https://doi.org/10.31891/2307-5732-2025-359-130
УДК 004.056

LYTOVCHENKO OLEKSII
Cherkasy State Technological University, Cherkasy, Ukraine
https://orcid.org/0000-0002-2405-042X
email: o.v.lytovchenko.asp22@chdtu.edu.ua
LAVDANSKYI ARTEM
Cherkasy State Technological University, Cherkasy, Ukraine
https://orcid.org/0000-0002-1596-4123
email: a.lavdanskyi@chdtu.edu.ua

REVIEW OF THE MODERN METHODS OF MASKING OF THE ACOUSTIC SPEECH SIGNALS FOR THE PRIVACY PURPOSES

In today's world, the importance of protecting speech privacy spans from safeguarding personal conversations to securing national interests. Among the effective strategies for mitigating speech eavesdropping is acoustic speech masking, which introduces background noise to reduce the intelligibility of private conversations. This article reviews and compares various modern methods of the acoustic speech signals masking for privacy purposes, with a primary focus on their effectiveness, subjective comfort, and complexity of implementation. The article begins the review from stationary colored noise maskers, white and pink noise, emphasizing the latter's favorable spectral characteristics and broader adoption in office environments due to its deeper, less intrusive sound. Then, the discussion turns to the speech-shaped noise (SSN) technique, which refines this approach by matching the spectrum of typical human speech, yielding equivalent privacy at lower sound pressure levels while being perceived as more natural. Further, the babble noise masking approach is reviewed, a mixture of multiple simultaneous human voices, is examined for its capacity to combine both energetic and informational masking, providing enhanced effectiveness over previously reviewed masking noises. Besides that, the article reviews adaptive speech-dependent masking and ultrasonic jamming approaches. The paper analyzes masking performance, annoyance levels, and deployment feasibility for each of the discussed approaches, highlighting the trade-offs between masking efficiency, comfort and simplicity. The paper concludes that while pink and SSN maskers are widely used in commercial environments, advanced methods like babble and adaptive maskers offer improved performance but are less common due to complexity.

Keywords: *speech masking, speech privacy, masking noise, colored noise, ultrasonic jamming.*

ЛИТОВЧЕНКО ОЛЕКСІЙ, ЛАВДАНСЬКИЙ АРТЕМ
Черкаський державний технологічний університет

ОГЛЯД СУЧАСНИХ МЕТОДІВ МАСКУВАННЯ АКУСТИЧНИХ МОВЛЕННЄВИХ СИГНАЛІВ З МЕТОЮ КОНФІДЕНЦІЙНОСТІ

У сучасному світі важливість захисту конфіденційності акустичного мовленнєвого сигналу охоплює широкий спектр потреб: від захисту особистих розмов до захисту національних інтересів. Серед ефективних стратегій запобіганню прослуховування є акустичне маскування мовлення, що додає фоновий маскувальний шум для зниження розбірливості приватних розмов. У цій статті розглядаються та порівнюються сучасні методи маскування акустичних мовленнєвих сигналів з метою забезпечення конфіденційності, з акцентом на їхню ефективність, суб'єктивний комфорт учасників розмов та складність втілення. Стаття починає розгляд зі стаціонарних "кольорових" шумів-маскувачів, білого та рожевого шумів, підкреслюючи сприятливі спектральні характеристики рожевого та більш ширше застосування в офісних умовах через його глибше та менш нав'язливе звучання. Далі обговорюється маскувальний шум, що має форму мовленнєвого сигналу (SSN), що є удосконаленням попереднього підходу, нагадуючи спектр людської мови та забезпечуючи еквівалентну конфіденційність при нижчих рівнях звукового тиску, завдяки чому даний маскувальний шум сприймається як більш природний. Наступним розглядається використання маскувального шуму, що базується на суміші кількох одночасних людських голосів (babble noise). Даний шум досліджується на предмет його здатності поєднувати як енергетичне, так і інформаційне маскування, забезпечуючи підвищену ефективність порівняно з попередньо розглянутими маскувальними шумами. Крім того, у статті розглядаються адаптивні підходи до маскування (такі, що залежать від мовленнєвого акустичного сигналу) та ультразвукового маскування. Для кожного з обговорюваних у статті підходів, аналізується ефективність маскування, рівень "подразнення", що спричиняють дані методи маскування, та можливості щодо розгортання, підкреслюючи компроміси між ефективністю маскування, комфортом та простотою. У статті робиться висновок, що хоча маскування на основі рожевого шуму та SSN широко використовуються в комерційних приміщеннях, більш досконалі методи, такі як маскування з використанням babble noise та адаптивні підходи до маскування, пропонують більшу ефективність, але є менш поширеними через свою складність.

Ключові слова: *маскування мовлення, захист мовлення, маскуючий шум, кольоровий шум, ультразвукове глушіння.*

Стаття надійшла до редакції / Received 19.09.2025
Прийнята до друку / Accepted 05.11.2025

Introduction

In the modern era, speech privacy has evolved beyond a matter of mere inconvenience - it now encompasses risks ranging from reputational harm to threats to national security. Addressing these risks is essential and demands the implementation of specialized measures, such as purpose-built architectural designs, advanced sound insulation, and devices that employ acoustic masking algorithms.

One of the speech security measures, acoustic speech masking, adds background sound to cover private conversations so eavesdroppers have trouble understanding them. The good result of applying this security measure means an unintended listener cannot identify words or meaning of speech.

According to [1] and [2], the "cover" principle is exactly this: emit a masking noise continuously in the room to hide conversation. Open-office studies show many modern offices lack sufficient privacy. Sound masking systems

(commercial units) aim to increase privacy and reduce intelligibility at typical hearing distances.

In open-plan settings, guidelines [4] suggest ambient masking levels on the order of 40–48 dBA sound pressure level (SPL) depending on usage (around 38 - 43 dBA in private offices, 44–48 dBA in open areas). These levels correspond roughly to just-barely-discernible background sound. The goal is to reduce the speech-to-mask signal-to-noise ratio so that normal conversation cannot be clearly understood more than a few feet away.

This article reviews acoustic speech signals masking techniques and devices used to block unwanted listening. First, the article describes conventional stationary noise masking (white/pink colored noise), then "speech-shaped" static noise, then "babble" (multi-talker) noise, followed by speech-dependent adaptive masking ("seeded" noise), and finally ultrasonic jamming approaches. For each, this article discusses masking effectiveness, annoyance factor, implementation complexity, and the performance. While some solutions are widely used in offices (e.g. pink noise emitters), others are research-grade (e.g. phoneme-aware ultrasonic jammers) and aim at high-security environments.

Review and analysis

Stationary colored-noise masking. The simplest approach to mask the speech is to apply a stationary random noise with a fixed spectrum.

White noise has equal power per unit frequency, which is acoustically perceived by humans as a "hiss". Pink noise has equal power per octave (~ -3 dB/octave). The pink noise rolls off in the treble part of the spectrum, so it is perceived as a "deeper" sound and less "shrill".

In practice, pink noise is widely chosen for sound masking [1, 3, 5]. The pink noise spectrum provides more low-frequency energy (similar to human voice) and avoids the sharp high-frequency "hiss" of white noise. Cambridge Sound Management notes that white noise "sounds 'hissy'", and is generally a poor choice for the speech masking, whereas pink noise is subjectively quieter while still covering the speech band [1, 5].

In fact, in a controlled study subjects rated white and pink equally annoying when loud enough to mask speech, but pink was slightly preferred. Still, the optimum masking sound actually matches the speech spectrum [5, 6].

In contrast to the white noise, the pink noise, which has about 3 dB/octave decay, reduces treble energy and is therefore preferred for speech masking [5]. Stationary noises are relatively trivial to generate. For example, a relatively simple digital signal processing (DSP) algorithm can generate the white noise and pass it through a $1/f$ filter to produce the pink noise. Variants like brown noise (which has -6 dB/octave roll off) have very deep rumble, but few studies compare them for speech masking.

The aforementioned types of noise mask human speech by utilizing the principle of energetic overlap within the speech frequencies. Speech intelligibility studies show that adding a broadband noise raises the signal-to-noise ratio (SNR) from the listener's perspective. In simple words, to achieve "moderate"-level privacy is it needed to reduce the SNR of speech at a distant listener to near 0–5 dB. For example, if normal speech at 1 m is ~65 dBA, raising ambient noise to ~55–60 dBA will make it hard to follow. In practice, a well-implemented sound masking system raises background to ~40–48 dBA, which for typical office layouts can cut intelligibility to minimal levels beyond a few meters [1, 4].

Noises with more low-frequency content provide more energetic masking per dB at frequencies where speech has energy. In [9] authors found that speech-shaped noise (discussed next) produces the same distraction at ~3 dB lower level compared to a standard -5 dB/octave noise like pink.

In summary, pink noise requires slightly less power than white to achieve a given loss of intelligibility, but both are comparable as energetic maskers.

Stationary noises can be intrusive if too loud. White noise at privacy levels often sounds harsh; pink noise is generally deemed less strident [5]. Objective measures (speech privacy indices) neglect listener comfort, but surveys show that occupants dislike masking sounds that vary or hiss. According to [6], listeners found white and pink noise equally annoying when both were loud enough to mask speech. Cambridge experts observe that even pink noise is "hissy" at high volume, and that the best sound for privacy is one shaped like speech [5].

Thus in practice a compromise volume and spectrum are chosen. A slight amplitude modulation or use of very low frequencies (like waterfall sounds) can make masking more tolerable, though stability (steady sound) is usually desired.

The white/pink noise generators are relatively the simplest from the reviewed in the article methods. They either can be analog circuits or digital audio loops. Hardware sound masking systems typically store a pink noise sample, or generate it on the fly. From the computational point of view, this process is trivial. However, deployment of this approach requires a calibrated speaker array to achieve uniform coverage [4]. In contrast, white noise played from a single source (e.g. a speaker on a desk) will not mask well beyond immediate vicinity. In modern offices, fixed overhead panels emit the noise. When it comes to this approach, the fine-tuning of levels and spectrum is the complex part.

Stationary masking noises treat all speech equally. Adding more talkers mainly raises total speech energy. Two simultaneous talkers at equal level give ~3 dB more overall speech sound, effectively reducing SNR by 3 dB if the noise is fixed.

In practice, an office with many speakers often already has higher ambient babble, so a baseline masker is set to suppress that combined noise. Designers usually set masking for worst-case occupancy (e.g. fully occupied call center).

There is little published data on precise performance vs. number of talkers, but one can infer that each doubling of concurrent speech requires ~3 dB more mask level to maintain the same privacy. The important part here which must be taken into account is that the stationary masking does not adapt per speaker – ones simply raise the general noise floor.

Speech-shaped noise (SSN) masking. This approach can be viewed as a refinement to approaches mentioned above. Speech-shaped noise (SSN) is a steady random signal whose ***long-term*** spectrum matches that of a typical human voice.

For example, in [2] and [9] authors generated masking noise by filtering white noise to follow the long-term average speech spectrum.

This concentrates masking energy where the ear expects speech, improving efficiency. Empirical studies confirm the advantage: one study found that a speech-shaped masker could achieve the same intelligibility reduction at ~3 dB lower level than standard pink noise [9].

In other words, SSN produces a given privacy effect while leading to less acoustic pollution. In [2] authors conclude SSN "results in better masking power when the background includes speech".

Unlike babble, SSN is still stationary (it has no intelligible content), but its spectral coloring makes it more potent. It sounds smoother and more "natural" than flat white noise; some people perceive it as resembling a waterfall or fan hum. Crucially, Cambridge Sound notes that optimal masking sounds often mimic the speech spectrum to blend with the environment [5]. Recent work suggests using lower-pitch SSN to reduce annoyance. In [2] it is shown that lowering the fundamental frequency of the SSN leads to significantly less annoyance while retaining masking effectiveness.

SSN is generally more effective than generic pink noise. For example, if 45 dBA of the pink noise gave 80% speech garbling at 5 m, SSN could achieve the same at ~42 dBA. It mostly still provides energetic masking (overlapping the vocal formant bands). Unlike reactive methods, SSN is not speaker-specific – it just covers the full speech frequency range statically. One must still use a high enough level to degrade all conversations. In practice SSN is engineered so that its spectrum decays about like human speech. Some advanced systems dynamically shape the noise, but basic SSN is precomputed.

Because SSN is tuned to speech, it tends to be less jarring than white noise or pink noise. By matching the timbre of voices, listeners often find it more soothing or, at least, more neutral.

Authors of [2] specifically found speech-spectrum noise to be subjectively less annoying than broadband noise while still being efficient. However, any continuous noise at sufficient level will cause some distraction.

Generating SSN is only slightly more complex than pink noise. One can start with a large sample of speech (e.g. many minutes of talk), compute its long-term spectrum, then filter white noise to match that spectrum. The required DSP is modest (real-time convolution or IIR filtering). Some systems let technicians upload voice samples to "tune" the masker to a particular language or voice. The deployment is otherwise identical to colored noise. SSN is usually provided by commercial maskers as a fixed spectrum in firmware (or an audio file). Computationally it is still very lightweight (a few filter multiplies per sample).

SSN, like other stationary methods, does not adapt to the number of speakers. It masks whichever voices contribute to the mix. If more people talk, the speech floor rises and intelligibility still drops as intended. In fact, SSN may hold up a bit better than pink when multiple voices are present, since the masking "shadows" the speech spectrum exactly.

Babble (multi-talker) noise masking. Babble noise is literally recorded or synthesized overlapping voices (e.g. 6–10 talkers at once), used as a masker [7, 8]. It is an example of informational masking, since the masker itself contains speech-like fluctuations. In [6] authors describe white noise as purely energetic versus babble as introducing informational masking that takes attention and memory.

In practice, playing a babble recording (e.g. indistinct chatter, a coffee-shop ambience) adds a realistic speech backdrop.

Babble is often more effective than white noise. In controlled tests in [6], it was found that many-voice babble masked target words about 3 dB better than white noise. A single additional concurrent talker (two voices instead of one) can itself mask 6 dB better than white noise.

In essence, a non-meaningful speech background can make it very hard to pick out one conversation. This is consistent with auditory masking theory: a fluctuating masker (like other voices) reduces the "glimpses" of target speech. Babble is thus a strong masker – arguably the strongest if it reasonably simulates the acoustic scene of overlapping talkers.

Babble noise is often perceived as more distracting than steady noise. The very reason it masks well (it sounds like speech) makes it cognitively intrusive. Interestingly, studies report that subjects find babble and other noises about equally annoying when at masking levels [6].

However, many people prefer a conversation-like humming over a high-pitched hiss. Cafeteria-like babble can sometimes seem more natural, but it may also draw attention and thus cause annoyance or stress. In any case, babble cannot be played too loudly or it becomes intelligible itself – which defeats privacy.

Besides speech masking for privacy purposes, it appears that babble noise is used sometimes as a background track in relaxation and wellness apps.

Babble masking requires storing multi-talker recordings (or on-the-fly mixing of clean speech clips). It can be as simple as looping a taped group conversation. Some systems allow mixing voices algorithmically. Computationally, it's light (just audio playback). However, one must be careful with looping and stereo; repeating the same recording can be noticeable over time. Hardware is straightforward (just audio output). There is minimal DSP complexity – the challenge is curating appropriate mask material.

Babble is multi-speaker speech, so if actual speakers in the room increase, the distinction between "mask" and "signal" blurs. If the environment already has many voices, adding babble may have diminishing returns (it just adds more similar voices). In a room where two people talk, a recording of 6-voice babble may be sufficient; if ten people talk, one would need an even louder babble or risk intelligibility. No formal scaling law is known, but one could approximate that each doubling of active talkers (and thus room intelligibility) requires ~3 dB more babble. Importantly, babble noise masking is usually applied to calm single-talk scenarios (e.g. someone on a phone in an office), not to noisy conferences.

Adaptive or speech-dependent masking. Instead of reproduction of a fixed acoustic speech masking noise, adaptive speech masking uses knowledge of the actual speech to generate more targeted masking noise. This category includes "speech-dependent" or "seeded" masking: the system listens to or predicts the voice and then emits a correlated obscuring signal.

One example is active noise cancelling of speech. Patent [10] outlines a device using non-acoustic sensors to detect a user's spoken words just before they emit into the air. A processor then inverts or scrambles those speech segments and plays them from small speakers near the face (or in the hand phone) in real time [10]. By timing and phasing these anti-speech sounds, the mask partially cancels the original voice at a nearby listener. The patent [10] describes inverting recent 10 ms pitch periods and mixing them back to produce "unintelligible distortion" – effectively lowering the speech level and garbling the message.

In practice, such a system would require tight hardware integration: precise sensors to capture subvocal vibrations, and a very fast DSP loop to generate the obscuring waveform with microsecond timing. The patent notes that voiced speech can be predicted several hundred microseconds in advance from throat motion, allowing near-perfect timing [10].

This approach is highly effective conceptually, but at present is complex and experimental. No widely available products use such methods, and performance in multi-talker settings is unproven. Similar ideas have been floated in various forms, but [10] is the most detailed.

When properly executed, adaptive masking can be very strong. The patent claims it "creates a reduction in speech-sound level at the listener and a distortion of speech meaning, leading to a reduction in recognition" [10]. In other words, it can achieve low intelligibility even at moderate volumes, because it directly attacks the signal. Compared to white noise or pink noise masking, adaptive cancellation could theoretically require much lower volume. However, it is highly speaker-specific and only masks the targeted voice. If one speaker is talking, it mainly obscures that one voice – other voices would not trigger cancellation.

Adaptive masking could be designed to emit only brief bursts of noise when speech occurs, potentially reducing overall acoustic pollution. However, if poorly done it could sound like a robotic echo or warble ("parroting" effect). Scheme in [10] primarily inverts speech itself, which listeners might perceive as bizarre distortions of the speaker's voice. There is no published study of listener annoyance for such a system.

This is the one of the most complex categories of the speech masking approaches. It requires real-time signal processing on live audio, or even sensor fusion. DSP must run at audio rates. Hardware must include extra sensors or fast mics/speakers. In practice, this might be implemented as an app or a dedicated device near the speaker's mouth. A simpler form is smartphone apps that play pink noise only when you speak, but those work poorly. The required computational complexity is high, and low-latency hardware design is critical.

Ultrasonic inaudible jamming methods. An innovative direction is to use inaudible signals to confuse microphones or eavesdroppers. Since humans typically hear up to ~20 kHz, ultrasonic (>20 kHz) or even "near-infrasonic" methods have been explored. The key principle is that many microphones have nonlinear frequency responses: an ultrasonic carrier can produce audible byproducts in the mic, jamming the recording. For example, one method is to emit a strong 20-22 kHz tone. The microphone's nonlinearity generates intermodulation noise in the speech band, effectively overlaying a rough "fuzz" without humans hearing it. Such ultrasonic jammers have been proposed to silently mask conversations.

The system in [11] uses an array of ultrasonic transducers to project "phoneme-based" ultrasonic noise into the space. The inaudible noise interacts with the microphone to insert phoneme-like disturbances. InfoMasker generates the ultrasound signal so that speech recognizers' accuracy falls below 50% even at SNR = 0 (i.e. jammed volume). This represents informational masking via ultrasonics: rather than reproducing broadband ultrasound, InfoMasker encodes specific confusing patterns to defeat denoising [11].

Another example is described in [12], which adapts the ultrasonic jammer in real time based on the detected speech to maximize interference.

Ultrasonic masking is inaudible to people, so it causes essentially no acoustic pollution. Neither InfoMasker nor simple ultrasound tones annoy nearby humans (though some dogs or electronics might be sensitive to such signals). This is a major appeal: the privacy can be achieved without raising the audible background.

Ultrasonic jammers have limited range and reliability. As [11] note, ultrasound must be very powerful to be effective – beyond a few meters the signal attenuates and returns to inaudibility. InfoMasker's author points out that existing ultrasound solutions "can remain inaudible only at limited energy and cover only short distance", meaning they easily lose effect or can be "removed" by a determined listener [11].

The ultrasound methods can be very powerful locally, but require careful design and still may be defeated by advanced recording setups. When properly aimed, ultrasound jamming can obliterate speech capture by microphones. InfoMasker claims to reduce automatic speech recovery (ASR) accuracy below 50% at SNR = 0 [11].

The [13] proposed using multiple ultrasonic arrays to cover a region and optimized their placement; their experiments showed a large area could be defended by coordinating emitters.

Audibly, the ultrasonic speech jamming approaches essentially produce no acoustic pollution. Humans don't hear tones over 20kHz, so the room stays "quiet." (in practice, any nonlinear transduction means a little "buzz" might be heard at the speakers, but this is usually negligible). The one concern is pets – many animals can hear ultrasound. Apart from that, ultrasonic jammers produce no human-disruptive noise.

These systems require special hardware. Ultrasonic jammers use high-frequency transducers or speaker arrays (more expensive than typical ceiling speakers). InfoMasker's prototype uses four 40 kHz transducers and a custom amplifier [11]. Computationally, generating ultrasound is light, but designing the waveform to confuse speech (as InfoMasker does) needs prior computation (e.g. selecting phoneme patterns).

Ultrasonic methods generally do not depend on how many people are talking – the jammer affects all microphones it reaches equally. If two people talk in a room with a single ultrasonic emitter, both voices will be equally jammed (assuming they are both near the emitter's coverage).

Summary

Acoustic methods of acoustic speech signal masking offer several trade-offs.

Simple colored noise (white or pink noise) is easy to deploy and can be relatively effective in reducing intelligibility of the speech. Pink noise is usually preferred over white noise due to more precise speech band coverage and less harsh sound. Though, the approach with colored noise as a masking noise requires to be audible to conversation participants, this may cause acoustic irritation, and not as effective as other approaches.

Speech-shaped noise refines prior approach by matching the speech spectrum. It is more efficient and tends to be somewhat more listener-friendly.

Babble noise method (overlapping voices) provides strong energetic and informational masking (often outperforming white noise, for example). On the other hand, the babble noise must be utilised with care since it can be distracting in itself.

Adaptive/real-time masks like the patented SpeechMask concept promise very high privacy by canceling or scrambling actual speech, but at great hardware complexity and with limited published real-world evaluation so far.

Finally, ultrasonic jamming can protect privacy invisibly to human ears. But it must be taken into account that the range of this group of methods can have limited area of effect. Moreover, sophisticated attack techniques may lead to overcoming these methods.

In choosing a speech masking method, one must balance effectiveness versus annoyance and cost.

Table-top white noise generators can improve privacy somewhat, but integrated office systems now typically use speech-shaped masking at calibrated levels. For government or corporate secure rooms, specialized techniques (e.g. ultrasound or phase-cancellation devices) may be justified.

Notably, most techniques are designed around a certain number of speakers – adding more simultaneous talkers generally requires a corresponding increase in mask level or multiple targeted emitters.

The literature suggests that stationary masking (pink or speech-shaped) effectively reduces intelligibility to acceptable levels when well-calibrated, albeit at the cost of raising the ambient noise floor.

Emerging methods like InfoMasker demonstrate that clever, speech-informed masks can push performance further without audible volume.

In all cases, careful design (speaker placement, spectrum, and level) is crucial to maximize privacy while minimizing occupant discomfort.

References

1. Anand S. Compromising Speech Privacy under Continuous Masking in Personal Spaces [Electronic resource] / S. Anand, Payton Walker, Nitesh Saxena. – [S. l. : s. n.]. – Mode of access: https://nsaxena.engr.tamu.edu/wp-content/uploads/sites/238/2020/10/aws-pst19.pdf (date of access: 16.06.2024). – Title from screen.

2. Arkadii Mykolaiovych Prodeus. Subjective evaluation of quality and intelligibility of speech distorted by synthesized noise [Electronic resource] / Arkadii Mykolaiovych Prodeus, Andrii Vitaliiovych Vityk, Daniil Yuriiovych Didenko // Microsystems Electronics and Acoustics. – 2017. – Vol. 22, no. 6. – P. 56–63. – Mode of access: https://doi.org/10.20535/2523-4455.2017.22.6.101929 (date of access: 16.06.2024). – Title from screen.

3. Biamp Systems I. What is Sound Masking? [Electronic resource] / Inc Biamp Systems // Biamp.com. – Mode of access: https://techpubs-review-c.biamp.com/What_is_Sound_Masking_.htm (date of access: 16.06.2024). – Title from screen.

4. Exploring the role of singing, semantics, and amusia screening in speech-in-noise perception in musicians and non-musicians [Electronic resource] / Ariadne Loutrari [et al.] // Cognitive Processing. – 2023. – Vol. 25, no. 1. – P. 147–161. – Mode of access: https://doi.org/10.1007/s10339-023-01165-x (date of access: 18.06.2024). – Title from screen.

5. Horrall T. Optimum Masking Sound: White or Pink? [Electronic resource] / Thomas Horrall. – [S. l. : s. n.]. – Mode of access: https://cambridgesound.com/wp-content/uploads/2013/02/Color-of-Noise.pdf (date of access: 18.06.2024). – Title from screen.

6. Implement of a secure selective ultrasonic microphone jammer [Electronic resource] / Yike Chen [et al.] // CCF Transactions on Pervasive Computing and Interaction. – 2021. – Vol. 3, no. 4. – P. 367–377. – Mode of access: https://doi.org/10.1007/s42486-021-00074-2 (date of access: 19.06.2024). – Title from screen.

7. InfoMasker: Preventing Eavesdropping Using Phoneme-Based Noise [Electronic resource] / Peng Huang [et al.] // Proceedings 2023 Network and Distributed System Security Symposium. – 2023. – Mode of access: https://doi.org/10.14722/ndss.2023.24457 (date of access: 19.06.2024). – Title from screen.

8. Krishnamurthy N. Babble Noise: Modeling, Analysis, and Applications [Electronic resource] / N. Krishnamurthy, J. H. L. Hansen // IEEE Transactions on Audio, Speech, and Language Processing. – 2009. – Vol. 17, no. 7. – P. 1394–1407. – Mode of access: https://doi.org/10.1109/tasl.2009.2015084 (date of access: 19.06.2024). – Title from screen.

9. MacCutcheon D. Sound masking in schools: panacea or auditory Band-Aid? - Acoustic Bulletin [Electronic resource] / Douglas MacCutcheon // Acoustic Bulletin. – Mode of access: https://www.acousticbulletin.com/sound-masking-in-schools-panacea-or-auditory-band-aid/ (date of access: 19.06.2024). – Title from screen.

10. Renz T. Auditory distraction by speech: Sound masking with speech-shaped stationary noise outperforms −5 dB per octave shaped noise [Electronic resource] / Tobias Renz, Philip Leistner, Andreas Liebl // The Journal of the Acoustical Society of America. – 2018. – Vol. 143, no. 3. – P. 212–217. – Mode of access: https://doi.org/10.1121/1.5027765 (date of access: 20.06.2024). – Title from screen.

11. Sound masking by a low-pitch speech-shaped noise improves a social robot's talk in noisy environments [Electronic resource] / Hamed Pourfannan [et al.] // Frontiers in Robotics and AI. – 2024. – Vol. 10. – Mode of access: https://doi.org/10.3389/frobt.2023.1205209. – Title from screen.

12. Speech masking and cancelling and voice obscuration [Electronic resource] / J. F. Holzrichter. – Published on 16.06.2013. – Mode of access: https://patents.google.com/patent/US20130317809A1/en (date of access: 16.02.2025). – Title from screen.

13. UltraJam: Ultrasonic adaptive jammer based on nonlinearity effect of microphone circuits [Electronic resource] / Zhicheng Han [et al.] // High-Confidence Computing. – 2023. – Vol. 3, no. 3. – P. 100129. – Mode of access: https://doi.org/10.1016/j.hcc.2023.100129 (date of access: 18.02.2025). – Title from screen.