

[DOI 10.28925/2663-4023.2025.29.937](https://doi.org/10.28925/2663-4023.2025.29.937)

УДК 004.934.2: 534.442.6

Нужний Сергій Миколайович

кандидат технічних наук, доцент,

завідувач кафедри комп'ютерних технологій та інформаційної безпеки

Національний університет кораблебудування імені адмірала Макарова, м. Миколаїв, Україна

ORCID: 0000-0002-7706-0453

s.nuzhniy@gmail.com

ОЦІНЮВАННЯ РІВНЯ ЗАХИЩЕНОСТІ СКЛАДНОЗАШУМЛЕНОЇ МОВНОЇ ІНФОРМАЦІЇ ЗА КРИТЕРІЄМ ЗАЛИШКОВОЇ РОЗБІРЛИВОСТІ МОВИ

Анотація. У статті обґрунтовано наукові засади оцінювання рівня захищеності складнозашумленої мовної інформації у технічних каналах витоку інформації за критерієм залишкової розбірливості. Проаналізовано обмеження традиційних підходів до визначення захищеності мовлення, зокрема показників SNR, SPC, AI та SII, які не враховують нелінійні спотворення, частотну селективність середовища та ефекти акустичної реверберації. Запропоновано використання інтегральної метрики RII, що характеризує здатність порушника відновити інформативний зміст сигналу після впливу завад. Розроблено узагальнену модель обробки мовного сигналу, яка поєднує вейвлет-фільтрацію, сегментацію, аналіз трифонної структури та нейронну оцінку залишкової розбірливості. Представлено класифікацію основних типів завад, зокрема білого, ревербераційного, комбінаційного та мовоподібного шумів, та наведено підходи до їх математичного формування. Показано, що врахування частотно-часових характеристик середовища та індивідуальних ознак диктора підвищує точність моделювання небезпечних сценаріїв перехоплення мовної інформації. Запропонований підхід забезпечує можливість об'єктивного ранжування каналів витоку та формування раціональних рішень щодо активного та пасивного акустичного захисту. Методика може бути використана для проектування комплексних систем технічного захисту інформації, експертного аналізу загроз та оцінювання ефективності засобів пригнічення мовлення у складних акустичних умовах. Особливу увагу приділено моделюванню акустичної взаємодії між джерелом, огорожувальними конструкціями та технічними засобами розвідки, що дає змогу кількісно оцінити ймовірність часткового або повного відновлення змісту повідомлення. Розроблена методика враховує вплив фазових спотворень, структурних максимумів спектра, характеру адаптивної фільтрації та умов тонального маскування. Запропоновано, з позиції зниження залишкової інформативності мовлення, процедуру порівняльного аналізу ефективності різних типів завад.

Ключові слова: комплексна система захисту інформації, захисті мовної інформації, оцінка рівня захищеності мовної інформації.

ВСТУП

Побудова КСЗІ об'єкту критичної інформаційної інфраструктури (ОКІІ), його задачі та структура визначається нормативно-правовими документами [1–3], нормативними документами (НД) з технічного захисту інформації (ТЗІ) [4–6], наказами Адміністрації Держспецзв'язку [7] та ін.

До переліку обов'язкових робіт, що повинні підтвердити стан та якість роботи КСЗІ, є проведення інструментального контролю [7]. По відношенню до вібраційного та акустичного технічних каналів витоку інформації (ТКВІ) критерієм захищеності є коефіцієнт, який визначається як відношення рівня небезпечного сигналу до рівня завади в точці можливого встановлення технічного засобу розвідки (ТЗР) за межами КЗ. В ЄС, США та інших англomовних краях, цей коефіцієнт отримав назву SNR (signal-to-noise



ratio). В Канаді використовують одну з його модифікацій – коефіцієнт приватності розмови SPC (speech privacy class) [8], [10].

Головним недоліком такого відходу є не врахування впливу спектрального складу мовного сигналу, «особливих» частотних смуг в тембрі голосу конкретного диктора, які можуть «з'явитись» і мати значну амплітуду. Їх виникнення пов'язане з «особливими» резонансами в голосовому тракті диктора і залежить від багатьох випадкових факторів. Також можлива поява таких частотних смуг, однак в набагато меншій мірі, і в наслідок впливу зовнішніх факторів. Так, наприклад, в результаті взаємодії таких фізичних явищ як реверберація, резонанс звуків на конструкційних елементах (в тому числі інтер'єр приміщення, меблі та інше), нелінійна модуляція тощо.

Для виправлення вказаного недоліку, США та ЄС ввели нормативний коефіцієнт, який дозволяє розрахувати рівень захищеності мовної інформації (РЗМІ) з врахуванням спектрального складу мовного сигналу та вагомості октав в передачі інформації – коефіцієнт (індекс) розбірливості мови SII (speech intelligibility index). Однак, як показано в [10] та [11], його використання при застосуванні процедур цифрової обробки (фільтрації, очистки), наприклад таких як вейвлет-перетворення, може дати негативний ефект. Так при досягненні досить високої якості очистки досліджуваного сигналу від завади на низьких та середніх частотах (до 4,5...5 кГц), на високих та надвисоких частотах спостерігається суттєве погіршення (зашумлення) досліджуваного сигналу [11], [13], [14]. Це призводить до визначення коефіцієнту SII, значення якого є некоректним і, відповідно, призводить до невірної оцінки рівня захищеності мовної інформації. [11]

Метою роботи є: розробка методу оцінки рівня захищеності складнозашумленої мовної інформації (РЗСЗМІ) за критерієм залишкової розбірливості мови диктора на основі нейронної мережі [16]

Для досягнення поставленої мети вирішувалися такі **наукові завдання** дослідження:

- Аналіз сучасного стану проблем оцінки РЗМІ методами цифрової обробки та очищення (фільтрації) зашумлених мовних сигналів.
- Теоретичне обґрунтування структури нейронної мережі для системи визначення та оцінки РЗСЗМІ за критерієм залишкової розбірливості мови диктора.

АНАЛІЗ ОСТАННІХ ДОСЛІДЖЕНЬ ТА ПУБЛІКАЦІЙ

Узагальнення наявних результатів дає підстави стверджувати, що жоден із існуючих підходів не забезпечує надійного прогнозування рівня захищеності складнозашумленої мовної інформації з позицій залишкової розбірливості мови диктора. Під поняттям «складнозашумленої мовної інформації» в роботі розуміється мовний сигнал, у якому відбувається одночасне накладання кількох різнорідних типів завад (або завад, створених системами постановки активних завад), що порушують як фізичну, так і лінгвістичну структуру мовлення та значно ускладнюють його реконструкцію традиційними методами. Нормативні документи [1–8] задають загальні вимоги та критерії, але не визначають процедур, які б коректно враховували спектрально-формантні особливості мовного сигналу та можливості сучасних алгоритмів відновлення мовлення. Наукові дослідження [10–19] фіксують обмеження класичних індексів розбірливості, показують суперечності між формальними метричними показниками й реальними результатами суб'єктивного сприйняття, однак не пропонують завершеної методики оцінювання рівня захищеності саме складнозашумленої мовної інформації. Це підтверджує актуальність поставленої в роботі задачі розроблення методу оцінки рівня



захищеності складнозашумленої мовної інформації за критерієм залишкової розбірливості мови диктора із використанням нейромережових моделей, адаптованих до спектрально-формантної структури українського мовлення [15].

Окрему групу публікацій складають роботи, присвячені побудові моделей оцінювання залишкової розбірливості мовлення в зашумлених приміщеннях із урахуванням просторових, спектральних та часових характеристик поля. Дослідження, де поєднано суб'єктивні тести розбірливості з об'єктивними показниками та моделюванням акустичних полів, показують, що існуючі індекси (SII, STI, AI) недостатньо чутливі до специфіки віброакустичних каналів і до ефектів нелінійної цифрової обробки [16], [17]. Автори наголошують на необхідності побудови спеціалізованих метричних моделей, які б враховували не лише спектральну енергію в октавних смугах, а й детальні параметри формантної структури, часових варіацій та індивідуальних «тембральних відбитків» мовця.

У дослідженнях останніх років активно розвиваються глибинні нейронні мережі для відновлення та підсилення мовлення в умовах екстремально низького SNR. Архітектури типу DNN, CNN та їхні гібриди демонструють здатність реконструювати мовний сигнал навіть за умов сильного маскування, проте ці алгоритми часто змінюють природну спектрально-формантну структуру мовлення [18], [20]. Для задач технічного захисту інформації це означає, що потенційний порушник може використати подібні методи для підвищення інформативності перехоплених сигналів, тоді як штатні методики оцінювання рівня захищеності, засновані на SNR або SII, не враховують такого сценарію. [21] [22]

Значна увага в сучасній літературі приділяється задачам реконструкції мовлення за вібраційними або оптичними відбитками, що моделюють роботу технічних засобів розвідки. Показано, що навіть низькоамплітудні вібрації конструкцій можуть містити достатньо інформації для відтворення мовного сигналу, якщо застосовувати спеціалізовані алгоритми обробки та машинного навчання [9], [17], [21], [22]. Ці результати свідчать про те, що формальна «низька чутливість» каналу за традиційними метричними показниками не гарантує відсутності можливості відновлення мовної інформації з позицій потенційного порушника.

У віброакустичних і акустичних каналах витоку інформації традиційно використовують SNR як базовий критерій оцінки захищеності, а в міжнародній практиці також застосовується коефіцієнт приватності мовлення SPC. Однак низка досліджень вказує, що при складній просторовій та частотній структурі завад енергетичні критерії не відображають реальної здатності засобів розвідки відновлювати мовний сигнал. Резонансні властивості будівельних конструкцій, реверберація приміщення та взаємодія хвильових процесів на межах поділу середовищ призводять до формування «особливих» частотних смуг із підвищеною інформативністю, які недостатньо описуються інтегральними індексами [9], [12].

Окремий напрям досліджень стосується методів підвищення якості та розбірливості сильно зашумленого мовлення. У монографічних роботах із теорії та практики підсилення мовних сигналів розглянуто широкий спектр алгоритмів – від класичних спектральних віднімальних фільтрів до вейвлет- та субсмугових методів, а також сучасних нейромережових підходів [11]. Експериментальні роботи показують, що, незважаючи на формальне підвищення відношення сигнал/шум, деякі алгоритми можуть знижувати розбірливість мовлення через спотворення формантної структури, міжсмужних кореляцій або надмірну згладженість спектра [11], [13]. Це особливо



критично для задач оцінки захищеності, де важливим є не лише «якість звучання», а саме збереження/втрата смислового змісту мовної інформації. [15]

Подальший розвиток отримали об'єктивні індекси якості та розбірливості, зокрема індекс розбірливості мови SII та пов'язані з ним стандартизовані методики вимірювань [8], [10], [13]. У роботах, присвячених порівнянню об'єктивних показників з суб'єктивною оцінкою якості мовлення, показано, що такі метрики добре корелюють з розбірливістю для стаціонарних завад та лінійних систем обробки мовного сигналу [14], [17]. Разом з тим, за умов нелінійної фільтрації, спектрального підсилення окремих смуг або істотної часово-частотної селекції сигналу спостерігаються суттєві відхилення між формальним значенням індексу SII і реальною здатністю людини розпізнавати мовлення.

У дослідженнях класичного етапу формування теорії розбірливості (articulation/intelligibility) було запропоновано низку метричних підходів до оцінки здатності каналу зв'язку зберігати інформативність мовного сигналу. Ранні роботи, присвячені артикуляційному індексу та факторам, що визначають розбірливість мовлення в телефонних і радіоканалах [9], [12], заклали основу для подальшого розвитку інтегральних критеріїв якості. Ці підходи орієнтовані переважно на лінійні системи передавання і не враховують нелінійних спотворень, а також складної спектрально-формантної структури реального мовлення в зашумлених приміщеннях.

Сучасний стан проблеми захисту мовної інформації характеризується поєднанням жорстких нормативних вимог та швидкого розвитку засобів цифрової обробки сигналів. Нормативне підґрунтя в Україні формують закони у сфері державної таємниці, захисту інформації в автоматизованих системах і персональних даних [1] – [3], а також комплекс документів з технічного захисту інформації (НД ТЗІ), що регламентують побудову, атестацію і контроль комплексних систем захисту інформації на об'єктах критичної інфраструктури [4–6]. Накази Адміністрації Держспецзв'язку встановлюють порядок проведення інструментального контролю, критерії приймання КСЗІ та підходи до документування результатів випробувань [7]. У сукупності ці документи формують вимоги до рівня захищеності мовної інформації, але не задають єдиної методики оцінювання залишкової розбірливості мови в умовах складних завад.

МЕТОДИ ОЦІНКИ РІВНЯ ЗАХИЩЕНОСТІ МОВНОЇ ІНФОРМАЦІЇ

Захист мовної/акустичної інформації є однією з пріоритетних задач забезпечення безпеки держави, її військової, політичної, економічної, наукової та ін. складових. Аналогічні задачі стоять і перед організаціями та фірмами. Проріхи в захисті призводять як до локальних так і до стратегічних збитків. Недостатній рівень захисту мовної інформації це не тільки неповага до себе та інтересів своєї фірми/організації, а до партнерів (співвиконавців, контрагентів та ін.).

Тому перед відповідними «структурами» (підрозділами) завжди стояла задача визначення рівня захищеності мовної інформації на конкретному об'єкті (кімнаті, наметі, споруді тощо.) та підтриманні його на необхідному рівні.

Історично так склалось, що люди в першу чергу покладались на можливості людського організму – можливості слухового апарату та здатність вирізняти окремі звуки (мову) з загального акустичного потоку. Це призвело до формування, так званих, суб'єктивних методів оцінки РЗМІ (рис.1).

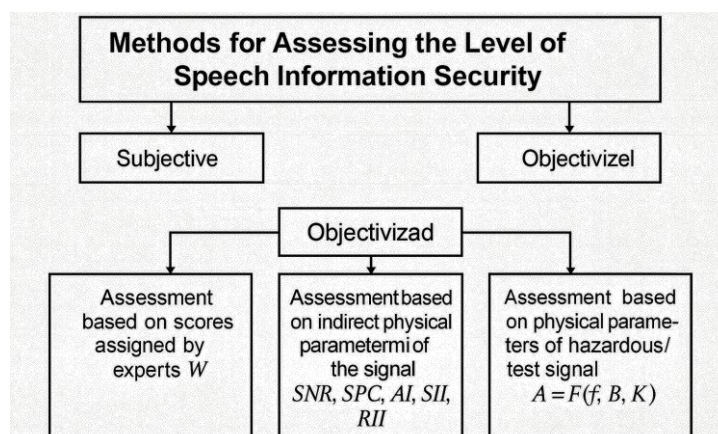


Рис. 1. Методика оцінки рівня захищеності мовної інформації

Один з найпоширеніших показників – коефіцієнт розбірливості мови W . Він заснований на проведенні експертизи, до якої залучаються диктор (або фонограма надиктованого ним тестового повідомлення) та група із 3-5 експертів-слухачів. Основною перевагою цього методу є розрахунок РЗМІ на основі обробки експертних оцінок і, відповідно, визначення реальної можливості розпізнавання порушником в перехопленому сигналі мовної інформації. Однак це й є основним недоліком методу. Він суттєво залежить від кваліфікації експертів-слухачів, їх настрою, стану здоров'я та ін. – тобто є суб'єктивним. З цього пункту випливає ще декілька особливостей – він потребує значних витрат робочого часу експертів і виведення з технологічного процесу досліджуваного приміщення та приміщень що його оточують з усіх сторін, а також зверху і знизу. У зв'язку з цим, цей метод, наразі, використовується майже тільки в лабораторних дослідженнях особливо відповідальних конструкційних елементів (приладів, систем, комплексів та ін.).

Розробка систем, які б забезпечували об'єктивне розпізнавання мовної інформації є надто складною задачею. Одні з перших науково обґрунтованих спроб відносяться до 17-о початку 19-о століть. Різними авторами були розроблені пристрої, що отримали назву «механічна голова» («talking head») або «розмовний пристрій» («talking device»), наприклад розмовний пристрій Euphonia). Деякі з них могли прослухати питання і відповісти на нього.

З розвитком ЕВМ, експертних систем (інформаційно-аналітичні системи, системи контролю доступу за голосовими сповіщеннями та багато інших), нейронних мереж і штучного інтелекту, питання розробки «електронного вуха» набули нового поштовху.

Втім ці методи мають ряд значних недоліків і основним з них є суттєва залежність від якості мовного сигналу і коефіцієнту SNR – вони не працюють з зашумленими сигналами (ЗС) ($SNR \geq 0$ дБ). Це призвело, на даний момент, до відсутності об'єктивних методів оцінки РЗМІ здатних працювати в умовах з від'ємними SNR ($SNR \leq -5$ дБ). [10]

Вирішенням проблеми, на даний час, стало використання об'єктизованих методів оцінки РЗМІ – тобто методів заснованих на інструментальному вимірюванні параметрів, які на пряму не пов'язані з РЗМІ та розрахунок, за певними аналітичними виразами, коефіцієнту, який і використовується для оцінки РЗМІ. До таких коефіцієнтів відносяться: (згадані вище) SNR, SPC, SII та індекс артикуляції AI (articulation index) і коефіцієнт залишкової розбірливості мови RII (residual intelligibility index). [10], [11]



ТИПИ СИГНАЛІВ ЗАВАДИ ТА МЕТОДИ ЇХ СИНТЕЗУ

Методи формування сигналу завади, в першу чергу, діляться в залежності від сигналу, що потребує захисту (рівень амплітуди сигналу завади до уваги не береться):

- сигналозалежні – створені/синтезовані на основі сигналу, що захищається;
- сигналонезалежні – створені без використання сигналу що захищається (автономні пристрої).

Всі типи завад, що використовуються в системах захисту інформації, поділяються на чотири групи:

1. Білий шум та спеціально сформовані на його основі сигнали (наприклад, його «кольорові» клони та «мовоподібний» сигнал);
2. Комбінаційні шум (використовується декілька незалежних джерел акустичного сигналу – отримали назву «мовний хор» або вечірка («party»));
3. Ревербераційний шум (використовується частотна або тимчасова реверберація);
4. Комбінований шум (використовується декілька з вказаних вище методів синтезу сигналу завади).

Доцільним є вказати, що 1) та 2) типи завад відносяться до сигналонезалежних методів формування сигналу завади, 3) тип завад є сигналозалежним, а 4) залежить від застосованих при його синтезі типів завад.

Ефективність використання того чи іншого типу завади визначається шляхом проведення інструментальних досліджень на об'єкті захисту або в лабораторії і розрахунком коефіцієнту РЗМІ.

Слід зауважити, що в Україні, США, ЄС та ін., в якості сигналу завади для захисту інформації, що містить конфіденційну інформацію та/або державну таємницю, дозволено використовувати тільки сигнали першої групи – білий шум та спеціально сформовані на його основі сигнали. Інші групи можна використовувати виключно для захисту комерційної інформації. [7]

Що ж до типів тест-сигналів, які використовуються при проведенні інструментальних досліджень, то на даний час використовуються:

1. Білий шум;
2. Моно- чи мультитоновий сигнал (наприклад, з розрахунку один чи декілька тонів на кожну $\frac{1}{3}$ -октавну смугу);
3. Мовоподібний сигнал (використовується нормована фонограма надиктованого диктором тестового повідомлення).

Тип тест-сигналу залежить від коефіцієнту який використовується для оцінки РЗМІ при проведенні досліджень. Так для SNR використовується білий шум, для AI, SII та RII – мовоподібний сигнал (фонограма мови диктора), а для та SPC можна використовувати як білий шум так і мовоподібний.

ПОПЕРЕДНЯ ОБРОБКА ЗАШУМЛЕНИХ СИГНАЛІВ

Додатково, на можливість розпізнавання тест-сигналу на фоні сигналу завади при оцінці РЗМІ, за умови що $SNR \leq -7 \dots -5$ дБ, суттєвий вплив має метод та процедура попередньої обробки сигналу (фільтрації). Так, при визначенні РЗМІ за коефіцієнтами SNR та SPC (за умови використання в якості тест-сигналу білого шуму) проведення попередньої обробки сигналу не допускається (див. рис.2). При застосуванні мовоподібного тест-сигналу (фонограми мови диктора) для визначенні РЗМІ за коефіцієнтами SPC, AI та SII проведення попередньої обробки, навіть для ЗС, в

офіційних методиках не розглядається. Більше того, передбачено проведення розрахунків коефіцієнтів SPC, AI та SII виключно для попередньо не оброблених сигналів. Для RII проведення попередньої обробки є доцільним, необхідність проведення визначається рівнем впливу сигналу завади на СЗ.

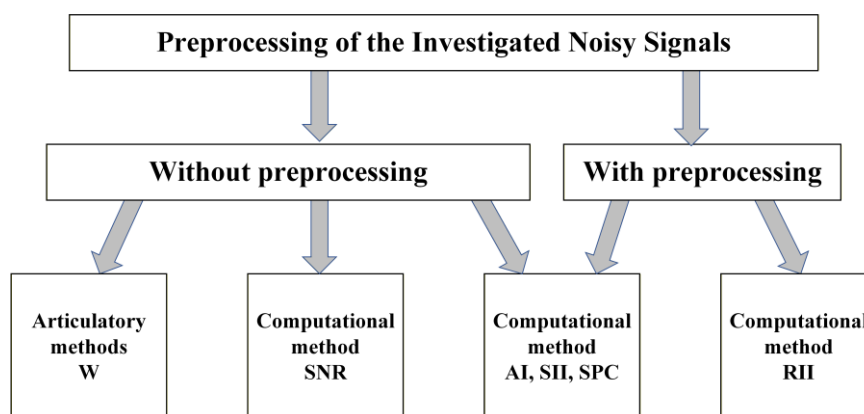


Рис. 2. Розподіл коефіцієнтів оцінки РЗАІ за попередньою обробкою досліджуваних СЗ

На рисунку подано класифікацію підходів до оцінювання рівня залишкової розбірливості мовної інформації. У схемі використано такі позначення:

МЕТОДИ ФІЛЬТРАЦІЇ ШУМОВОЇ ЗАВАДИ

Намагання покращити якість фонограми отриманого/перехопленого сигналу, за умови впливу на його шумової завади, завжди було одним із необхідних технологічних процесів. Однак, при використанні аналогових методів попередньої обробки, це можливо тільки в лабораторних умовах. При цьому перелік методів, які можуть дати позитивний ефект, є суттєво обмеженою (див. рис. 3). До них відносяться смугові та кореляційні методи.

При використанні цифрових методів обробки сигналів, перелік можливих до застосування методів є набагато різноманітним, як з точки зору рівня фільтрації так і вибірконості спектру. З врахуванням особливостей мовної інформації – її семантики, фонетики, умов формування в звуковому тракті людини та ін., виділяється декілька особливих методів попередньої обробки для ЗС, що пов'язані з частотно-часовим розподілом енергії в сигналі.

Найбільшого розповсюдження отримали вейвлет-перетворення, адаптивна фільтрація та кореляційний аналіз. Це зумовлюється тим, що, навіть при відносно не великих значеннях частотного спектру звукового сигналу (умовно, від 20 Гц до 20кГц) та його випадковості, зміну форми дотичної можна вважати квазістаціонарною (для часових періодів близьких до 0,1...0,2 типової протяжності фонем, тобто від 10 мс до 50 мс). Враховуючи (при рівномірному розподілу частот в спектрі сигналу завади типу білий шум) амплітуди окремих тонів сигналу завади в часовому і частотному вимірах є випадковими величинами, а отже і відповідні дотичні до такого сигналу на вказаних відтинках часу не можуть вважатись квазістаціонарними. Такий підхід на даний час став основою теоретичного апарату в обробці ЗС. [15]

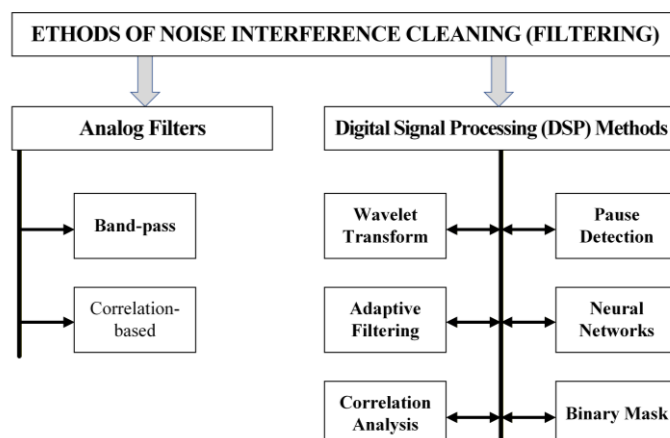


Рис. 3. Методи очистки (фільтрації) ЗС від шумової завади

Наступним елементом в підході до попередньої обробки ЗС стали методи виявлення пауз у мові диктора між словами. Це зумовлено тим, що в паузах відсутні квазістаціонарні сигнали, що відповідають елементам (енергії) спектру людської мови, а шуми, що пов'язані з навколишнім середовищем, мають більшу за часом і частотою стабільність (скрипи дверей, стільців та інших предметів, шерхіт вітру тощо.). В більшості методів попередньої обробки виявлені паузи використовують для сегментації мови на окремі блоки (ансамблі) і далі виконують попередню обробку до кожного з них окремо (див. рис. 4).

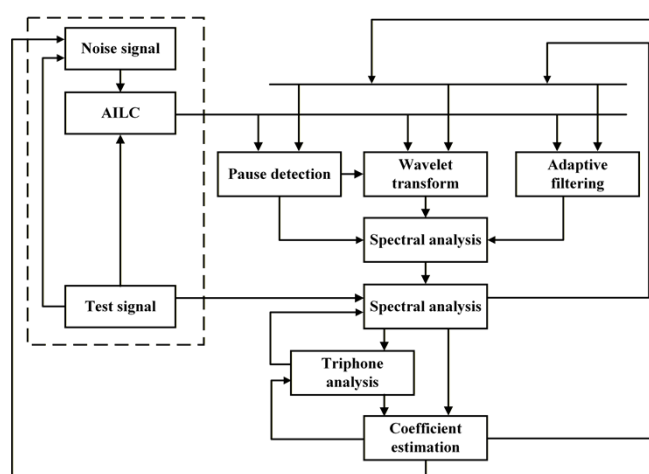


Рис. 4. Структурна модель системи оцінки рівня захищеності ЗС мовної інформації за критерієм залишкової розбірливості мови диктора

На рисунку наведено композиційну структуру системи оцінювання рівня захищеності мовної інформації. Використано такі скорочення:

- Wavelet-based pre-processing
- Pause detection
- Signal segmentation
- Triphone analysis
- Neural network model
- AILC (Acoustic Information Leakage Channel)
- попередня обробка. [11]
- виявлення пауз. [15]
- сегментація сигналу. [16]
- аналіз трифонів. [17]
- нейронна мережа. [16] [18]
- акустичний канал витоку інформації. [19]



- TMI (Technical Means of Intelligence) – технічні засоби розвідки. [20]
- Residual intelligibility estimation – оцінка залишкової розбірливості. [10]

Такий підхід дозволяє перейти до паралельної обробки сигналу, що суттєво підвищує якість та продуктивність роботи системи, однак потребує значно більших розрахункових потужностей (цей недолік виправляється, наприклад, використанням «майнінгових» станцій). Одночасно з цим, виявлення пауз додатково дає можливість збільшити якість основної обробки за рахунок застосування методу трифонів і перейти до оцінки рівня захищеності ЗС мовної інформації за критерієм залишкової розбірливості мови диктора.

При такому підході є доцільним поєднати попередню обробку з використанням, наприклад, вейвлет-перетворення та виявлення пауз, з аналізом трифонів, в якості основної обробки. Найбільш доцільним при цьому є використання нейронної мережі, що дозволить створити цілісну структуру.

ВИСНОВКИ

У результаті виконання роботи поставлену мету – розроблення науково-технічного методу оцінювання рівня захищеності складнозашумленої мовної інформації за критерієм залишкової розбірливості – досягнуто шляхом формування формалізованої моделі технічного каналу витоку акустичної інформації. Запропонований підхід забезпечує кількісне визначення залишкової інформативності мовного сигналу після впливу різнотипних завад і відповідає вимогам об'єктивності в задачах технічного захисту інформації.

У межах дослідження виконано такі науково-технічні завдання.

- 1) Проаналізовано існуючі критерії оцінювання інформативності (SNR, SPC, AI, SII, RII) та встановлено їх обмеження в умовах технічних каналів витоку акустичної інформації. Обґрунтовано використання показника RII як базового критерію.
- 2) Розроблено формалізовані моделі основних типів акустичних завад (білий, ревербераційний, комбінаційний, мовоподібний шум) та визначено їх спектрально-часові характеристики з позицій впливу на інформативність мовлення.
- 3) Сформовано комплекс процедур попередньої спектрально-часової обробки, що включає вейвлет-фільтрацію, сегментацію та трифонний аналіз. Забезпечено підвищення стійкості виділення інформативних компонентів.
- 4) Розроблено нейронну модель оцінювання залишкової розбірливості на основі спектрально-часових ознак, яка забезпечує прогнозування рівня інформативності мовлення після впливу акустичних завад.

Отримані результати мають практичну науково-технічну цінність і можуть бути використані для оцінювання ефективності акустичного маскування, ранжування небезпечних сценаріїв перехоплення, аудиту та сертифікаційного контролю систем технічного захисту інформації, а також для проектування комплексів активного й пасивного шумозахисту та створення інтелектуальних систем контролю витоку мовної інформації.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Закон України «Про державну таємницю» (1994). <https://zakon.rada.gov.ua/laws/show/3855-12>
2. Закон України «Про захист інформації в інформаційно-телекомунікаційних системах» (1994). <https://zakon.rada.gov.ua/laws/show/80/94-вр>



3. Закон України «Про захист персональних даних» (2010). <https://zakon.rada.gov.ua/laws/show/2297-17>
4. НД ТЗІ 3.7-003-2023. (2023). *Методика моделювання загроз та оцінки ризиків у системах технічного захисту інформації*. Київ: DSSZZI
5. НД ТЗІ 3.6-003-2016. (2016). *Порядок проведення атестації комплексної системи захисту інформації*. Київ: DSSZZI
6. НД ТЗІ 3.5-001-2017. (2017). *Загальні вимоги щодо забезпечення технічного захисту інформації*. Київ: DSSZZI
7. Адміністрація Держспецзв'язку України. Наказ № 023 (2015/2023). «Про затвердження Порядку проведення інструментального контролю технічних каналів витоку інформації». Київ: DSSZZI
8. Acoustical Society of America. (1997). ANSI/ASA S3.5-1997: *Methods for calculation of the Speech Intelligibility Index*. <https://acousticalsociety.org>.
9. French, N. R., & Steinberg, J. C. (1947). Factors governing the intelligibility of speech sounds. *Journal of the Acoustical Society of America*, 19(1), 90–119. <https://doi.org/10.1121/1.1916407>
10. Hu, Y., & Loizou, P. C. (2008). *Evaluation of objective quality measures for speech enhancement*. IEEE Transactions on Audio, Speech & Language Processing, 16(1), 229–238. <https://doi.org/10.1109/TASL.2007.911054>
11. Loizou, P. C. (2013). *Speech enhancement: Theory and practice* (2nd ed.). CRC Press. <https://doi.org/10.1201/b14529>
12. Dudley, H. (1940). *The carrier nature of speech*. Bell System Technical Journal, 19(4), 495-515. URL: <https://archive.org/details/bstj19-4-495>.
13. Zhao, Y., Ni, Z., Chen, M., & Wang, B. (2018). *Convolutional-recurrent neural networks for speech enhancement*. IEEE/ACM Transactions on Audio, Speech, and Language Processing, 26(12), 2152–2161. <https://doi.org/10.1109/TASLP.2018.2871433>
14. Snyder, D., Garcia-Romero, D., Sell, G., Povey, D., & Khudanpur, S. (2018). *X-vectors: Robust DNN embeddings for speaker recognition*. Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 5329–5333. <https://doi.org/10.1109/ICASSP.2018.8461375>
15. Ткаченко, М & Федоренко, Р & Кондратенко, Ю & Зотова, І. (2019). *Методи автоматичної ідентифікації диктора за голосом*. Збірник наукових праць Центру воєнно-стратегічних досліджень НУОУ імені Івана Черняхівського. 131-135. <https://doi.org/10.33099/2304-2745/2018-3-64/131-135>
16. Rabiner, L. R., & Schafer, R. W. (1978). *Digital Processing of Speech Signals – Online complement*. <https://people.csail.mit.edu/rmd/Projects/SpeechProcessing/>
17. O'Shaughnessy, D. (2000). *Speech communications: Human and machine* (2nd ed.). IEEE Press. <https://doi.org/10.1109/9780470545397>
18. Rabiner, L. R. (1989). *A tutorial on hidden Markov models and selected applications in speech recognition*. Proceedings of the IEEE, 77(2), 257–286. <https://doi.org/10.1109/5.18626>
19. Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT Press. <https://www.deeplearningbook.org/>
20. Davis, A., Rubinstein, M., Wadhwa, N., Durand, F., & Freeman, W. (2014). *The visual microphone: Passive recovery of sound from video*. ACM SIGGRAPH 2014. <https://doi.org/10.1145/2601097.2601119>
21. U.S. Department of Defense. (2014). *DoD Instruction 5240.05: Technical surveillance countermeasures (TSCM) (Incorporating Change 2, August 27, 2020)*. <https://www.esd.whs.mil/Portals/54/Documents/DD/issuances/dodi/524005p.pdf>.
22. U.S. Department of Defense. (2017). *Technical Surveillance Countermeasures (TSCM) handbook*. <https://www.dni.gov/files/NCSC/documents/products/TSCM-Handbook.pdf>
23. Taal, C. H., Hendriks, R. C., Heusdens, R., & Jensen, J. (2011). *A short-time objective intelligibility measure for time-frequency weighted noisy speech*. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), 4214–4217. <https://doi.org/10.1109/ICASSP.2011.5944212>
24. Ephrat, A., Halperin, T., Peleg, S., & Freeman, W. T. (2018). *Looking to listen at the cocktail party: A visual-audio model for speech separation*. ACM Transactions on Graphics (SIGGRAPH), 37(4), 1–12. <https://doi.org/10.1145/3197517.3201397>

**Serhii Nuzhnyi**

Candidate of Technical Sciences, Associate Professor,
Head of the Department of Computer Technologies and Information Security
Admiral Makarov National University of Shipbuilding, Mykolaiv, Ukraine
ORCID: 0000-0002-7706-0453
s.nuzhnyi@gmail.com

ASSESSMENT OF THE SECURITY LEVEL OF HEAVILY NOISE-DISTORTED SPEECH INFORMATION BASED ON THE RESIDUAL INTELLIGIBILITY CRITERION

Abstract. This paper substantiates the scientific foundations for assessing the security level of heavily noise-distorted speech information in technical channels of information leakage using the residual intelligibility criterion. The limitations of traditional approaches to speech-security assessment are analyzed, including SNR, SPC, AI, and SII metrics, which do not account for nonlinear distortions, the frequency-selective properties of the environment, or acoustic reverberation effects. An integral metric, RII, is proposed to characterize an adversary's ability to reconstruct the informative content of a signal after interference exposure. A generalized speech-processing model is developed, combining wavelet filtering, segmentation, triphone-structure analysis, and neural residual-intelligibility estimation. A classification of major noise types is presented, including white, reverberant, combinational, and speech-like noise, along with approaches to their mathematical generation. It is demonstrated that incorporating the time-frequency characteristics of the environment and the speaker's individual features increases the accuracy of modeling hazardous scenarios of speech-information interception. The proposed approach enables objective ranking of leakage channels and supports rational decision-making regarding active and passive acoustic protection. The methodology can be applied to the design of comprehensive technical information-protection systems, threat expert analysis, and evaluation of the effectiveness of speech-suppression measures under complex acoustic conditions. Special attention is given to modeling the acoustic interaction between the source, enclosing structures, and technical reconnaissance devices, enabling quantitative assessment of the probability of partial or complete reconstruction of message content. The developed methodology accounts for phase distortions, structural spectral maxima, adaptive-filtering characteristics, and conditions of tonal masking. From the standpoint of reducing residual speech informativeness, a comparative analysis procedure is proposed for evaluating the effectiveness of different types of interference.

Keywords: Integrated information-protection system, speech-information protection, assessment of speech-information security level.

REFERENCES (TRANSLATED AND TRANSLITERATED)

1. Verkhovna Rada Ukrainy. (1994). *Zakon Ukrainy "Pro derzhavnu taïemnytsiu"*. <https://zakon.rada.gov.ua/laws/show/3855-12>
2. Verkhovna Rada Ukrainy. (1994). *Zakon Ukrainy "Pro zakhyst informatsii v informatsiino-telekomunikatsiïnykh systemakh"*. <https://zakon.rada.gov.ua/laws/show/80/94-vr>
3. Verkhovna Rada Ukrainy. (2010). *Zakon Ukrainy "Pro zakhyst personalnykh danykh"*. <https://zakon.rada.gov.ua/laws/show/2297-17>
4. NDTZI 3.7-003-2023. (2023). *Metodyka modeliuvannia zahroz ta otsinky ryzykiv u systemakh tekhnichnoho zakhystu informatsii*. Kyiv: DSSZZI.
5. NDTZI 3.6-003-2016. (2016). *Poriadok provedennia atestatsii kompleksnoi systemy zakhystu informatsii*. Kyiv: DSSZZI.
6. NDTZI 3.5-001-2017. (2017). *Zahalni vymohy shchodo zabezpechennia tekhnichnoho zakhystu informatsii*. Kyiv: DSSZZI.
7. Administratsiia Derzhspetsviazku Ukrainy. (2015/2023). *Nakaz № 023 "Pro zatverdzhennia Poriadku provedennia instrumentarnoho kontroliu tekhnichnykh kanaliv vytku informatsii"*. Kyiv: DSSZZI.
8. Acoustical Society of America. (1997). *ANSI/ASA S3.5-1997: Methods for calculation of the Speech Intelligibility Index*. <https://acousticalsociety.org>



9. French, N. R., & Steinberg, J. C. (1947). *Factors governing the intelligibility of speech sounds*. *Journal of the Acoustical Society of America*, 19(1), 90–119. <https://doi.org/10.1121/1.1916407>
10. Hu, Y., & Loizou, P. C. (2008). *Evaluation of objective quality measures for speech enhancement*. *IEEE Transactions on Audio, Speech & Language Processing*, 16(1), 229–238. <https://doi.org/10.1109/TASL.2007.911054>
11. Loizou, P. C. (2013). *Speech enhancement: Theory and practice* (2nd ed.). CRC Press. <https://doi.org/10.1201/b14529>
12. Dudley, H. (1940). *The carrier nature of speech*. *Bell System Technical Journal*, 19(4), 495–515. <https://archive.org/details/bstj19-4-495>
13. Zhao, Y., Ni, Z., Chen, M., & Wang, B. (2018). *Convolutional-recurrent neural networks for speech enhancement*. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 26(12), 2152–2161. <https://doi.org/10.1109/TASLP.2018.2871433>
14. Snyder, D., Garcia-Romero, D., Sell, G., Povey, D., & Khudanpur, S. (2018). *X-vectors: Robust DNN embeddings for speaker recognition*. In *ICASSP 2018: IEEE International Conference on Acoustics, Speech and Signal Processing* (pp. 5329–5333). <https://doi.org/10.1109/ICASSP.2018.8461375>
15. Tkachenko, M., Fedorenko, R., Kondratenko, Yu., & Zotova, I. (2019). *Metody avtomatichnoi identyfikatsii dyktora za holosom*. *Zbirnyk naukovykh prats Tsentru voienno-stratehichnykh doslidzhen NUOU imeni Ivana Cherniakhovskoho*, 131–135. <https://doi.org/10.33099/2304-2745/2018-3-64/131-135>
16. Rabiner, L. R., & Schafer, R. W. (1978). *Digital processing of speech signals – Online complement*. <https://people.csail.mit.edu/rmd/Projects/SpeechProcessing/>
17. O'Shaughnessy, D. (2000). *Speech communications: Human and machine* (2nd ed.). IEEE Press. <https://doi.org/10.1109/9780470545397>
18. Rabiner, L. R. (1989). *A tutorial on hidden Markov models and selected applications in speech recognition*. *Proceedings of the IEEE*, 77(2), 257–286. <https://doi.org/10.1109/5.18626>
19. Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT Press. <https://www.deeplearningbook.org>
20. Davis, A., Rubinstein, M., Wadhwa, N., Durand, F., & Freeman, W. (2014). *The visual microphone: Passive recovery of sound from video*. *ACM SIGGRAPH 2014*. <https://doi.org/10.1145/2601097.2601119>
21. U.S. Department of Defense. (2014). *DoD Instruction 5240.05: Technical surveillance countermeasures (TSCM) (Incorporating Change 2, August 27, 2020)*. <https://www.esd.whs.mil/Portals/54/Documents/DD/issuances/dodi/524005p.pdf>
22. U.S. Department of Defense. (2017). *Technical Surveillance Countermeasures (TSCM) handbook*. <https://www.dni.gov/files/NCSC/documents/products/TSCM-Handbook.pdf>
23. Taal, C. H., Hendriks, R. C., Heusdens, R., & Jensen, J. (2011). *A short-time objective intelligibility measure for time-frequency weighted noisy speech*. *ICASSP 2011: IEEE International Conference on Acoustics, Speech, and Signal Processing*, 4214–4217. <https://doi.org/10.1109/ICASSP.2011.5944212>
24. Ephrat, A., Halperin, T., Peleg, S., & Freeman, W. T. (2018). *Looking to listen at the cocktail party: A visual-audio model for speech separation*. *ACM Transactions on Graphics*, 37(4), 1–12. <https://doi.org/10.1145/3197517.3201397>

