

А.М. ОЛЕЙНИКОВ, канд. техн. наук, П.М. ШКОПОТКО

МЕТОДИ ФОРМУВАННЯ АКУСТИЧНИХ ЗАВАД ДЛЯ ЗАСОБІВ ПРОТИДІЇ НЕСАНКЦІОНОВАНОМУ ДОКУМЕНТУВАННЮ МОВНОЇ ІНФОРМАЦІЇ

Вступ

Стрімкий розвиток і мініатюризація технічних засобів акустичної розвідки (ТЗАР), таких як диктофони та радіозакладки, значно ускладнюють завдання захисту конфіденційної інформації [1, 2]. Мовна інформація, що передається під час ділових переговорів або нарад, є потенційним об'єктом для перехоплення, що зумовлює необхідність розробки ефективних методів протидії. Зростання чутливості сучасних мікрофонів та вдосконалення алгоритмів шумозаглушення в записуючих пристроях ставлять нові виклики перед фахівцями з інформаційної безпеки.

Одним з найпоширеніших і найефективніших напрямків протидії є активні акустичні методи, що ґрунтуються на створенні маскуючих звукових завад. Ефективність цих методів безпосередньо залежить від якості сформованої завади, її спектральної відповідності сигналу мови, енергетичної ефективності, часової синхронізації структури завади та акустичного сигналу, а також просторових умов проведення процедури протидії [3].

Акустичні методи протидії є перспективними, оскільки впливають безпосередньо на основний канал витоку інформації – акустичний. На відміну від акустичних, ультразвукові методи придушення можуть бути неефективними проти диктофонів, що не сприймають ультразвуковий діапазон. Електромагнітні ж методи, спрямовані на придушення електронних компонентів диктофона, можна нейтралізувати ефективним екрануванням пристрою, що захищається, як це реалізовано в сучасних смартфонах високого класу. Таким чином, акустичний канал є найбільш універсальним і складним для нейтралізації, що робить дослідження в цій галузі актуальними.

Мета роботи – аналіз, систематизація та розрахункове порівняння методів формування акустичних завад для засобів протидії несанкціонованому документуванню мовної інформації. Для досягнення мети поставлено наступні завдання: провести огляд та класифікацію методів формування акустичних завад; розробити методіку програмного моделювання для порівняння ефективності різних типів завад; визначити найбільш ефективні методи за критерієм мінімізації якості запису.

1. Класифікація та аналіз методів формування акустичних завад

У загальному випадку методи протидії можна класифікувати за принципом впливу на засіб документування (ЗДМІ). Серед них ключове місце посідають методи фізичного придушення, зокрема акустичні, що базуються на енергетичному маскуванні корисного сигналу. Основним критерієм ефективності є зниження співвідношення сигнал/шум (С/Ш) на вході мікрофона ЗДМІ нижче порогового рівня, необхідного для розпізнавання мови [4]. Фізичною основою цього процесу є явище акустичної інтерференції, коли звукові хвилі від джерела завади та джерела мови складаються в просторі, що призводить до спотворення корисного сигналу.

До основних типів акустичних завад відносять:

1) білий шум – шум з рівномірною спектральною щільністю потужності в усьому діапазоні частот. Недоліком є низька енергетична ефективність, оскільки значна частина потужності витрачається на частоти, що не несуть ключової інформації для розуміння мови. Також створює значний психологічний дискомфорт через свою монотонність та високу загальну гучність;

2) тональна завада – вузькосмугова завада, що формується як сума синусоїдальних сигналів на ключових частотах мовного діапазону (наприклад, 800, 1500, 2500 Гц). Недолік –

наявність «проміжків» у спектрі, через які може проходити корисний сигнал. Ефективність може бути підвищена шляхом динамічного зміщення частот тонів, що ускладнює фільтрацію;

3) мовоподібна завада («багатомовний гомін») – один з найефективніших типів завад. Існує кілька основних методів її формування:

- speech-shaped noise (SSN) – білий шум, пропущений через фільтр, спектр якого відповідає усередненому спектру мови (100–8000 Гц, підсилення в діапазоні 500–2000 Гц). Ефективний для стаціонарного маскування, проте менш ефективний у паузах мовлення;

- babble / «багатомовний» шум – суміш записів кількох дикторів (3–12) з різними затримками, що створює ефект гомону натовпу. Високоєфективний для маскування мови, оскільки імітує природне акустичне середовище. Його недоліком є можливість виділення окремих фраз при недостатній кількості дикторів;

- модульований за спектральною огиною шум (envelope-modulated noise) – SSN, модульований низькочастотним сигналом (2–8 Гц), що імітує фразову структуру мови, ефективно маскуючи паузи. Це дозволяє зменшити середню потужність завади, підвищуючи комфорт;

- псевдомова (pseudo-speech) – генерація потоку беззмістовних фонем або складів за допомогою синтезаторів мови (TTS, вокодер). Володіє високою «мовоподібністю». Формування шляхом випадкового перемішування сегментів реальної мови у часовій області також є одним з методів створення такого типу завади;

4) «формантна завада» (модифікований акустичний метод) – перспективний метод. Завада формується динамічно в реальному часі шляхом аналізу поточного мовного сигналу, виявлення формант (спектральних максимумів, що несуть основну інформаційну навантаження) та генерації вузькосмугового шуму саме в цих частотних областях [6]. Це забезпечує цільове придушення найбільш інформативних компонентів мови при мінімальній загальній потужності завади. Фізичною основою методу є те, що людський слух та системи розпізнавання мови найбільш чутливі саме до формантних частот.

Крім того, для підвищення ефективності можуть застосовуватись інші типи шумів, спектр яких ближчий до мовного, ніж у білого шуму. Рожевий шум має спектральну щільність, що зменшується зі зростанням частоти (3 дБ/октаву), і його спектр найбільш близький до спектральної огиною мови, що робить його перспективним для застосування. Коричневий (червоний) шум має ще крутіший спад (6 дБ/октаву), і його енергія зосереджена переважно в найнижчих частотах, де розташовані перші, найважливіші форманти.

1.1. Синхронізація завади з мовним сигналом

Критично важливим аспектом підвищення ефективності та комфорту є синхронізація завади з корисним мовним сигналом. Ідея полягає в тому, щоб активувати генератор завади лише в моменти мовлення. Це дозволяє [1]:

- значно знизити загальну енергію, споживану системою захисту;
- підвищити психологічний комфорт, оскільки в паузах між фразами відсутній завадний сигнал;
- ускладнити автоматичне фільтрування завади, оскільки вона стає нестационарною.

Найбільш ефективною є адаптивна синхронізація, коли завада формується безпосередньо з аналізу поточної мови, що забезпечує ідеальну часову та спектральну відповідність. Для цього використовуються алгоритми детектування мовної активності (VAD – Voice Activity Detection), які дозволяють точно визначати початок та закінчення мовлення.

Критичним параметром для будь-якого акустичного методу є відстань між джерелом завади та потенційним диктофоном. Згідно із законом зворотних квадратів, зменшення відстані дозволяє значно підвищити рівень завади в точці розташування мікрофона без

збільшення вихідної потужності генератора [3]. Це означає, що розміщення джерела завади в безпосередній близькості від диктофона (на відміну від віддаленого джерела мови) є ключовим для енергетичної ефективності.

2. Матеріали та методика розрахункового дослідження

Для кількісної оцінки ефективності різних типів завад було проведено програмне моделювання в середовищі Python 3 з використанням бібліотек для обробки звуку (Librosa, SciPy) та об'єктивної оцінки якості мови (PESQ). Моделювання дозволило контролювати всі параметри в ідеальних умовах, виключаючи вплив акустики приміщення та не ідеальності обладнання.

В якості еталонного сигналу використовувався аудіозапис української мови тривалістю 20 с, що містив фрази різної складності та фонемний склад (частота дискретизації 16 кГц, роздільність 16 біт). Було програмно згенеровано чотири типи завад: білий шум, тональна завада, мовоподібна завада (шляхом випадкового перемішування сегментів мови в часовій області) та формантна завада (шляхом аналізу спектру мови методом лінійного передбачення (LPC) та генерації шуму в області формантних частот). Для коректного порівняння ефективності, що виключає вплив різниці в потужності, енергія усіх завад була нормалізована до однакового рівня перед їх накладанням на мовний сигнал.

Ефективність оцінювалася за допомогою об'єктивної перцептивної метрики PESQ (Perceptual Evaluation of Speech Quality, ITU-T P.862), яка прогнозує суб'єктивну якість мови за 5-бальною шкалою (від 1 – дуже погано, до 5 – відмінно). Для завдання придушення несанкціонованого запису нижчий бал PESQ свідчить про вищу ефективність завади. Додатково досліджувався вплив відстані на ефективність придушення шляхом моделювання затування звуку згідно із законом зворотних квадратів.[6]

Окрім оцінки якості звучання, для більш глибокого аналізу захищеності мовної інформації було введено другу метрику – STOI (Short-Time Objective Intelligibility). Це показник короткочасної об'єктивної розбірливості мови, який базується на кореляції часових огинаючих чистого та спотвореного сигналів у коротких часових інтервалах. Значення STOI лежать у діапазоні від 0 до 1, де 0 означає повну нерозбірливість (інформація захищена), а 1 – ідеальну чіткість. Використання двох метрик дозволяє оцінити не лише, наскільки "неприємним" є шум для слуху (PESQ), а й наскільки ефективно він руйнує зміст розмови (STOI).

Схема розрахункового дослідження передбачала моделювання розташування джерела звукового сигналу (1), джерела акустичної завади (2) та засобу документування мовної інформації (3) (рис. 1).

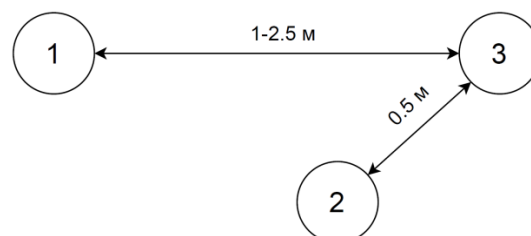


Рис. 1. Схема розташування джерела мови (1), джерела формантної завади (2) та засобу документування мовної інформації (3) при проведенні дослідження ефективності придушення несанкціонованого запису мовної інформації

Відстань між джерелом звукового сигналу та джерелом акустичної завади змінювалася в діапазоні 1–2,5 м, а відстань між джерелом акустичної завади та пристроєм документування – 0,5 м для оцінки залежності ефективності придушення від дистанції.

3. Результати дослідження ефективності придушення несанкціонованого запису мови

На першому етапі було проведено оцінювання ефективності завад за критерієм якості мови PESQ при фіксованому відношенні сигнал/шум ($SNR = -6$ дБ). Результати моделювання показали наступні значення якості (рис. 2): тональна завада – 1,68; білий шум – 1,25; мовоподібна завада – 1,22; формантна завада – 1,13.

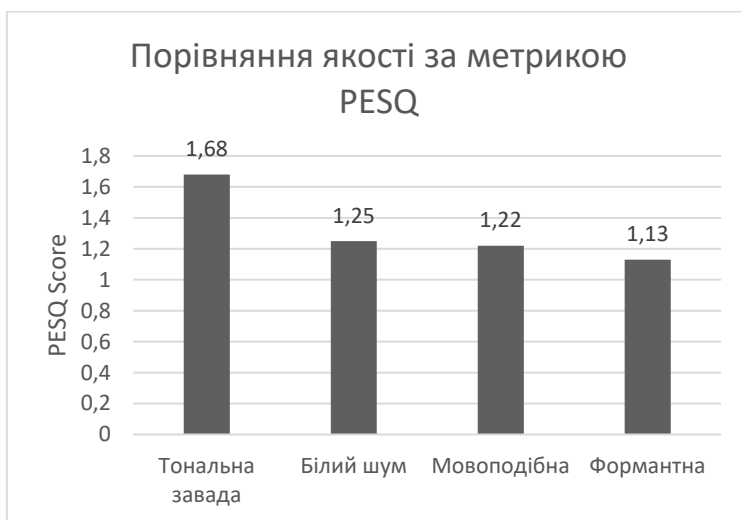


Рис. 2. Порівняння якості мови за метрикою PESQ

Як видно з отриманих даних, найменше значення PESQ (що відповідає найбільшому спотворенню) забезпечує формантна завада. Це пояснюється тим, що енергія завади зосереджена в тих самих частотних смугах, що й енергія мовного сигналу, створюючи ефект частотного маскування.

На другому етапі дослідження було проведено аналіз за критерієм розбірливості STOI, який є критично важливим для задач захисту інформації. Результати розрахунків продемонстрували ще більшу різницю в ефективності методів (рис. 3): тональна завада – 0,78 (висока залишкова розбірливість); білий шум – 0,62; мовоподібна завада – 0,51; формантна завада – 0,42.

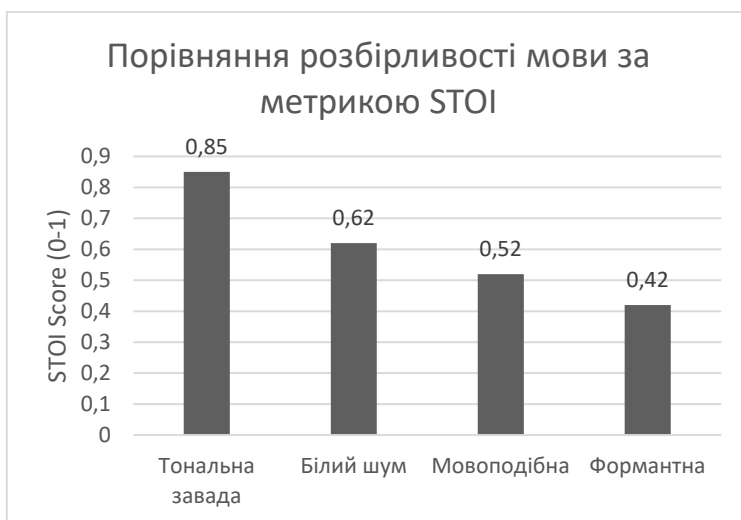


Рис. 3. Порівняння розбірливості мови за метрикою STOI

Аналіз діаграми показує, що хоча за метрикою PESQ різниця між білим шумом і формантною завадою є помірною, за метрикою розбірливості STOI перевага формантної завади є значною. Значення STOI 0.42 свідчить про перехід через поріг розбірливості, за яким відновлення змісту розмови стає неможливим без застосування спеціальних методів реставрації фонограм.

Окремим етапом стало дослідження впливу просторового фактору на захищеність переговорів. Було змодельовано ситуацію віддалення диктофона від джерела завади на відстань від 1 до 2,5 м. Залежність індексу розбірливості STOI від відстані наведено на рис. 3.

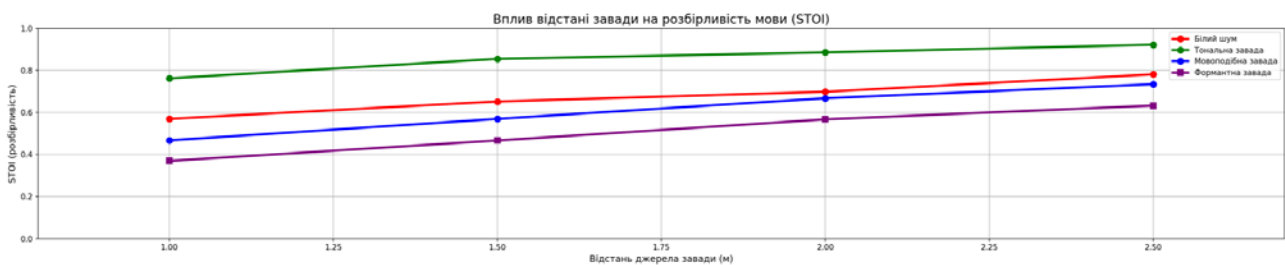


Рис. 3. Вплив відстані завади на розбірливість мови

Графік демонструє, що зі збільшенням відстані рівень звукового тиску завади падає, що призводить до зростання показника STOI (погіршення захисту). Проте, формантна завада демонструє найкращу стійкість: навіть на відстані 2,5 м показник STOI утримується на рівні, нижчому за поріг впевненого сприйняття мови. Традиційні ж методи (білий шум, тональна завада) на цій відстані втрачають свою ефективність (криві стрімко йдуть вгору), що створює загрозу витоку інформації.

Висновки

1. Проведено аналіз методів активного акустичного захисту мовної інформації. Встановлено, що підвищення ефективності систем захисту можливе шляхом переходу від енергетичних методів (білий шум) до структурних, що враховують спектральні особливості мови.
2. Розроблено програмну модель генерації завад та проведено їх двокритеріальне оцінювання. Використання метрики PESQ дозволило оцінити деградацію якості звучання, а метрика STOI – втрату змістовної складової мови.
3. Експериментально доведено безумовну перевагу запропонованої формантної завади. Вона забезпечує найнижчі показники якості ($PESQ = 1,13$) та мінімальну розбірливість ($STOI = 0,42$) порівняно з іншими дослідженими сигналами.
4. Дослідження залежності від відстані показало, що формантна завада дозволяє збільшити радіус захищеної зони. На відстані до 2,5 м вона зберігає низький рівень розбірливості мови, тоді як ефективність білого шуму та тональних завад на такій дистанції стає недостатньою.

Список літератури:

1. Засоби та системи технічного захисту інформації: Навчальний посібник для студентів ЗВО / І. С. Антіпов, А. М. Олейніков, Ю. В. Ликов, В. Д. Кукуш, І. О. Милютченко. 2-е вид., перероб. і доп. Харків : ХНУРЕ, 2024. 266 с.
2. Гоков О.М. Фізичні основи технічних засобів розвідки : навч. посіб. Харків : ХНЕУ ім. С. Кузнеця, 2022. 255 с.
3. Fielder G. D. Acoustical and Vibrational Methods of Information Protection // Journal of the Audio Engineering Society. 1998. Vol. 46, No. 7/8. P. 652–665.
4. Щербатий В. Д. Методи та засоби активного захисту від несанкціонованого перехоплення мовної інформації / В. Д. Щербатий, М. В. Коваль // Вісник Нац. техн. ун-ту України "КПІ". Сер. Приладобудування. 2019. Вип. 57(1). С. 105–111.

5. Lykov Y., Oliinyk A., Zavolodko G. Method of increasing the noise immunity of acoustic suppression systems for speech information // *Radioelectronics & Informatics*. 2023. No. 3. P. 43–48.

6. ITU-T Recommendation P.862: Perceptual evaluation of speech quality (PESQ): An objective method for end-to-end speech quality assessment of narrow-band telephone networks and speech codecs // International Telecommunication Union, 2001.

Надійшла до редколегії 11.01.2026

Прийнята до друку після рецензування 23.04.2026

Публікація (оприлюднення) 30.04.2026

Відомості про авторів:

Олейніков Анатолій Миколайович – канд. техн. наук, професор, Харківський національний університет радіоелектроніки, професор кафедри комп'ютерної радіоінженерії та систем технічного захисту інформації, Україна; e-mail: anatoly.oleynikov@nure.ua; ORCID: <https://orcid.org/0000-0002-4458-8833>

Шкопотко Павло Михайлович – Харківський національний університет радіоелектроніки, магістр кафедри комп'ютерної радіоінженерії та систем технічного захисту інформації, Україна; e-mail: pavlo.shkopotko@nure.ua; ORCID: <https://orcid.org/0009-0006-5080-1772>