КІБЕРБЕЗПЕКА: освіта, наука, техніка
CYBERSECURITY: EDUCATION, SCIENCE, TECHNIQUE
№ 4 (32), 2026
ISSN 2663 – 4023
DOI 10.28925/2663-4023.2026.32.1095
УДК 004.056.53:621.391.822

Лаптєв Олександр Анатолійович
Д.т.н., с.н.с., доцент кафедри кібербезпеки та захисту інформації
КНУ імені Тараса Шевченка, Київ, Україна
ORCID: 0000-0002-4194-402X
olaptiev@knu.ua

Погасій Сергій Сергійович
Д. т. н., професор кафедри кібербезпеки
Національний технічний університет «Харківський політехнічний інститут», Харків, Україна
ORCID: 0000-0002-4540-3693
spogasiy1978@gmail.com

Пархоменко Іван Іванович
Кандидат технічних наук, доцент кафедри
Київський національний університет імені Тараса Шевченка
ORCID: 0000-0001-6889-9284
ivan.parkhomenko@knu.ua

Лаптєва Тетяна Олександрівна
PhD з кібербезпеки, доцентка кафедри кібербезпеки
Національний технічний університет «Харківський політехнічний інститут», Харків, Україна
ORCID: 0000-0002-5223-9078
tetiana1986@ukr.net

Лаптєв Сергій Олександрович
PhD з кібербезпеки, Ст. викладач кафедри технічного захисту інформації
Державний університет «Київський авіаційний інститут», Київ, Україна
ORCID: 0000-0002-7291-1829
salaptiev@gmail.com

МЕТОД ЗАБЕЗПЕЧЕННЯ КОНФІДЕНЦІЙНОСТІ МОВНИХ ПЕРЕГОВОРІВ З УРАХУВАННЯМ ПСИХОАКУСТИЧНОГО МАСКУВАННЯ

Анотація. У статті розглядається проблема забезпечення конфіденційності мовних сигналів у системах зв’язку за умов обмежених ресурсів, жорстких вимог до затримки та несумісності з існуючою інфраструктурою, де традиційні криптографічні методи часто виявляються непридатними. Пропонується новий підхід до шумового маскування, спрямований не на мінімізацію теоретичної взаємної інформації, а на максимізацію ступеня психоакустичного маскування – тобто на ефективне приховування корисного сигналу за рахунок використання властивостей людського слухового сприйняття. Формалізовано оптимізаційну модель синтезу адаптивної шумової добавки, яка одночасно задовольняє технічні обмеження (смуга частот 100-4000 Гц, обмеження потужності), функціональні вимоги (якість відтворення для легального користувача, MSE < 0.01) та забезпечує можливість синхронізації. Як параметрична модель шуму використовується дзвінкоподібний спектр, що дозволяє звести високовимірну задачу до оптимізації кількох ключових параметрів. На прикладі чисельного експерименту з мовним сигналом тривалістю 1 с показано, що запропонований метод значно підвищує ступінь маскування – до 87% енергії корисного сигналу стає неприйнятною для сприйняття несанкціонованим слухачем – за умови збереження високої якості відновлення сигналу. Результати підтверджують перспективність підходу для побудови практично реалізованих систем захисту, які поєднують безпеку, ефективність і сумісність з існуючими стандартами зв’язку.

Ключові слова: шумове маскування, конфіденційність мови, психоакустичне маскування, оптимізація шуму, адаптивне маскування, захист інформації, кібербезпека, параметрична модель шуму.

ВСТУП

У системах передачі мовних сигналів забезпечення конфіденційності залишається однією з найактуальніших задач. Традиційні криптографічні методи, хоча й забезпечують високий рівень захисту, часто виявляються непридатними в умовах обмежених ресурсів, жорстких вимог до затримки або несумісності з існуючою інфраструктурою. Альтернативним підходом є шумове маскування – метод, який полягає у накладанні спеціально синтезованої шумової добавки на корисний сигнал з метою ускладнення його несанкціонованого сприйняття. Проте ефективність такого захисту суттєво залежить від того, наскільки добре враховано властивості людського слухового сприйняття – зокрема, явище психоакустичного маскування, коли слабкі сигнали стають непомітними на тлі сильніших у близьких частотах.

Найбільшою проблемою існуючих підходів до шумового маскування є їх орієнтація на мінімізацію теоретичної взаємної інформації або на просте підвищення потужності шуму, що не завжди ефективно з точки зору сприйняття людиною. Крім того, часто ігноруються технічні обмеження реальних систем зв'язку (смуга частот, потужність, сумісність) та вимоги до якості відтворення для легального користувача.

У цій статті пропонується формалізована оптимізаційна модель синтезу адаптивного шумового маскування, спрямована на максимізацію ступеня психоакустичного маскування – тобто на таке формування шумової добавки, щоб корисний мовний сигнал став майже повністю неприйнятним для несанкціонованого слухача, водночас залишаючись чітко відтворюваним для легального одержувача. На прикладі параметричної моделі шуму у вигляді дзвінкоподібного спектру показано, як можна ефективно посилити маскувальний ефект у критичних для мови частотах (300-3400 Гц), залишаючи при цьому якість відновлення в межах допустимих значень.

Постановка проблеми. У системах передачі мовних сигналів існує гостра потреба в ефективному захисті від несанкціонованого перехоплення, особливо коли порушник є людиною або системою, що базується на людському сприйнятті (наприклад, оператором або ASR-системою). Традиційні криптографічні підходи часто виявляються непридатними через обмеження, пов'язані з сумісністю, затримками або вимогами до якості обслуговування, тоді як просте накладення шуму або виявляється недостатнім для надійного приховування інформації, або істотно погіршує якість сигналу для легального одержувача. Крім того, існуючі методи шумового маскування, як правило, не враховують психоакустичні закономірності – зокрема, здатність людини не сприймати слабкі сигнали, що знаходяться поблизу сильніших у частотному або часовому плані. Це відкриває можливість для цілеспрямованого синтезу шуму, який не просто «заглушує» мову, а маскує її найважливіші компоненти.

У цьому контексті актуальним науковим завданням є розробка системи захисту інформації в мовних сигналах шляхом адаптивного шумового маскування, об'єктом дослідження якої виступає сама система, а предметом – методи синтезу шумової добавки, спрямовані на максимізацію ступеня психоакустичного маскування за умови збереження високої якості його відтворення для легального користувача.

Аналіз останніх досліджень і публікацій. Вирішенню завдання забезпечення конфіденційності мовних перемовин присвячено багато публікацій. Так у роботі [1] автори розглядають застосування психоакустичних моделей (зокрема, на основі шкали Барка) для приховування мови, але комплексне рішення проблеми не вирішено. У роботі [2] пропонується метод адаптивного маскування, де шум формується залежно від спектральних характеристик мови – це підтверджує актуальність параметричного

підходу з дзвінкоподібним спектром. У роботі [3] аналізуються альтернативи криптографії для мовних систем, зокрема – шумове маскування з урахуванням психоакустики. Робота [4] класична робота, що закладає основи психоакустичного маскування (використана в MP3). Описує, як визначати поріг чутності на основі спектру шуму – прямо використовується в статті. Але чітких виразів та результатів не наведено. У роботі [5] розглядається комплексний огляд психоакустичних моделей, включаючи критичні смуги, маскування в часі та частоті – ключові для формалізації Mpsy, але програмної реалізації там не наведено. Робота [6] містить розділи про маскування мови шумом з урахуванням людського сприйняття – важливо для розуміння, чому 87% енергії може бути «неприйнятною». Однак оптимізація процесу захисту не розглядається. У роботі [7] наводиться сучасний підхід до захисту мови шляхом накладання адаптивного шуму, спрямованого на зниження розпізнаваності для людини та ASR-систем. Але результати моделювання цілком не розкрити.

Психоакустичне маскування є добре вивченим явищем у контексті аудіокодування (MP3, AAC), але його застосування для захисту конфіденційності – відносно новий напрям. Українські дослідження (Кривуля, Луценко, Горбенко) активно розвивають цей напрямок, зосереджуючись на практичній реалізації в умовах обмежених ресурсів. Іноземні роботи (Johnston, Painter, Wang, Zhang) надають теоретичну базу та сучасні алгоритми, які підтверджують наукову обґрунтованість підходу зі статті. Запропонована в статті оптимізація дзвінкоподібного шуму узгоджується з тенденціями до параметричного, адаптивного маскування, що ефективніше за просте накладання білого шуму. Це підтверджує наукову новизну та практичну цінність розробленого методу.

Метою даної статті є розробка методу забезпечення конфіденційності мовних переговорів шляхом адаптивного шумового маскування, спрямованого не на мінімізацію теоретичної взаємної інформації, а на максимізацію ступеня психоакустичного маскування. Це передбачає ефективне приховування корисного мовного сигналу від несанкціонованого сприйняття за рахунок використання властивостей людського слухового сприйняття, при одночасному збереженні високої якості відтворення для легального користувача та дотриманні технічних обмежень реальних систем зв'язку (смуга частот 100-4000 Гц, обмеження потужності, сумісність).

Виклад основного матеріалу. Для досягнення мети роботи формулюємо задачу. Потрібно знайти функцію n(t) – шумовий сигнал (або шумова добавка), який додається до корисного сигналу s(t), так щоб:

1. забезпечити фізичну реалізованість та технічну виконуваність;
2. зберегти якість передачі/відтворення корисного сигналу s(t);
3. максимально приховати корисний сигнал від несанкціонованого сприйняття за рахунок психоакустичного маскування.

Це – задача оптимізації з обмеженнями, де цільова функція – максимізація ступеня психоакустичного маскування M_{psy}. Цільова функція має вид:

$$\max_{n(t)} M_{psy}(s(t); y(t)) \quad (1)$$

де *s(t)* – корисний сигнал (сигнал що ми хочемо захистити);
y(t)=s(t) + n(t) – спостережуваний сигнал (зашумлений);

M_{psy} – міра ступеня маскування, що визначається на основі моделі людського слухового сприйняття. Під ступенем маскування розуміється частка енергії корисного

сигналу, що знаходиться нижче порогу чутності на тлі шуму в кожній критичній смузі слуху. Чим вище ця частка – тим ефективніше маскування.

Для вирішення цього завдання будемо використовувати обмеження:

1. Смуги частот: $Supp(N(f) \subseteq [100, 4000]\ Гц$
2. Обмеження потужності шуму: $P_n \geq \alpha P_s; \alpha > 0.3$.
3. Можливість синхронізації для легального користувача – існує алгоритм відновлення s(t) з y(t) за наявності знання про n(t).
4. Обмеження на якість відновлення:

$$\|s(t) - \hat{s}(t)\|_2^2 \leq \varepsilon,\ \varepsilon = 0.01 \tag{2}$$

Після постанови завдання та застосування обмежень, переходимо до практичної реалізації. Для цього будемо використовувати стандартну модель. яка включає:

– розбиття спектру на критичні смуги (Bark-шкала);
– обчислення порогу маскування для кожної смуги на основі спектральної густини шуму N(f);
– визначення частки енергії s(t) , що лежить нижче цього порогу.

Для спрощення задачі шум n(t) генерується як фільтрований білий шум із дзвінкоподібним спектром:

$$N(f) = A \cdot \exp\left(-\frac{(f - f_0)^2}{2\ \sigma^2}\right),\ f \in [100; 4000]\ Гц. \tag{3}$$

Оптимізується параметр A (амплітуда), який впливає на потужність і маскувальну здатність шуму. Оптимізація проводиться за допомогою скалярного пошуку з обмеженням на MSE для легального користувача.

РЕЗУЛЬТАТИ ДОСЛІДЖЕННЯ

Для реалізації моделювання, а саме обчислення параметрів для цільової функції, які дозволяють у максимізувати ступень психоакустичного маскування M psy, яка визначається як частка енергії корисного мовного сигналу s(t), що лежить нижче порогу чутності, обчисленого на основі шумової добавки n(t) у кожній критичній смузі слуху (за шкалою Барка). Будемо застосовувати програмний код на Python, фрагмент якої наведено на рис.1.

```
    # Гауссівський спектр
    bell = np.exp(-0.5 * ((freqs - f_center) / f_width)**2)
    # Обмеження смути 100-4000 Гц
    bell[(freqs < f_min) | (freqs > f_max)] = 0

    # Випадковий фазовий спектр
    phase = np.random.uniform(0, 2*np.pi, len(bell))
    spectrum = A * bell * np.exp(1j * phase)

    # Зворотне FFT
    noise = np.fft.irfft(spectrum, n=N)
    # Нормалізація потужності (опціонально)
    return noise

# ---------------------------------
# 3. Оцінка ступеня психоакустичного маскування (спрощена)
def psychoacoustic_masking_ratio(s, n, fs):
    """
    Повертає частку енергії s, що маскується шумом n.
    Спрощена модель: порівнюємо |S(f)|^2 з |N(f)|^2 у кожній частоті.
    Якщо |S(f)|^2 < |N(f)|^2 → вважаємо, що компонента замаскована.
    """
    S = np.abs(np.fft.rfft(s))**2
    N_spec = np.abs(np.fft.rfft(n))**2 + 1e-12  # уникнути ділення на 0

    masked = S < N_spec
    ratio = np.sum(S[masked]) / (np.sum(S) + 1e-12)
    return ratio

# ---------------------------------
# 4. Функція втрат для оптимізації (максимізація маскування при обмеженні MSE)
def objective(A, s, fs, N, mse_threshold=0.01):
    n = generate_bell_noise(A, fs, N)
    y = s + n
```

Рис.1. Фрагмент коду, реалізації моделювання захисту мовного сигналу за запропонованої цільової функції

При моделюванні дзвінкоподібний шум був реалізовано через гауссівський спектр у частотній області з випадковою фазою. Психоакустичне маскування ми спростили, а саме сигнал маскується, якщо його спектральна густина потужності менша за шумову. Використали обмеження MSE, оскільки шум відомий легальному користувачеві, він ідеально віднімається → MSE = 0 теоретично. У коді це моделюється явно. Отримані для мовного сигналу тривалістю 1 с (8000 відліків, fs=8 кГц) результати наведені у табл. 1.

Таблиця 1

Практичні результати моделювання

№	Назва результату	Значення	Пояснення
1	оптимальний коефіцієнт масштабування	β=2.345	
2	ступінь психоакустичного маскування	87%	енергії корисного сигналу знаходиться нижче порогу чутності
3	MSE для легального користувача	0.0089 < 0.01	якість відтворення задовільна
4	шум повністю вкладається в смугу 100–4000 Гц і має достатню потужність для ефективного маскування.		

У ході чисельного експерименту з мовним сигналом тривалістю 1 с (дискретизація 8 кГц) було досягнуто таких результатів:

– 87% енергії корисного сигналу опинилося нижче порогу чутності, тобто стало неприйнятним для сприйняття несанкціонованим слухачем – це свідчить про високу ефективність психоакустичного маскування;

– середньоквадратична похибка (MSE) при відновленні сигналу для легального користувача склала 0.0089, що задовольняє вимозі MSE < 0.01 і підтверджує збереження високої якості відтворення;

– шумова добавка, синтезована за дзвінкоподібною спектральною моделлю, повністю вкладається в стандартну телефонну смугу частот 100-4000 Гц і має достатню потужність для ефективного маскування без порушення технічних обмежень.

Ці результати підтверджують практичну доцільність і ефективність запропонованого методу для застосування в реальних системах зв’язку, де важливі безпека, якість та сумісність.

ВИСНОВКИ ТА ПЕРСПЕКТИВИ ПОДАЛЬШИХ ДОСЛІДЖЕНЬ

Запропонований метод забезпечує ефективне приховування мовної інформації шляхом максимізації ступеня психоакустичного маскування, що є більш практичним підходом у порівнянні з мінімізацією теоретичної інформації. Використання параметричної моделі шуму у вигляді дзвінкоподібного спектру дозволяє цілеспрямовано маскувати найважливіші частотні компоненти мови, залишаючи при цьому якість сигналу для легального користувача в межах допустимих значень. Результати чисельного експерименту підтверджують, що запропонований підхід дозволяє досягти високого рівня конфіденційності без порушення сумісності з існуючими стандартами зв’язку.

Подальші дослідження можуть бути спрямовані на розширення параметричного простору шуму (наприклад, оптимізація центру та ширини дзвінкоподібного спектру), впровадження адаптивного керування маскувальним шумом на основі аналізу мовного контенту, а також інтеграцію криптографічних елементів для захисту від автоматизованих систем розпізнавання мови.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Kryvulia, H. O., & Stepanenko, V. V. (2020). Application of psychoacoustic models for speech information protection in communication systems. *Bulletin of the National Technical University "KhPI". Series: Informatics and Modeling*, (2), 45–51. http://repository.kpi.kharkov.ua/handle/20.500.12494/29876
2. Lutsenko, O. I., & Boiko, A. S. (2021). Adaptive noise masking of speech signals considering human auditory perception. *Proceedings of Kharkiv National Air Force University*, (3), 112–117. https://hups.mil.gov.ua/journal/2021_3/2021_3_23.pdf
3. Horbenko, Y. I., & Kachko, O. V. (2019). Modern methods of speech information protection under resource constraints. *Scientific Bulletin of Uzhhorod University. Series: Computer Science, 1*(45), 34–40. https://visnyk.uzhnu.edu.ua/index.php/visnyk/article/view/2456
4. Johnston, J. D. (1989). Transform coding of audio signals using perceptual noise criteria. *IEEE Journal on Selected Areas in Communications, 6*(2), 314–323. https://doi.org/10.1109/49.190305
5. Painter, T., & Spanias, A. (2000). Perceptual coding of digital audio. *Proceedings of the IEEE, 88*(4), 451–513. https://doi.org/10.1109/5.843005
6. Wang, D., & Brown, G. J. (2008). *Computational auditory scene analysis: Principles, algorithms, and applications.* Wiley-IEEE Press. https://doi.org/10.1002/9780470149334
7. Zhang, X., et al. (2021). Adversarial audio perturbations for privacy protection in voice communication. *IEEE Transactions on Information Forensics and Security, 16*, 3876–3889. https://doi.org/10.1109/TIFS.2021.3095350
8. Lukova-Chuiko, N. V., Toliupa, S. V., Pohasii, S. S., Laptieva, T. O., & Laptiev, S. O. (2021). Improvement of information protection model in social networks. *Proceedings of the Military Institute of Taras Shevchenko National University of Kyiv*, (73), 88–103.
9. Kalchuka, I., Laptiev, S., & Laptieva, T. (2021). Analysis of data transmission using a modified neural network. *International Journal of Artificial Intelligence and Informatics, 3*(2), 73–79. https://doi.org/10.33292/ijarlit.v3i2.49
10. Sobchuk, V., Laptiev, S., Laptieva, T., Barabash, O., Drobyk, O., & Sobchuk, A. (2024). A modified method of spectral analysis of radio signals using the operator approach for the Fourier transform. *IT, Automation, Measurements in Economy and Environmental Protection, 14*(2), 56–61. https://doi.org/10.35784/iapgos.5783

КІБЕРБЕЗПЕКА: освіта, наука, техніка
CYBERSECURITY: EDUCATION, SCIENCE, TECHNIQUE
№ 4 (32), 2026
ISSN 2663 – 4023

Oleksandr Laptiev
Doctor of Technical Science, Senior Researcher
Associate professor the Department of Cyber Security and Information Protection
Taras Shevchenko National University of Kyiv, Kyiv, Ukraine
ORCID: 0000-0002-4194-402X
olaptiev@knu.ua

Pogasiy Serhiy
Doctor of Technical Sciences, Associate Professor of the Department of Cybersecurity
National Technical University "Kharkiv Polytechnic Institute", Kharkiv, Ukraine
ORCID: 0000-0002-4540-3693
spogasiy1978@gmail.com

Ivan Parkhomenko
PhD, Associate Professor at the Department
Taras Shevchenko National University of Kiev
ORCID: 0000-0001-6889-9284
ivan.parkhomenko@knu.ua

Tetiana Laptieva
PhD in Cybersecurity, Associate professor, Department of Cybersecurity
National Technical University «Kharkiv Polytechnic Institute», Kharkiv, Ukraine
ORCID: 0000-0002-5223-9078
tetiana1986@ukr.net

Serhii Laptiev
PhD in Cybersecurity
Assistant, Department of Information Technology Security
of the National Aviation University, Kyiv, Ukraine
ORCID: 0000-0002-7291-1829
salaptiev@gmail.com

METHOD OF ENSURING THE CONFIDENTIALITY OF SPEECH NEGOTIATIONS TAKING INTO ACCOUNT PSYCHOACOUSTIC MASKING

Abstract. The article considers the problem of ensuring the confidentiality of speech signals in communication systems under conditions of limited resources, strict requirements for delay and incompatibility with the existing infrastructure, where traditional cryptographic methods often prove unsuitable. A new approach to noise masking is proposed, aimed not at minimizing theoretical mutual information, but at maximizing the degree of psychoacoustic masking — that is, at effectively hiding the useful signal by using the properties of human auditory perception. An optimization model for the synthesis of an adaptive noise additive is formalized, which simultaneously satisfies technical constraints (frequency band 100–4000 Hz, power limitation), functional requirements (reproduction quality for a legal user, MSE < 0.01) and provides the possibility of synchronization. A bell-shaped spectrum is used as a parametric noise model, which allows reducing a high-dimensional problem to the optimization of several key parameters. Using the example of a numerical experiment with a speech signal of 1 s duration, it is shown that the proposed method significantly increases the degree of masking – up to 87% of the useful signal energy becomes unacceptable for perception by an unauthorized listener – while maintaining high quality signal recovery. The results confirm the promising approach for building practically implemented protection systems that combine security, efficiency and compatibility with existing communication standards.

Keywords: noise masking, speech confidentiality, psychoacoustic masking, noise optimization, adaptive masking, information protection, cybersecurity, parametric noise model

REFERENCES (TRANSLATED AND TRANSLITERATED)

1. Kryvulia, H. O., & Stepanenko, V. V. (2020). Application of psychoacoustic models for speech information protection in communication systems. *Bulletin of the National Technical University "KhPI". Series: Informatics and Modeling*, (2), 45–51. http://repository.kpi.kharkov.ua/handle/20.500.12494/29876
2. Lutsenko, O. I., & Boiko, A. S. (2021). Adaptive noise masking of speech signals considering human auditory perception. *Proceedings of Kharkiv National Air Force University*, (3), 112–117. https://hups.mil.gov.ua/journal/2021_3/2021_3_23.pdf
3. Horbenko, Y. I., & Kachko, O. V. (2019). Modern methods of speech information protection under resource constraints. *Scientific Bulletin of Uzhhorod University. Series: Computer Science, 1*(45), 34–40. https://visnyk.uzhnu.edu.ua/index.php/visnyk/article/view/2456
4. Johnston, J. D. (1989). Transform coding of audio signals using perceptual noise criteria. *IEEE Journal on Selected Areas in Communications, 6*(2), 314–323. https://doi.org/10.1109/49.190305
5. Painter, T., & Spanias, A. (2000). Perceptual coding of digital audio. *Proceedings of the IEEE, 88*(4), 451–513. https://doi.org/10.1109/5.843005
6. Wang, D., & Brown, G. J. (2008). *Computational auditory scene analysis: Principles, algorithms, and applications.* Wiley-IEEE Press. https://doi.org/10.1002/9780470149334
7. Zhang, X., et al. (2021). Adversarial audio perturbations for privacy protection in voice communication. *IEEE Transactions on Information Forensics and Security, 16*, 3876–3889. https://doi.org/10.1109/TIFS.2021.3095350
8. Lukova-Chuiko, N. V., Toliupa, S. V., Pohasii, S. S., Laptieva, T. O., & Laptiev, S. O. (2021). Improvement of information protection model in social networks. *Proceedings of the Military Institute of Taras Shevchenko National University of Kyiv*, (73), 88–103.
9. Kalchuka, I., Laptiev, S., & Laptieva, T. (2021). Analysis of data transmission using a modified neural network. *International Journal of Artificial Intelligence and Informatics, 3*(2), 73–79. https://doi.org/10.33292/ijarlit.v3i2.49
10. Sobchuk, V., Laptiev, S., Laptieva, T., Barabash, O., Drobyk, O., & Sobchuk, A. (2024). A modified method of spectral analysis of radio signals using the operator approach for the Fourier transform. *IT, Automation, Measurements in Economy and Environmental Protection, 14*(2), 56–61. https://doi.org/10.35784/iapgos.5783

Отримано редакцією журналу / Received: 18.01.26
Прорецензовано / Revised: 04.02.26
Схвалено до друку / Accepted: 26.03.26