Audio Surveillance

Silas Wang

Abstract– With the rise of technology, audio surveillance and acoustic attacks have had a profound impact on the development of modern society. Most notably, with the rise of voice assistants and the growing privacy implications surrounding those products. This paper is a historical survey on acoustic surveillance technology and acoustic weapons.

Keywords: Audio surveillance, microphone jamming, acoustic attacks

I. History

Since the birth of ancient civilizations, the need for military espionage has been discussed and documented at length. Some instances of this include in *The Art of War*, where Chinese philosopher Sun Tzu claims "One who knows the enemy and knows oneself will not be endangered in one hundred engagements," hinting at the need for an intelligence department. Additionally, Indian statesman and philosopher Chanakya also wrote for the need for a society to secretly collect and process information to maintain and expand the state's power and security. As a result of this, many ancient civilizations had their own unique policies and procedures on spying. [1]

Sound related attacks were utilized in ancient times for psychological manipulation or to gain a strategic advantage in warfare. A good example of this takes place in the fourth century B.C. when Alexander the Great first encountered war elephants, which sent his horses into hysteria from their trumpeting. Similarly, during his conquest of India in 326 B.C, his men took advantage of the elephant's poor eyesight and sensitive hearing by playing dissonant sounds to send the elephants fleeing. In Ancient China, horseback archers of the Steppes would carry "screaming arrows" that would whistle through the air, terrifying their enemies. China's usage of gunpowder to create booming sounds in A.D. 250 have also been documented in ancient military manuals to disorient their foes. [2]

A. Wiretapping

With the introduction of wired communication comes a slew of methods to incercept the communication signals that flow through them. During the American Civil War, military signal operators such as William Foster and Charles Gaston were hailed as heroes for their work in wiretapping and disinformation. Since the inception of Morse Code and the telegraph, there has been security concerns relating to information being transmitted electronically. Samuel Morse, for instance, has proposed in 1837 to bury communication lines to prevent wiretapping. [3]

Even when wired communication has found its footing within American society, the first American to be convicted for wiretapping would come long after the telegraph's inception. D.C. Williams was a stockbroker in California who actively intercepted signals from manufacturing firms and mining companies in the Sacremento and San Francisco region. Whenever he came across confidential information– a pending patent application, a quote, sale-in-process, he would notify a natioanl network of stockbrokers who would make financial decisions based on Williams' statements. Eventually, an anonymous tip to federal agents would shut down his operation, leading to the arrest, trial, and sentencing for his actions.

Reporters at the time hailed this sentence as a landmark decision, proclaiming that "a new chapter in crime" has dawned society. [3] The aftermath of this decision marked the beginning of the U.S. government's increased outreach in surveillance on its constituents. Until the 1920s, wiretapping was only used by as a tool to root out union activity. But by the Prohibition Era, wiretapping has become a principle tool for investigating crime–and a stream of revenue in some places like New York–as "private ears" would be hired to covertly spy on spouses. [4] While the public recognized that wiretapping was a necessary evil for the sake of national security, but wiretapping to enforce criminal law or as a "private ear" was looked down upon. By the time the Watergate Scandal and Edward Snowden's whistleblowing was made public, the public began becoming far more worried about the government covertly listening to them. [4]

B. The Thing

Recognized as one of the earliest prototypes of the audio bugs and the predecessor of radio frequency identification, The Thing was a gift from the Soviet Union to W. Avrell Harriman, the US Ambassador to the Soviet Union on August 4, 1945. Developed by Leon Theremin, The Thing was comprised of a nine-inch antennae connected to a capacitive membrane hidden inside a wooden frame of the U.S. Seal. Select reports claim that the device itself was underneath the beak of the eagle on the seal.

Fig. 1: A replica of The Thing hidden within a wooden seal [Source: [5]]

Alongside the antennae lies some insulation and a cavity with a tuning post, giving the ability for the Soviets to

passively amplify the signal that the antennae would send out. The combination of the membrane and the tuning post created a condenser microphone whose signal would get amplified by the cavity through resonance, and transmitted by an antennae to a receiver. The receiver would then send power at the cavity's resonant frequency, powering the device and transmitting audio. This device was accidentally discovered in 1951 by a British radio operator who intercepted an American conversation on an open Soviet Air Force channel. As a result of this, numerous American agencies including the Central Intelligence Agency, the Federal Bureau of Investigation, and the Naval Research Laboratory conducted investigations, analyzed the device, and ran secret research programs to develop their own passively powered covert listening device. A similar system known by the codename SATYR was inspired by The Thing. It was developed and used by the British, Australians, Americans, and Canadians throughout the 1950s.

C. Modern Acoustic Technologies

With the introduction of advanced technology, modern acoustic attacks and devices have expanded the effectiveness of sound-based devices through its applications outside of warfare or its change in lethality through the use of technology. This following section is a collection of acoustic devices and developments from the 20th century onwards.

During World War II, Nazi Germany was developing a sonic weapon known as the "Schallkanone," which roughly translates to sound cannon. This device consisted of large parabolic reflectors that was connected to a methane-oxygen combustion chamber, which upon combustion would release focused pressure waves of roughly 44Hz. It was reported that from 50 meters away, roughly 30 seconds of exposure would kill a human, while between 200-300 meters, it will cause severe nausea, pain, and vertigo by vibrating critical organs like the inner ear. Between 50-200 meters, the pressure waves could harm intractable organs like the kidneys and liver. Despite its lethality, the Schallkanone was never deployed on the battlefield due to its immense size and the inability for low-frequency pressure waves to be directionalized.

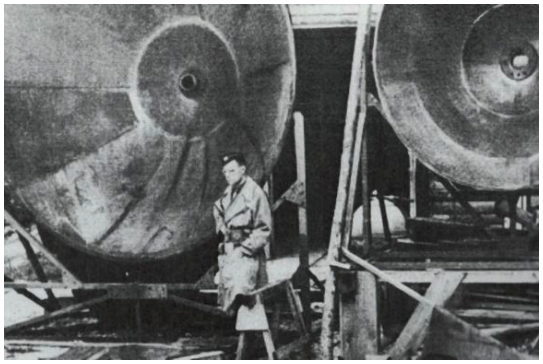

Fig. 2: An image of the German Schallkanone (sound canon) [Source: [6]]

The United States deplooyed a propaganda campaign on the Viet Cong forces known as Operation Wandering Soul during the Vietnam War. It was an attempt to lower morale and increase defections psychologically. Audio clips were broadcasted throughout the jungle using helicopters carrying speakers playing audio files with eerie recordings and altered voice recordings of South Vietnamese soldiers representing a "Wandering Soul." In Vietnamese culture, Wandering Souls represent the souls of deceased who were improperly buried and thus continue to wander the Earth. Audio recordings such as the well-known Ghost Tape Number 10 were played throguhout the night to keep the Viet Cong awake. Results were reportedly mixed, as while the Viet Cong easily shot down the loudspeakers that played these recordings, it did inevitably reveal their positions.

The development of Sonar, Radar, and Lidar has been heavily influenced by warfare, but its applications even extend beyond the premises it was built for. Sonar, for instance was devised by American engineers in the 1940s as a way to detect objects underwater through measuring how long a sound wave travels to the system. Sonar can be categorized into active and passive systems. The former sends sound wave "pings" and calculates the distance based on when the system hears a reflection of the ping, while the latter can determine the direction of an object based on the acoustic makeup the system picks up. Conceptually, sonar is very similar to radar and lidar; which uses radio and light waves instead of sound waves respectively. Radar, Sonar, and Lidar have many applications that extend beyond the military, including mapping ocean floors, weather prediction, and computer vision for autonomous vehicles.

D. LRADs

The Long-Range Acoustic Device (LRAD) is an incredibly directional speaker array capable of emitting high power, intelligible, and directional sound through beamforming and acoustic lensing. Applications within LRADs involve crowd control, long distance communication, and as a sonic weapon. By nature of the Inverse Square Law, implying the need for high sound pressure levels to be generated for long distance communication. Devices from the American LRAD company Genesys reportedly are capable of broadcasting 160dB-SPL of sound energy at one meter away, enabling for navies to communicate with vessels thousands of meters away. Some LRADs have a siren functionality that have been used as a sonic weapon for crowd control. Sound transmission will emit high power tones between 2000 - 4000Hz, the human ear's most sensitive frequency range that carries most speech intelligibility. At high sound pressure levels, LRADs can cause nausea, dizziness, and hearing damage. LRADs have been deployed throughout developed nations like the United States during protests to broadcast warning messages without disrupting the surrounding neighborhood. Recent allegations regarding the use of an LRAD in Serbia have surfaced after protesters at an anti-government rally in Serbia briefly heard "frightening sounds from hell," causing a small stampede amongst the protestors.

The directionality of an LRAD can be created through beamforming and are assessed through measuring the angle by which the cone of sound forms upon playing a 1-2kHz tone. The principle of beamforming lies within a phased

parametric array of piezoelectric ultrasonic transducers where each transducer playes a signal that is uniquely out of phase. As a result of this, the sum of every transducer signal will form a constructive interference at a given distance, broadcasting a directional sound wave where only members within the cone of constructive interference can hear the sound wave.

II. Methods of Audio Surveillance

To precede a discussion of the design and evluation of a microphone jammer, it is important to recognize the applications that an ultrasonic microphone jammer can be used for. As a result of this, the following section outlines select research developments and technologies that have applications in audio surveillance to show the importance of data privacy and the motivation for using a device such as the ultrasonic microphone jammer.

A. Microphones

Microphones are a cornerstone piece of technology utilized to capture high quality sound, with applications in audio surveillance. A key characteristic of any microphone lies within its polarity. That is, its sensitivity to sound pressure waves based on the direction at which the sound wave hits the microphone diaphragm. While pressure-gradient based diaphragm designs such as the cardioid capsule have been very common within disciplines such as audio engineering and sound for film, the ability for directional microphones to record an isolated polar direction has led to the development of various hyper-directional microphones, such as the phased microphone array and the parabolic microphone. [7]

The parabolic microphone consists of a microphone whose diaphragm is pointed into a bowl-shaped reflector. The position of the microphone is based on the focal point, described as $f = \frac{r^2}{4d}$, where r is the radius of the reflector, and d is the depth [8]. Because sound waves that reflect into a parabolic reflector will converge at the focal point, a microphone that picks up the sound that converges at the focal point will create a highly sensitigve long-range microphone. The frequency response of the parabolic microphone depends on its reflector, as a typical 1m parabolic dish will struggle to directionally capture sounds with a wavelength $\lambda > 30$cm, roughly equating to sounds below 1 kHz. [9]

The phased microphone array is an array of microphones that is specifically arranged such that each microphone in the array is out of phase. As a result of this, every microphone recording an incoming sound wave can be directionalized through delay compensating each out-of-phase microphone, revealing the general direction of a given object or person. [10]

An idea that reportedly originated from Leon Theremin's work with the Soviet surveillance system *Buran*, the laser microphone can detect sound vibrations on reflective objects and recreate the vibrations by picking up the laser's signal from a receiver. If an invisible low-power infrared laser is used, an adversary can covertly spy on a victim without there being any alterations within the room. [11]

B. Acoustic Cryptanalysis

Alongside the steadfast development of technology comes novel ways to extract information from audio files of inherently mundane tasks. For instance, a proof-of-concept paper was published documenting the design and implementation of a machine learning system that listened to an audio file of handwriting, and recognizing written words with roughly 50-60% accuracy [12]. Similar research has been conducted with keylogging, in which a machine learning model can be trained on the sounds of a keyboard's keystrokes, which can then correctly identify word with roughly 95% accuracy. [13]

With the rising popularity of smart voice assistants, publications have revealed potential vulnerabilities in devices that interface with voice assistants such as Amazon Alexa, Siri, and Cortana. Because of their "wake-up" system that only activates the voice assistant when a certain word or phrase is detected, there has been a lot of discussion on solutions that would prevent a voice assistant from constantly recording information–and by extension, providing an outlet of information for adversaries to eavesdrop to. [14]

References

[1] ""History of Espionage"." Accessed: Aug. 20, 2025. [Online]. Available: https://en.wikipedia.org/wiki/Laser_microphone

[2] A. Mayor, "The Use of Sound as Strategy in Ancient Medieval Warfare," *Brewminate*, Aug. 2022, Accessed: Aug. 20, 2025. [Online]. Available: https://brewminate.com/the-use-of-sound-as-strategy-in-ancient-and-medieval-warfare/

[3] B. Hochman, *The Listeners: A History of Wiretapping in the United States*. Cambridge, MA: Harvard University Press, 2022.

[4] A. White, "A Brief History of Surveillance in America." Accessed: Aug. 20, 2025. [Online]. Available: https://www.smithsonianmag.com/history/brief-history-surveillance-america-180968399/

[5] ""The Thing (listening device)"." Accessed: Aug. 20, 2025. [Online]. Available: https://en.wikipedia.org/wiki/The_Thing_(listening_device)

[6] "Sound Cannon." Accessed: Aug. 20, 2025. [Online]. Available: https://www.nevingtonwarmuseum.com/sound-cannon.html

[7] H. Robjohns, "Understanding & Using Directional Microphones." Accessed: Aug. 20, 2025. [Online]. Available: https://www.soundonsound.com/techniques/understanding-using-directional-microphones#:~:text=After%20a%20pressure%20wave%20arrives,rear%2C%20creating%20a%20pressure%20gradient.

[8] "The complete guide to parabolic microphones." Accessed: Aug. 20, 2025. [Online]. Available: https://diymics.com/parabolic-microphones/

[9] "Parabolic microphone." Accessed: Aug. 20, 2025. [Online]. Available: https://en.wikipedia.org/wiki/Parabolic_microphone

[10] "Phased-microphone array." Accessed: Aug. 20, 2025. [Online]. Available: https://nrc.canada.ca/en/research-development/nrc-facilities/phased-microphone-array

[11] "Laser Microphone." Accessed: Aug. 20, 2025. [Online]. Available: https://en.wikipedia.org/wiki/Laser_microphone

[12] T. Yu, H. Jin, and K. Nahrstedt, "Audio based Eavesdropping of Handwriting via Mobile Devices," pp. 444–455, 2016, doi: 10.1145/2971648.2971765.

[13] "Acoustic Cryptanalysis." Accessed: Aug. 20, 2025. [Online]. Available: https://en.wikipedia.org/wiki/Acoustic_cryptanalysis

[14] G. Zhang, C. Yan, X. Ji, T. Zhang, T. Zhang, and W. Xu, "DolphinAttack: Inaudible Voice Commands," in *Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security*, Dallas, Texas, USA: Association for Computing Machinery, 2017, pp. 1039–1052. doi: 10.1145/3133956.3134052.