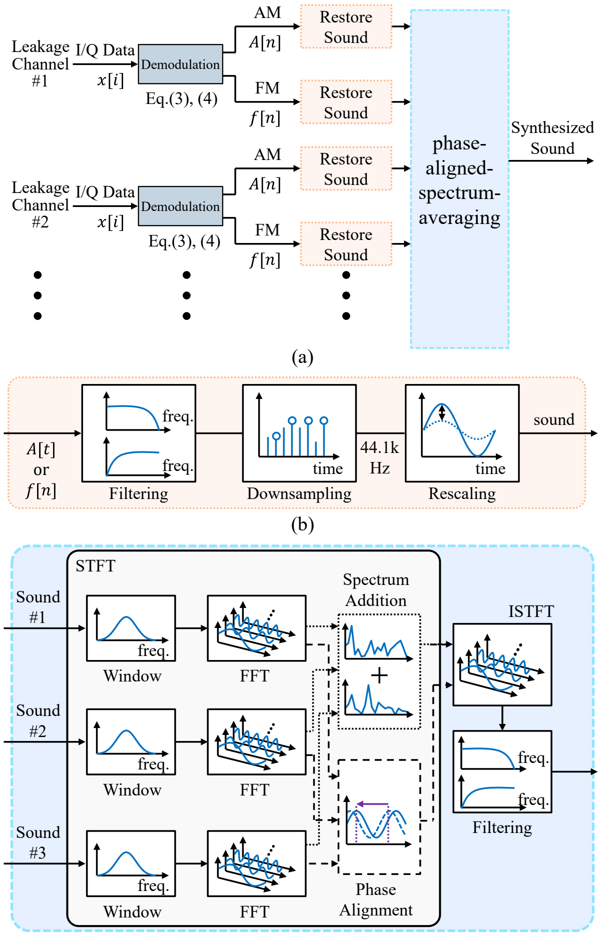


Fig.1 Block diagram of sound reconstruction by phase-aligned-spectrum-averaging including virtual leakage channels. (a) overview, (b) restore sound, (c) phase-aligned-spectrum-averaging

III. Reconstruction of Audio Information Through Leaked EM Waves Using Proposed Method

A. Experimental Conditions and Setup

In this section, we experimented to measure the conducted emissions from a speakerphone and evaluate them by comparing the reconstruction quality of audio information. We compared sound information under the following two conditions:

(a) Reconstructed sound synthesized by two signals demodulated by AM at different receive frequencies. (Conventional Method)

(b) Reconstructed sound synthesized by two signals demodulated by AM and FM at a single receive frequency. (Proposed Method)

We derived the signal-to-noise ratio (SNR) and spurious-free dynamic range (SFDR) as evaluation parameters from the reconstructed audio in the frequency domain.

$$[\mathrm{SNR}]_{\mathrm{dB}} = 20\log\frac{P_{signal}}{P_{noise}} \tag{5}$$

$$[\mathrm{SFDR}]_{\mathrm{dB}} = 20\log\frac{P_{signal}}{P_{spurious}} \tag{6}$$

where P_{signal} is the intensity at a frequency that includes audio information, P_{noise} is the root-mean-square (RMS) value of the spectrum without components of DC and fundamental and harmonics of P_{signal}, $P_{spurious}$ is a high peak next to the fundamental and harmonics of P_{signal}.

Fig. 2 shows the experimental environment. In this experiment, the sound signal leaking from the power line of the speakerphone by the conducted emissions was measured by a current probe and amplified by a low noise amplifier (LNA). The signal was then separated by a splitter and input into two spectrum analyzers. Measurements were performed in an anechoic chamber for reproducibility, and the sound was played back via Bluetooth. In this experiment, we excited Gaussian noises of -25 dBm from the function generator (FG) through a coupler and attenuator to consider the degradation of the reconstruction quality caused by emissions from non-target devices.

I/Q signals were acquired with two spectrum analyzers, with receive frequencies f_{CF} of 5 MHz and 12 MHz. The sampling rate was 1 MSa/s, the bandwidth was 800 kHz, and the measurement time was 5 seconds. The time window for the STFT was a triangular window of 50 ms, with an overlap of 50%. The threshold value P_{th} used for denoising was the RMS value of the spectrum in each time frame.

B. Reconstruction of Audio Information Using Proposed Method

1) Experiment using a 4 kHz sinusoidal sound

The first experiment evaluated the restored sound obtained when a 4 kHz sinusoidal wave was reproduced. Fig. 3 shows the spectrum of the reconstructed sound with AM and FM demodulation applied to the measured signals at each received frequency. The 4 kHz peak was observed in all the recovered sounds, confirming that multiple sound signals can be extracted from a single leakage channel using different demodulation methods. In addition, a comparison of the AM/FM demodulation results shows that AM demodulation has a strong P_{signal} (4 kHz) intensity and a strong background noise component. Conversely, in FM demodulation, the intensity of P_{signal} (4 kHz) was weak, but the background noise component was hardly observed.

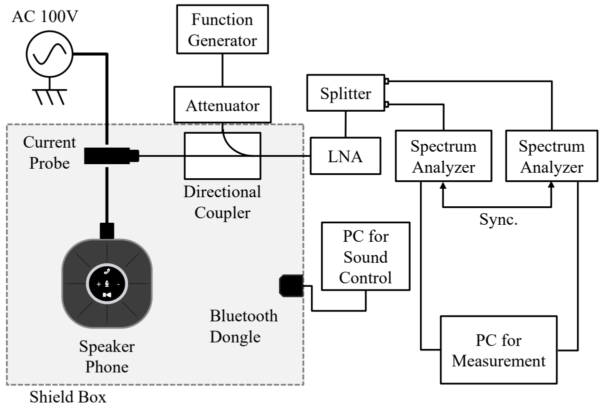


Fig2. Measurement setup

TABLE I MEASUREMENT SETUP

Spectrum Analyzer	Rohde & Schwartz FSV
Function Generator	KEYSIGHT 81160A
Speakerphone	Anker Power Conf
Attenuator	Mini-Circuits VAT-20+ (DC-6 GHz)
Directional Coupler	Mini-Circuits ZFDC-20-5-S+ (0.1-2000 MHz)
Low Noise Amplifier	R&K LA101-0S (9 kHz-1000 MHz)
Splitter	Mini-Circuits ZFRSC-42-S+ (DC-4200 MHz)
Shield Box	MICRONIX Taurus MY1510
Current Probe	F.C.C. F-2000-12mm (10 MHz-3GHz)

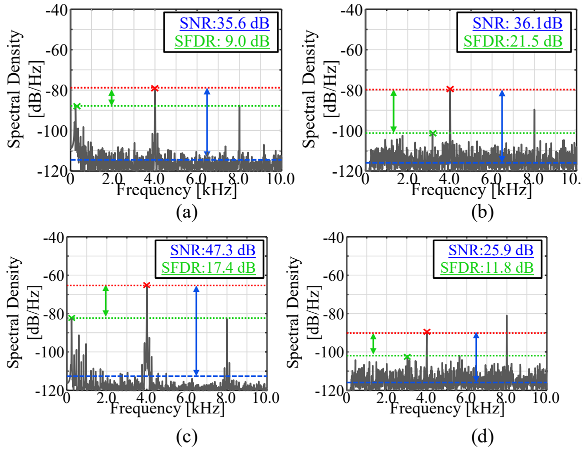
Figs. 4 and 5 show the spectra after the synthesis, and Table II lists the calculated SNR and SFDR, respectively. Here, the frequency resolution was 10 Hz. The results of synthesizing the sound information obtained by observing a

TABLE II SNR AND SFDR IN EXPERIMENTS OF SINUSOIDAL SOUND

	(1)	(2)	(3)	(4)	(1) + (3)	(1) + (2)	(3) + (4)
f_{CF}	5 MHz	5 MHz	12 MHz	12 MHz	5/12 MHz	5 MHz	12 MHz
Demodulation	AM	FM	AM	FM	AM	AM/FM	AM/FM
Spectrum Averaging	No	No	No	No	Yes	Yes	Yes
Method	[5], [6]	[5], [6]	[5], [6]	[5], [6]	[7]	Proposed	Proposed
SNR (Improvement)	35.6 dB	36.1 dB	47.3 dB	25.9 dB	51.3 dB (+4.0 dB)	50.6 dB (+14.5 dB)	51.7 dB (+4.4 dB)
SFDR (Improvement)	9.0 dB	21.5 dB	17.4 dB	11.8 dB	26.4 dB (+9.0 dB)	26.1 dB (+5.0 dB)	20.4 dB (+3.0 dB)

TABLE IIISNR AND SFDR IN EXPERIMENTS OF CHIRP SOUND

	(1)	(2)	(3)	(4)	(1) + (3)	(1) + (2)	(3) + (4)
f_{CF}	5 MHz	5 MHz	12 MHz	12 MHz	5/12 MHz	5 MHz	12 MHz
Demodulation	AM	FM	AM	FM	AM	AM/FM	AM/FM
Spectrum Averaging	No	No	No	No	Yes	Yes	Yes
Method	[5], [6]	[5], [6]	[5], [6]	[5], [6]	[7]	Proposed	Proposed
SNR (Improvement)	23.6 dB	19.5 dB	31.8 dB	15.7 dB	38.8 dB (+7.0 dB)	35.1 dB (+11.5 dB)	37.5 dB (+5.7 dB)
SFDR (Improvement)	2.1 dB	4.1 dB	4.3 dB	0.2 dB	3.6 dB (-0.7 dB)	5.3 dB (+1.2 dB)	4.6 dB (+0.3 dB)



(a) (b) (c) (d)

Fig. 3 Spectrum of reconstructed sound of sinusoidal with 4 kHz in leakage channel (a) AM, f_{CF} = 5 MHz, (b) FM, f_{CF} = 5 MHz, (c) AM, f_{CF} = 12 MHz, (d) FM, f_{CF} = 12 MHz

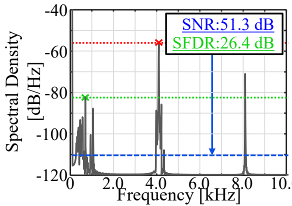
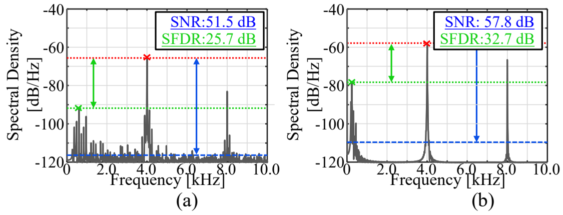
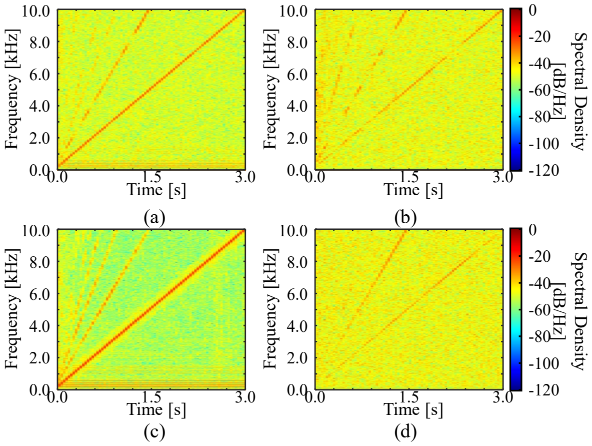


Fig. 4 Spectrum of combined sound in different leakage channels at f_{CF} = 5 and 12 MHz (conventional method)



(a) (b)

Fig. 5 Spectrum of reconstructed sound in a single channel (proposed method), (a) AM/FM, f_{CF} = 5 MHz, (b) AM/FM, f_{CF} = 12 MHz



(a) (b) (c) (d)

Fig. 6 Spectrum of reconstructed sound of chirp (100 Hz-1 kHz) in leakage channels (a) AM, f_{CF} = 5 MHz, (b) FM, f_{CF} = 5 MHz, (c) AM, f_{CF} = 12 MHz, (d) FM, f_{CF} = 12 MHz

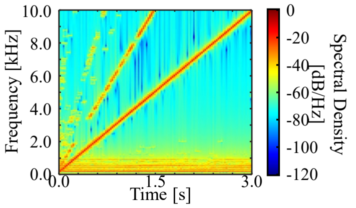
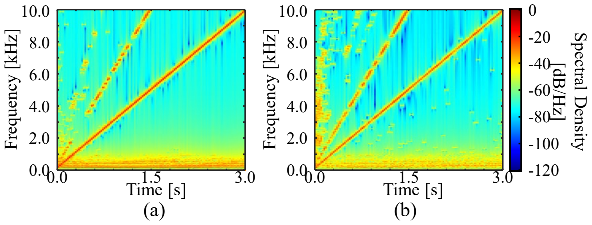


Fig. 7 Spectrum of combined sound in different leakage channels at f_{CF} =5 and 12 MHz (conventional method)



(a) (b)

Fig. 8 Spectrum of reconstructed sound in a single channel (proposed method), (a) AM/FM, f_{CF} = 5 MHz, (b) AM/FM, f_{CF} = 12 MHz

single leakage channel of the proposed method and applying AM/FM demodulation (Condition b) show an improvement of approximately 4 to14 dB in the SNR and 3 to 5 dB in the SFDR. These values are equivalent to or better than those obtained with multiple leakage channels using conventional methods (Condition a). This shows that even when the number of leakage channels that can be observed is limited because of instrument limitations, recovery accuracy can be improved by applying multiple demodulation schemes to a single leakage channel and generating a virtual leakage channel.

2) Experiment using a chirp sound (100 Hz to10 kHz)

We evaluated the reconstructed audio obtained when the chirp was reproduced. The chirp is a sweeping sound ranging from 100 Hz to 10 kHz for three seconds. Fig. 6 shows the time-frequency responses of sound reconstructed by applying FM and AM demodulation to each leaked channel. We found chirp trends in the reconstructed sound, similar to the sinusoidal experiments. Moreover, although the P_{signal} is high in AM demodulation, noise is also strong at approximately 100 Hz to1 kHz, and the opposite trend was observed in FM demodulation.

Figs. 7 and 8 indicate the time-frequency responses of sound synthesized by phase-aligned-spectrum-averaging in the case of Conditions a and b. Table III also lists the SNR and SFDR average derived from each time index, as with the sinusoidal evaluation. SNR improved by approximately 5 to 11 dB in both conditions when multiple leakage channels were measured (Condition a) and when the demodulation scheme was altered at a single leakage channel (Condition b), respectively. We found that the SNR and SFDR values were almost the same as those of conditions a and b, as with the sinusoidal experiments. Hence, we demonstrate that improving reconstruction quality by increasing the number of channels virtually focused on demodulation methods is feasible.

Meanwhile, in the time-frequency responses, we can see differences in the spectrum intensities in harmonic chirps (overtones) and noises below 1 kHz. Therefore, another evaluation method that considers the reconstruction quality when actual words are played back is required.

IV. Conclusion

This study focused on the threat of electromagnetic information leakage from I/O devices that process sound information, such as speakerphones. We developed a method to virtually increase the number of leakage channels even when measurement resources are limited and recover sound information equivalent to that obtained when multiple leakage channels are measured.

We focused on the modulation method of the leakage electromagnetic waves to virtually increase the number of leakage channels and extracted sound information from a single leakage channel by applying AM/FM demodulation to the measured leakage electromagnetic waves. By synthesizing these, we confirmed that the quality of the restored sound information could be improved.

To demonstrate the effectiveness of the proposed method, we restored sound information using the proposed method on the observed signal when sinusoidal waves and a chirping sound were played back. These values are equivalent to those obtained by restoring the sound information using multiple leakage channels, demonstrating the effectiveness of the proposed method.

However, when actual sound is the target, evaluating the attacker's ability to obtain information requires a different metric than the one used in this study. The proposal of such a metric is a topic for future research, alongside investigating how the EM waves leak across different frequency bands in various demodulation schemes.

Acknowledgment

This work was supported in part by the JST FOREST Program under Grant JPMJFR206L and in part by JSPS KAKENHI under Grant 19H01104 and Grant 21K19772.

References

[1] W. van Eck, "Electromagnetic radiation from video display units: An eavesdropping risk?," *Computers and Security*, vol. 4, no. 4, pp. 269–286, 1985.

[2] T. Kitazawa, K. Arai, Y. Kim, D. Fujimoto, and Y. Hayashi, "A Novel Remote Visualization of Screen Images Against High-Resolution Display With Divided Screens Focusing on the Difference of Transfer Function of Multiple Emanations," *IEEE Trans. Electromagn. Compat.*, vol. 64, no. 6, pp. 1941–1948, Dec. 2022.

[3] D.-H. Choi, E. Lee, T. Nam, and J.-G. Yook, "Recent Trends in Image Information Recovery Using Leaked Electromagnetic Wave from Electronic Equipment," *IEEE Electromagnetic Compatibility Magazine*, vol. 11, no. 3, pp. 77–83, rd 2022.

[4] T. Tosaka, K. Taira, Y. Yamanaka, A. Nishikata, and M. Hattori, "Feasibility study for reconstruction of information from near field observations of the magnetic field of laser printer," in *2006 17th International Zurich Symposium on Electromagnetic Compatibility*, Feb. 2006, pp. 630–633.

[5] M. Kinugawa, D. Fujimoto, and Y. Hayashi, "Electromagnetic Information Extortion from Electronic Devices Using Interceptor and Its Countermeasure," *TCHES*, pp. 62–90, Aug. 2019.

[6] S. Fukushima, D. Fujimoto, and Y. Hayashi, "Evaluation of EM Information Leakage from Speakerphones under Remote Work Environments," 2021 Symposium on Cryptography and Information Security (SCIS2021), vol. 2D2-2, 2021.

[7] S. Fukushima, D. Fujimoto, and Y. Hayashi, "Fundamental Study on Evaluation of EM Information Leakage from Smart Speakers with Different Installation Environments and Its Countermeasures," IEICE Technical Report, vol. 120, no. 401, pp. 130–135, 2021.

[8] J. Choi, H.-Y. Yang, and D.-H. Cho, "TEMPEST Comeback: A Realistic Audio Eavesdropping Threat on Mixed-signal SoCs," in *Proceedings of the 2020 ACM SIGSAC Conference on Computer and Communications Security*, New York, NY, USA: Association for Computing Machinery, 2020, pp. 1085–1101.

[9] M. Kinugawa, Y.-I. Hayashi, T. Mizuki, and H. Sone, "Study on information leakage of input key due to frequency fluctuation of RC oscillator in keyboard," *IEICE Trans. Commun.*, vol. E96.B, no. 10, pp. 2633–2638, 2013.

[10] P. De Meulemeester, B. Scheers, and G. A. E. Vandenbosch, "Differential Signaling Compromises Video Information Security Through AM and FM Leakage Emissions," *IEEE Trans. Electromagn. Compat.*, vol. 62, no. 6, pp. 2376–2385, Dec. 2020.

[11] M. Prvulovic, A. Zajić, R. L. Callan, and C. J. Wang, "A Method for Finding Frequency-Modulated and Amplitude-Modulated Electromagnetic Emanations in Computer Systems," *IEEE Trans. Electromagn. Compat.*, vol. 59, no. 1, pp. 34–42, Feb. 2017.