International Journal of Creative and Open Research in Engineering and Management
ISSN: 3108-1754 (Online)
Volume 02 Issue 10 October-2026 | Impact Factor: 3.5



Acoustic Interference System for Audio Privacy Protection

Mrs. Reena D Kulkarni
Assistant Professor
Dept. of ECE
K.S. School of Engineering and Management
Bengaluru, India
reenadk@kssem.edu.in

Chaithu Priya K
Student
Dept. of ECE
K.S. School of Engineering and Management
Bengaluru, India
chaithupriya@kssem.edu.in

G Raghavendra Rao
Student
Dept. of ECE
K.S. School of engineering and Management
Bengaluru, India
raghavendragr@kssemedu.in

Lingutla Rakesh
Student
Dept. of ECE
K.S. School of Engineering and Management
Bengaluru, India
lrakesheng2005@gmail.com

Namburi Lehya
Student
Dept. of ECE
K.S. School of Engineering and Management
Bengaluru, India
lehyan@kssem.edu.in

How to Cite this Article:
K, C. P., Rao, G. R., Rakesh, L. & Lehya, N. (2026). Acoustic Interference System for Audio Privacy Protection. International Journal of Creative and Open Research in Engineering and Management, 02(10), 1-9. https://doi.org/10.55041/ijcope.v2i10.244

License:
This article is published under the terms of the Creative Commons Attribution 4.0 International License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.
© The Author(s). Published by International Journal of Creative and Open Research in Engineering and Management.





https://doi.org/10.55041/ijcope.v2i10.244

***Abstract*—The proliferation of smart IoT devices with always-on microphones creates significant privacy vulnerabilities through unauthorized audio recording. This project presents an "Acoustic Interference System for Audio Privacy Protection" that prevents eavesdropping by leveraging the inherent hardware nonlinearity of commercial off-the-shelf (COTS) microphones. The system utilizes an array of ultrasonic transducers to emit near-ultrasonic waves (20–40 kHz) that are inaudible to human ears.**

Upon reaching a target microphone, these signals undergo self-demodulation at the diaphragm, generating a low-frequency "shadow signal" within the audible spectrum. This induced noise effectively masks legitimate speech, rendering recordings unintelligible to human eavesdroppers and automatic speech recognition engines. By targeting the physical layer of the recording hardware, the system provides a robust defense that cannot be easily bypassed by software filters. Experimental evaluations demonstrate that the system successfully disrupts unauthorized recordings while maintaining a quiet user environment

Keywords — IoT, Eavesdropping, nonlinearity, Ultrasonic Transducers, Self-demodulation, Shadow signal.

I. INTRODUCTION

The rapid proliferation of microphone-equipped smart devices, including **voice assistants** like Amazon Alexa and Google Home, smartphones, and wearable electronics, has revolutionized human-machine interaction. However, this ubiquity introduces significant **privacy risks**, as these "always-on" microphones can be exploited for unauthorized eavesdropping and secret recordings without user consent. Reports of devices recording private conversations without the required wake word have sparked global concern, highlighting the urgent need for robust physical-layer defenses. Traditional audio privacy methods often fall short of modern requirements, as physical mute buttons compromise the convenience of hands-free interaction, and audible jamming disrupts normal human communication. To address these limitations, recent engineering research leverages the inherent **hardware nonlinearity** found in commercial off-the-shelf (COTS) microphones. Although microphones are designed as linear systems, the physical properties of their diaphragms and pre-amplifiers are imperfect, causing them to exhibit nonlinear behavior when exposed to high-frequency signals. By emitting **near-ultrasonic waves**, typically between 16 kHz and 40 kHz, an interference system can exploit this nonlinearity to generate a low-frequency **"shadow signal"** within the audible band of the microphone's circuit. This shadow signal, while completely inaudible to human ears, is captured by the recording device as a loud masking noise that renders recorded speech unintelligible

Advanced acoustic interference systems have evolved from simple energetic masking to sophisticated strategies like **informational masking**. While energetic masking relies solely on the power ratio of noise to speech, informational masking utilizes phoneme-based noise that mimics the linguistic structure of human speech. This technique is particularly effective against **automatic speech recognition** (ASR) engines and human eavesdroppers, as it prevents denoising algorithms from easily separating the noise from the target signal. Furthermore, some systems implement **selective jamming**, which employs a phoneme-level detection pipeline using Recurrent Neural Networks (RNN) and Hidden Markov Models (HMM) to identify wake words. Once the onset of a legitimate command is detected, the system suspends jamming with negligible latency (typically under 60 ms), allowing the device to remain responsive while unintended speech remains obfuscated.



The implementation of these systems requires overcoming the directional constraints of ultrasound, which propagates in a straight line. Engineers have addressed this by developing **ultrasonic transducer arrays** arranged in ring-layouts, dome-shaped structures, or hexagonal configurations to achieve omnidirectional protection. Additionally, the use of **curved reflection layers**—such as iron woks or parabolic shields—effectively scatters ultrasonic waves to enlarge the protected area. Other designs focus on **gain suppression**, emitting high-volume signals that force a microphone to reach its **Acoustic Overload Point** (AOP), causing the signal to clip and lose its semantic features.

Finally, the hardware for these systems is often built using low-cost microcontrollers, such as the **STM32 or Raspberry Pi**, making them suitable for portable or wearable form factors. System performance is typically evaluated using metrics such as **Perceptual Evaluation of Speech Quality** (PESQ), **Word Error Rate** (WER), and **Speech Recognition Vocabulary Accuracy** (SRVA). These engineering efforts collectively provide a robust and non-intrusive physical-layer solution for safeguarding audio privacy in an increasingly connected world

II. Related Work

The research begins with InfoMasker, which addresses the limitation of traditional energetic masking by introducing **informational masking**. It utilizes **phoneme-based noise** that mimics the linguistic structure of human speech, making it robust against denoising algorithms and reducing automatic speech recognition (ASR) accuracy to below 50%. To address the physical limitations of jamming coverage, **JamSys** proposes a coverage optimization scheme using multiple ultrasonic arrays. It utilizes the **Modified Harmony Search Algorithm (MHSA)** to determine the optimal location and number of arrays to maximize the jammed area in a given room.

A significant evolution in the field is the **Patronus** system, which introduces **selective unscrambling**. While it blocks unauthorized devices using ultrasonic-induced "scrambles," it allows authorized devices to record clearly by delivering a "secret key" (frequency sequence) that permits noise cancellation via **NLMS adaptive filters**.Building on the phoneme-based approach, subsequent work develops a proactive anti-eavesdropping system with **controlled recording privilege**, using transformer-based models to recover speech for legitimate users while maintaining high-interference noise for others.

The vulnerability of modern voice-activated systems is further explored in research on **near-ultrasonic interference**. This study demonstrates that signals in the 16 kHz to 22 kHz range can exploit the inherent properties of **MEMS microphones** to disrupt ASR algorithms effectively across various distances and phoneme structures. **MicShield** optimizes this interference for smart assistants by proposing **selective jamming**. It uses a phoneme-level detection pipeline to identify the onset of wake words (e.g., "Alexa"), suspending jamming momentarily to allow commands while continuously obfuscating unintended private speech.

Innovative hardware designs have also emerged, such as **MicFrozen**, which moves beyond simple noise injection to **active signal cancellation**. By injecting an anti-speech signal that is the inverse of the speaker's voice, it effectively neutralizes the audio at the adversary's microphone without affecting the user's local environment. Practical implementation using low-cost microcontrollers is demonstrated by a system based on the **STM32F103**, which utilizes a modular design and an ultrasonic transducer array arranged in a spherical configuration to create a conical directional sound field for privacy protection.

To overcome the directional constraints of ultrasound, researchers developed a **dome-shaped ultrasonic jammer**. This prototype uses 67 transducers arranged on a curved surface to mitigate the directivity of traditional flat parametric speakers, significantly extending the range and angle of interference. Other specialized research, such as **DARE-GP**, focuses on protecting privacy against **Speech Emotion Recognition (SER)**. It uses genetic programming to generate **Emotion Obfuscating Noises (EONs)** that perturb spectral features used for emotion detection while preserving the intelligibility of the speech transcript.

Recent advancements also include **representation learning** for audio privacy, which integrates source separation and robust adversarial learning to filter privacy-sensitive speech data while maintaining the performance of non-speech event detection. The **Ultrajam** system focuses on **adaptivity**, utilizing a **Differential Gain Control (DGC)** mechanism to adjust the power of jamming signals across different bandwidths based on frequency response and acoustic attenuation.

Finally, the portability of these systems is addressed through **wearable microphone jamming**. This research presents a bracelet-like device that uses a ring layout of transducers and natural hand gestures to eliminate "blind spots" in jamming coverage. Complementing these portable solutions, **UltraArray** utilizes **acoustic parametric array theory** and double sideband (DSB) modulation to achieve long-distance jamming, reaching effectiveness at ranges of over 5 meters for most commercial smartphones

.

III. Proposed Methodology

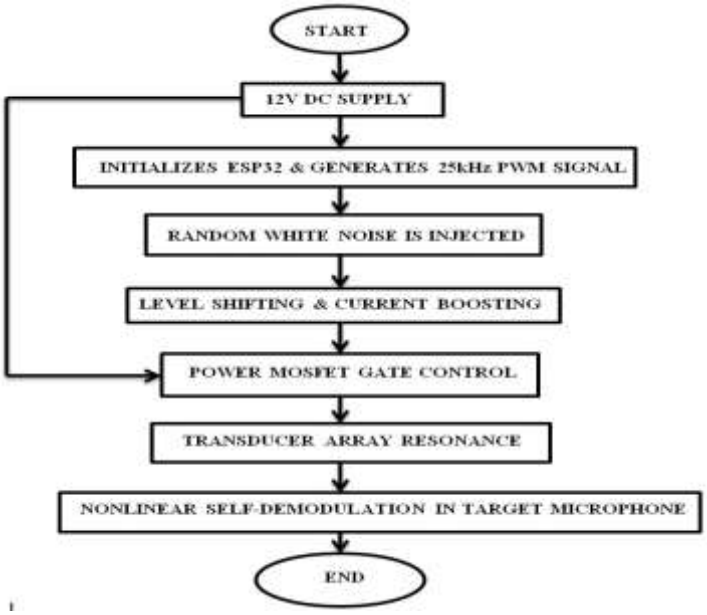


Fig 1: Flowchart of Acoustic Interference System for Audio Privacy Protection.



The operational methodology of the proposed acoustic interference system follows a strictly defined, linear workflow, transitioning from digital signal synthesis to physical acoustic disruption. The flowchart illustrates the step-by-step execution required to successfully establish a localized privacy zone and neutralize unauthorized recording devices.

The workflow initiates with the primary signal generation phase. At the core of this operation is the ESP32 microcontroller setup. Upon system activation, the ESP32 utilizes its highly precise internal hardware timers to synthesize a continuous 25 kHz Pulse Width Modulation (PWM) signal. This specific frequency is strategically selected because it operates entirely within the ultrasonic spectrum, rendering the primary carrier wave completely inaudible to human bystanders, thereby fulfilling the requirement for a non-disruptive privacy solution.

Once the digital carrier signal is generated, the workflow progresses to the amplification and emission phase. Because the microcontroller cannot output sufficient current to drive physical speakers, the signal is routed through a dedicated power amplifier stage. This amplifier is physically connected to a dense array of 20 ultrasonic piezoelectric transducers. The amplifier forces high-current electrical pulses into the transducer array, triggering rapid electro-mechanical oscillation. As a result, the array generates a highly directional, high-intensity ultrasonic sound wave that propagates through the air toward the target microphone. The utilization of a 20-transducer grid ensures that the acoustic emission is uniform, eliminating spatial blind spots within the targeted privacy zone.

The critical phase of the methodology occurs when the acoustic energy reaches the eavesdropping device. The workflow illustrates that as the directional ultrasonic waves physically strike and interact with the target microphone's internal components—specifically the diaphragm and the analog pre-amplifier—the system exploits a fundamental hardware flaw. Modern Micro-Electro-Mechanical Systems (MEMS) microphones are inherently non-linear at extreme, high-frequency sound pressure levels. Instead of simply ignoring the inaudible 25 kHz sound, the hardware is overwhelmed and inadvertently acts as an acoustic demodulator.

In the final stage of the workflow, this exploitation of hardware non-linearity yields the desired destructive result. The target's pre-amplifier circuit physically strips away the high-frequency carrier wave, forcing the hidden, low-frequency acoustic energy down into the baseband (the audible human hearing spectrum). Consequently, the targeted smartphone or recording device generates massive, overwhelming baseband jamming signals internally. This localized noise permanently buries any human speech present in the room, causing total recording corruption and successfully safeguarding the user's acoustic privacy.

IV CONCLUSION

The The development of the **"Acoustic Interference System for Audio Privacy Protection"** successfully addresses the critical privacy vulnerabilities introduced by always-on microphones in smart IoT devices,. By exploiting the inherent **hardware nonlinearity** found in commercial microphones, the system utilizes an array of **ultrasonic transducers** to emit inaudible signals that generate audible "shadow signals" within the recording circuitry, effectively masking legitimate speech,

Throughout this research, we have progressed from traditional energetic masking to implementing sophisticated techniques such as **informational masking** and **active signal cancellation**, which ensure robustness against modern denoising and signal recovery attacks. Hardware innovations, specifically the deployment of **spherical and dome-shaped transducer arrays**, have successfully overcome the directional constraints of high-frequency waves to provide broad protection coverage. Furthermore, the integration of **selective jamming** allows for continuous privacy while maintaining the responsiveness of voice-activated assistants to legitimate commands. Experimental evaluations across diverse environments confirm that these systems significantly reduce speech intelligibility for unauthorized recorders while remaining imperceptible to humans,,. Ultimately, this project establishes a portable and reliable **physical-layer defense** for safeguarding private conversations in an increasingly connected world.

VII.REFERENCES

[1] Peng Huang, Yao Wei, Peng Cheng, Zhongjie Ba, Li Lu, Feng Lin, Fan Zhang, and Kui Ren, "InfoMasker: Preventing Eavesdropping Using Phoneme-Based Noise", 2023, IEEE International Conference on Computer Communications (INFOCOM), DOI: 10.1109/INFOCOM48880.2022.9796708

[2] Hao Shen, Weiming Zhang, Han Fang, Zehua Ma, and Nenghai Yu, "JamSys: Coverage Optimization of a Microphone Jamming System Based on Ultrasounds", 2019, IEEE Access, DOI: 10.1109/ACCESS.2020.2918261

[3] Lingkun Li, Manni Liu, Yuguang Yao, Fan Dang, Zhichao Cao, and Yunhao Liu, "Patronus: Preventing Unauthorized Speech Recordings with Support for Selective Unscrambling", 2020, Proceedings of the 20th ACM Conference on Embedded Networked Sensor Systems (SenSys), DOI: 10.1145/3560905.3568528

[4] Peng Huang, Yao Wei, Peng Cheng, Zhongjie Ba, Li Lu, Feng Lin, Yang Wang, and Kui Ren, "Phoneme-Based Proactive Anti-Eavesdropping with Controlled Recording Privilege", 2024, IEEE Transactions on Information Forensics and Security, DOI: 10.1109/TIFS.2024.3353594

[5] Forrest McKee and David Noever, "Safeguarding Voice Privacy: Harnessing Near-Ultrasonic Interference", 2023, arXiv pre-print (Computers and Society), DOI: 10.48550/arXiv.2307.03152



[6] K. Sun, C. Chen, and X. Zhang, "“Alexa, Stop Spying on Me!”: Speech Privacy Protection Against Voice Assistants," in *Proceedings of the 18th ACM Conference on Embedded Networked Sensor Systems (SenSys)*, 2020.

[7] Ming Gao, Yike Chen, Yajie Liu, Jie Xiong, Jinsong Han, and Kui Ren, "Cancelling Speech Signals for Speech Privacy Protection against Microphone Eavesdropping", 2023, Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies (IMWUT), DOI: 10.1145/3631435

[8] Chuan Yang and Wenge Huang, "Design of a Novel Audio Shielding System Basedon STM32", 2025, 5th International Conference on Artificial Intelligence, Automation and High Performance Computing (AIAHPC),DOI: 10.1109/AIAHPC66801.2025.1129028

[9] Noriyuki Hayasaka and Kazuhiro Kondo, "Dome-Shaped Ultrasonic Jammer for speech privacy protection", 2023, Inter-Noise 2023

[10] Brian Testa, Yi Xiao, Harshit Sharma, Avery Gump, and Asif Salekin, "Privacy against Real-Time Speech Emotion Detection via Acoustic Adversarial Evasion", 2023, Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, DOI: 10.1145/3610926

[11] Diep Luong, Minh Tran, Shayan Gharib, Konstantinos Drossos, and Tuomas Virtanen, "Representation Learning for Audio Privacy Preservation Using Source Separation and Robust Adversarial Learning", 2023, IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA), DOI: 10.1109/WASPAA58266.2023.10248123

[12] Zhicheng Han, Jun Ma, Chao Xu, and Guoming Zhang, "UltraJam: Ultrasonic Adaptive Jammer based on nonlinearity effect of microphone circuits", 2023, High-Confidence Computing, DOI: 10.1016/j.hcc.2023.100129

[13] Yuxin Chen, Huiying Li, Shan-Yuan Teng, Steven Nagels, Zhijing Li, Pedro Lopes, Ben Y. Zhao, and Haitao Zheng, "Wearable Microphone Jamming", 2020, Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI), DOI: 10.1145/3313831.3376133

[14] Xiaosong Ma, Yubo Song, Shang Gao, Bin Xiao, and Aiqun Hu, "You Can Hear But You Cannot Record: Privacy Protection by Jamming Audio Recording", 2021, IEEE International Conference on Communications(ICC),DOI: 10.1109/ICC42927.2021.9500456