

<https://doi.org/10.31891/2307-5732-2026-363-51>

УДК 004.021

БОНДАРЕНКО МАКСИМ

Харківський національний університет радіоелектроніки

<https://orcid.org/0000-0002-2500-7626>

e-mail: maksym.bondarenko@nure.ua

ІВАЩЕНКО ГЕОРГІЙ

Харківський національний університет радіоелектроніки

<https://orcid.org/0000-0003-1027-5262>

e-mail: heorhii.ivashchenko@nure.ua

МЕТОДИ ОНОВЛЕННЯ ПРОФІЛЯ КОРИСТУВАЧА В СИСТЕМАХ БАГАТОВЕКТОРНОЇ ГОЛОСОВОЇ ІДЕНТИФІКАЦІЇ

У роботі досліджено підходи щодо підтримки актуальності векторних профілів користувачів у системах голосової ідентифікації. Запропоновано рішення, яке передбачає оновлення профілю та коригування вагових коефіцієнтів векторних представлень. Проведено порівняльний аналіз ефективності методів інкрементального додавання векторів, часткового й повного заміщення ознак, а також багатовекторних підходів з фіксованими та змінними значеннями вагових коефіцієнтів. Обчислювальні експерименти підтвердили ефективність запропонованих схем оновлення профілю та дозволили оцінити їхній вплив на точність і робастність голосової ідентифікації.

Ключові слова: системи голосової ідентифікації, фізіологічні зміни голосу, голосовий профіль користувача, вектор ознак, багатовекторний банк, адаптивне оновлення.

BONDARENKO MAKSYM, IVASHCHENKO HEORHII

Kharkiv National University of Radioelectronics

UPDATING USER PROFILES METHODS IN MULTIVECTOR VOICE IDENTIFICATION SYSTEMS

The article presents the results of a comparative analysis of approaches to updating and adapting vector user profiles in speech identification systems, along with a new method. The article focuses on maintaining the stability and accuracy of recognition, taking into account voice changes between sessions, due to physiological factors, variations in articulation and acoustic recording conditions. Methods such as incremental vector addition, partial and complete feature replacement, and multi-vector approaches with fixed or adaptive weighting are analysed.

The proposed method assigns a weight coefficients based on the similarity between new voice samples and existing profile vectors, thereby balancing adaptability with robustness to noise and atypical voice samples. Experiments were conducted to evaluate different mechanisms for updating user profiles and to assess their impact on the accuracy, stability, and robustness of identification.

Experimental results show that multi-vector approaches with distinct weighting schemes, including the proposed method, more effectively account for vocal changes between sessions and variable acoustic conditions. This process reduces the false-rejection rate compared to fixed-profile and single-vector updating approaches. The choice of update mechanism significantly influences the dynamics of profile adaptation, system stability in the presence of noise and physiological voice changes, and the balance between recognition accuracy and computational costs.

The results of this development have practical significance for the design and configuration of speech recognition systems that rely on user profiles for long-term performance. This paper offers recommendations for selecting an update approach based on operational conditions and the desired balance between recognition accuracy, robustness, and computational efficiency.

Keywords: voice identification systems, physiological voice changes, user voice profile, feature vector, multi-vector bank, adaptive updating.

Стаття надійшла до редакції / Received 12.02.2026

Прийнята до друку / Accepted 28.02.2026

Опубліковано / Published 26.03.2026



This is an Open Access article distributed under the terms of the [Creative Commons CC-BY 4.0](https://creativecommons.org/licenses/by/4.0/)

© Бондаренко Максим, Іващенко Георгій

Постановка проблеми у загальному вигляді та її зв'язок із важливими науковими чи практичними завданнями

Акустичні параметри голосу визначаються анатомо-фізіологічними особливостями мовця і залежать від його фізіологічного стану, що включає рівень втоми, психоемоційний стан, наявність захворювань дихальних шляхів або голосового апарату, а також вікові зміни [1]. Ці фактори можуть призводити до варіацій основного тону, тембру голосу та спектральних характеристик мовлення. Внаслідок цього акустичні параметри голосу одного й того ж користувача, зафіксовані у різні часові проміжки, можуть значно відрізнитися [2].

Додаткову нестабільність голосового сигналу одного користувача визначають умови запису та технічні характеристики системи фіксації. До таких чинників належать акустичні властивості приміщення (наявність резонансу, відбиття звуку, фонові шуми), параметри мікрофона та його розташування, а також характеристики каналу передавання сигналу [3]. Крім того, вплив мають методи попередньої обробки, зокрема фільтрація, кодування та нормалізація амплітуди [4-5]. Сукупність усіх цих факторів формує значну змінність голосових характеристик одного користувача.

Безпосереднє порівняння голосових сигналів, отриманих у різні моменти часу, є складним завданням, оскільки їхні акустичні властивості можуть суттєво відрізнитися (рис. 1). У системах голосової ідентифікації ці сигнали перетворюються у векторні представлення, які формалізують основні ознаки голосу та формують профіль користувача для подальшого зіставлення. Висока варіативність таких представлень ускладнює точне порівняння нових зразків із існуючими профілями та підвищує складність побудови надійних систем ідентифікації.

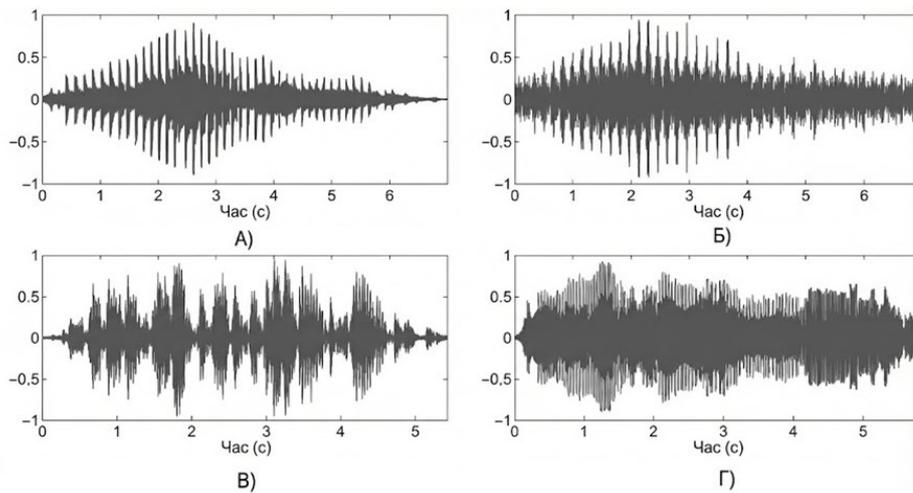


Рис. 1. Часове представлення (осцилограми) голосових сигналів у різних умовах: а) – голос у нормальних умовах (без шуму); б) – голос в умовах вуличного шуму; в) емоційно змінений голос; г) оброблений голос (компресія сигналу)

Для врахування нестабільності голосових характеристик використовуються багатовекторні голосові профілі користувачів, у яких для одного користувача зберігається декілька векторних представлень його голосу (рис. 2) [6]. Ці вектори можуть оновлюватися частково або повністю, що дозволяє профілю адаптуватися до змін голосових ознак у часі та за різних умов запису. При цьому різним векторним представленням можуть надаватися різні коефіцієнти значущості, що дозволяє виділяти ознаки, найбільш характерні для користувача.

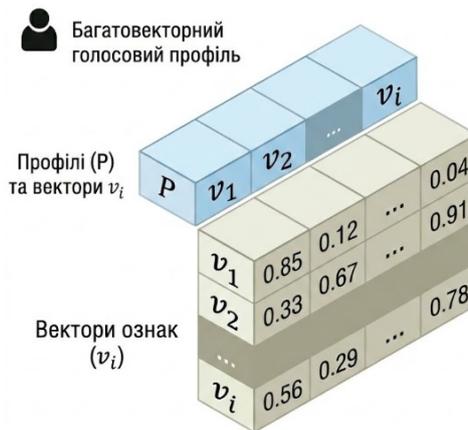


Рис. 2. Структура багатовекторного голосового профілю, сформованого з послідовності векторів ознак

Багатовекторний підхід забезпечує більш точне зіставлення нових голосових зразків із існуючими профілями та підвищує надійність ідентифікації навіть за умов значної варіативності голосу, спричиненої фізіологічними, мовленнєвими або зовнішніми факторами. Багатовекторний підхід дозволяє не лише враховувати різні вектори, а й формувати стратегії їх оновлення та зважування, що дає змогу адаптувати системи голосової ідентифікації до практичних вимог і задач, компенсувати несталість голосових записів та знижувати ризики підробки голосу в умовах широкого поширення технологій синтезу та зміни голосу [7].

Аналіз досліджень та публікацій

Актуальні напрямки розвитку систем обробки голосових сигналів відображені у сучасних дослідженнях [8-11], присвячених багатовекторним підходам до оновлення профілю користувача, у яких профіль розглядається як динамічна сукупність векторних представлень голосу одного користувача, що відображають різні стани та варіації його акустичних характеристик, сформовані під впливом фізіологічних змін, мовленнєвої манери та умов реєстрації сигналу. У таких підходах оновлення профілю описується як безперервний процес його актуалізації у часі, що може здійснюватися шляхом зсуву наявного набору векторів із поступовим введенням нових представлень, розширенням множини векторів для охоплення ширшого спектра голосових варіацій або повною заміною застарілих зразків актуальними. При цьому пріоритет надається способам урахування внеску окремих векторних представлень у формування узагальненого профілю користувача. Такі підходи можуть базуватися на рівнозначному врахуванні всіх векторів або на диференційованих схемах, де більш інформативним або актуальним записам надають більший вплив. Такий підхід до оновлення профілю дозволяє розглядати його як адаптивну структуру, здатну адекватно відображати динаміку голосових характеристик і забезпечувати стійку роботу системи за умов їх суттєвої варіативності.

У дослідженні [8] запропоновано підхід до профілювання мовця на основі одновимірної згорткової нейронної мережі (1D CNN), який дозволяє одночасно оцінювати вік, зріст та гендер з коротких аудіо-фрагментів. Особливістю підходу є те, що фільтри першого шару мережі є тренуваними, тобто під час навчання їхні ваги коригуються на основі вхідних даних, щоб виділяти найбільш інформативні ознаки мовця. Це дозволяє створити персоналізовані набори фільтрів, підібрані для виділення характеристик голосу користувача. Підхід підтримує інкрементальне оновлення профілю мовця, що робить його перспективним для систем, де регулярно додаються нові голосові фрагменти або характерна нестабільність умов мовлення.

Дослідження [9] описує систему UzAssistant, призначену для автоматизованого обслуговування користувачів у контакт-центрах. Система формує персоналізовані голосові профілі на основі коротких голосових зразків, використовуючи MFCC-ознаки та GMM-моделі. Особливістю підходу є інкрементальне оновлення профілю мовця: нові голосові фрагменти додаються до існуючого профілю, що дозволяє системі адаптуватися до змін голосу користувача або умов запису та підвищує точність розпізнавання. Завдяки інкрементальному оновленню система здатна підтримувати актуальність голосових профілів користувачів у довгостроковій перспективі, що підвищує надійність і точність розпізнавання у роботі з великими потоками голосових звернень.

Дослідження [10] розглядає донавчання та порівняльний аналіз моделей синтезу мовлення Tacotron 2, Deep Voice 3 та FastSpeech 2 в умовах обмежених обчислювальних ресурсів і малих обсягів навчальних даних. Автори демонструють, що використання попередньо навчених моделей і тонкого налаштування параметрів дає змогу ефективно адаптувати систему до нових мовленнєвих даних без потреби у великому корпусі записів і багатопроцесорному середовищі. Використання навчання з перенесенням і часткового переналаштування ваг дозволяє уточнювати векторні представлення мовлення користувача та розширювати профіль ознак без повного перенавчання моделі. Така стратегія зменшує обчислювальні витрати, пришвидшує збіжність і зберігає стабільність базових характеристик голосового представлення.

У дослідженні [11] розглядаються методи динамічної адаптації профілів у потокових системах. Запропонований метод базується на безперервному оновленні групи векторних представлень голосу, що належать одному користувачу або мають схожі характеристики шляхом додавання нових вхідних векторів (x-vectors/d-vectors) до існуючих усереднених векторів (центроїдів), які представляють кожну групу. Алгоритм використовує порогову метрику відстані для прийняття рішення про приєднання нового сегмента до існуючого профілю або створення нового. Це дозволяє системі в режимі реального часу автоматично компенсувати зміни в каналі передачі даних або положенні мікрофона. Експериментальні дані підтверджують ефективність методу для зниження часової затримки при збереженні конкурентної точності.

Результати аналізу сучасних досліджень свідчать, що з розвитком систем голосової ідентифікації проблема оновлення профілю ознак користувача залишається актуальною. Існуючі підходи демонструють різну ефективність залежно від умов експлуатації, характеру змін голосових характеристик та обраної стратегії інтеграції нових даних (схеми оновлення) у профіль користувача. Сучасні системи голосової ідентифікації функціонують в умовах нових викликів, пов'язаними з поширенням технологій синтезу та підробки голосу, що зумовлює розробку нових методів оновлення профілю користувача, здатних ефективно адаптуватися до змін голосу та забезпечувати високу точність і надійність ідентифікації у сучасних практичних застосуваннях.

Формулювання цілей статті

Метою роботи є: розробка та порівняння методів оновлення профілю користувача у системах голосової ідентифікації на основі багатовекторного банку ознак, із метою забезпечення актуального стану профілю користувача в умовах варіативності акустичного середовища та характеристик мовлення.

Виклад основного матеріалу

Сучасні системи голосової ідентифікації передбачають попередню обробку голосового сигналу (видалення шумів, нормалізацію амплітуди, компенсацію акустичних викривлень), сегментацію на короткі фрейми та виділення векторних представлень, що формалізують спектральні, частотні та часові характеристики голосу. Вектори можуть містити інформацію про темп, інтонаційні патерни, артикуляційні особливості та фонетичні маркери, забезпечуючи повне представлення голосу для порівняння з еталонними профілями.

Використання одного статичного вектора на користувача забезпечує швидке зіставлення, але не забезпечує точності при міжсесійних змінах голосу, варіаціях акустичного середовища або наявності шуму. Багатовекторні підходи дозволяють відображати різні стани голосу та умови запису, усуваючи обмеження класичних методів: втрату актуальності профілю, низьку робастність і обмеженість у відображенні індивідуальних голосових ознак користувача.

При багатовекторному підході оновлення профілю здійснюється шляхом часткової або повної заміни векторів у наявній множині профілю на вектори нового сигналу. Набір векторів профілю користувача представлений множиною векторів $P = \{v_1, v_2, \dots, v_N\}$.

При надходженні нового векторного представлення x_{new} профіль може оновлюватися за схемою повного заміщення старого вектора:

$$v_j \leftarrow x_{new}, \text{ якщо } \arg \min_i d(x_{new}, v_i), \quad (1)$$

де $d(\cdot, \cdot)$ – обрана метрика схожості (наприклад, косинусна відстань або евклідова). Такий підхід

ефективний у стабільних умовах запису та при помірних змінах голосу, але не дозволяє диференціювати внесок різних векторів у підсумковий профіль.

Оновлення профілю може включати зміни окремих коефіцієнтів значущості a_i для кожного векторного представлення, що дозволяє збільшити вплив найбільш релевантних ознаки. Узагальнений профіль формується як зважене поєднання векторів:

$$v_{agg} = \sum_{i=1}^N \alpha_i v_i, \quad \sum_{i=1}^N \alpha_i \neq 1. \quad (2)$$

Коефіцієнти можуть бути сталими, тобто $a_i = const$ або мати адаптивні значення, які можна обчислювати на основі схожості нового сигналу до кожного вектора профілю:

$$\alpha_i = \frac{\exp(-\lambda d(x_{new}, v_i))}{\sum_{j=1}^N \exp(-\lambda d(x_{new}, v_j))} \quad (3)$$

де λ – параметр, що регулює розподіл ваги значущості, а $d(\cdot, \cdot)$ – метрика схожості. У цьому випадку більш нові або більш релевантні вектори отримують більший внесок у профіль, а менш інформативні чи застарілі вектори впливають менше.

Інкrementальне оновлення (розширення) профілю передбачає, що нові вектори не заміщують існуючі, а додаються до множини P для охоплення ширшого спектра станів голосу. При цьому існуючі вектори можуть поступово зменшувати свій коефіцієнт значущості a_i за функцією часу або частоти використання:

$$\alpha_i(t+1) = \beta \alpha_i(t), \quad 0 < \beta < 1. \quad (4)$$

Такий підхід дозволяє профілю поступово адаптуватися до змін голосу користувача, накопичуючи інформацію про попередні стани для забезпечення стабільної ідентифікації. Він особливо ефективний у сценаріях потокових систем або при обробці численних голосових фрагментів, записаних за різних умов.

Інший варіант оновлення базується на зсуві та комбінованій адаптації, де існуючі вектори заміщуються новими частково, а коефіцієнти перераховуються залежно від відповідності нового сигналу та накопичених ознак:

$$v_i \leftarrow (1 - \alpha_i) v_i + \alpha_i x_{new}, \quad (5)$$

де a_i визначає швидкість адаптації і може бути різною для кожного вектора. Високі значення a_i встановлюються більш релевантним або свіжим зразкам, що дозволяє швидко оновлювати профіль у випадках різких змін голосу, наприклад, через захворювання або шумне середовище.

Запропоновано метод оновлення багатовекторного профілю користувача, який передбачає інтеграцію нового вектора v_{new} з існуючими в профілі. Кожен наявний вектор v_i враховується при формуванні оновленого профілю. Середній стан профілю, який використовується для обчислення ваг векторів профіля з розмірністю N , визначається як:

$$C = \frac{1}{N} \sum_{i=1}^N v_i. \quad (6)$$

Абсолютна вага кожного вектора обчислюється на основі його відстані до середнього стану профілю, що дозволяє враховувати його внесок при оновленні профілю:

$$w_i = \frac{1}{d(v_i, C)}, \quad w_i \geq 0. \quad (7)$$

У такій схемі вектори, які найбільш узгоджені з середнім станом профілю і відображають типові характеристики профілю, отримують більший внесок, тоді як вектори, що значно відхиляються, впливають на оновлений профіль менше. Оновлений профіль формується як:

$$P = \gamma v_{new} + \sum_{i=1}^N w_i v_i, \quad (8)$$

де $\gamma \geq 0$ визначає абсолютний вплив нового сигналу, без необхідності нормалізації сумарної ваги. Такий підхід дозволяє гнучко регулювати внесок нового та старих векторів, зберігати ознаки голосу користувача та одночасно адаптувати профіль до змін акустичного середовища чи фізіологічних варіацій голосу.

Для оцінки ефективності багатовекторних підходів до оновлення профілю користувача у системах голосової ідентифікації необхідно провести комплексне експериментальне дослідження. Слід реалізувати та порівняти різні методи оновлення профілю: просте інкрементальне додавання векторів, часткове або повне заміщення наявних ознак, а також підходи з диференційованими адаптивними коефіцієнтами. Експерименти повинні відтворювати типові умови системи, включаючи міжсесійну варіативність голосу, зміну акустичного середовища, різні рівні шуму та варіації емоційного стану. Особливу увагу слід приділити як поступовим змінам голосу, так і різким зсувам через зміну обладнання або умов запису. Потрібно оцінити вплив кожного підходу на точність ідентифікації, стабільність профілю та робастність системи, а також з'ясувати, як механізм оновлення і спосіб обчислення коефіцієнтів впливають на баланс між адаптивністю профілю та збереженням його основних ознак.

Результати експериментальних досліджень

Експериментальне дослідження ефективності підходів до оновлення профілю ознак мовця реалізовано у відтворюваному обчислювальному середовищі на базі Jupyter Notebook з використанням Python. Обчислювальний конвеєр розгорнуто з підтримкою CPU та GPU, де етапи формування спікерських векторів ознак і операцій зіставлення виконуються на графічному прискорювачі. Програмна реалізація побудована із

використанням бібліотек NumPy та SciPy для числової обробки й операцій у векторних просторах, Librosa і TorchAudio для аналізу та підготовки аудіосигналів, а також PyTorch, SpeechBrain і ruannotate.audio для отримання векторних представлень мовця на основі попередньо навчених нейромережових моделей. Розрахунок метрик та статистична обробка результатів виконані засобами scikit-learn [12]. Параметри обробки та конфігурації моделей зафіксовані, що забезпечує однакові умови для всіх варіантів оновлення профілю.

Конверсер обробки мовлення реалізовано як послідовність стандартизованих етапів: уніфікація частоти дискретизації, амплітудна нормалізація, спектральне перетворення та екстракція ознак. Для формування векторних представлень використано MFCC-ознаки та нейромережові вектори ознак мовця типу x-vector і d-vector, отримані з моделей SpeechBrain [13]. Для аналізу сигнал поділяється на короткі інтервали, після чого формуються вектори, що відображають характеристики окремих фрагментів або сесій.

В якості тестових даних використано мовленнєві корпуси VoxCeleb1 та VoxCeleb2 [14], що містять багатосесійні записи великої кількості мовців у різних акустичних умовах і каналах запису. Для кожного мовця сформовано початковий профіль на основі обмеженої стартової вибірки записів. Наступні записи подаються послідовно та використовуються для поетапного оновлення профілю різними підходами. Така схема симулює довготривалу експлуатацію системи та дозволяє безпосередньо порівняти поведінку методів при накопиченні нових голосових зразків і зміні умов реєстрації.

Профіль користувача реалізовано у вигляді багатовекторного банку, що містить набір векторних представлень голосу разом із параметрами їх пріоритетизації. У межах експериментів реалізовано кілька підходів оновлення: інкрементальне додавання нових векторів до наявного набору (1); оновлення зі зсувом фіксованого розміру банку з витісненням найстаріших записів (2); агреговане оновлення з перерахунком узагальненого представлення профілю (3); а також адаптивне оновлення з перерахунком коефіцієнтів значущості окремих векторів (4). Пріоритетизація векторів налаштована у кількох варіантах: із рівними коефіцієнтами (5), з підвищеним впливом новіших записів (6), а також запропоноване рішення, у якому ваги визначаються рівнем подібності до поточного вхідного зразка (7). Окремо реалізовано схеми оновлення, у яких коефіцієнти нормуються (8), і схеми, де їх вплив задається незалежно без обмеження сумарного значення (9).

Оцінювання ефективності виконано за метриками FAR, FRR та EER, а також за динамікою зміни цих показників у процесі послідовного оновлення профілю. Додатково проаналізовано стабільність профілю, яка полягає у збереженні узгодженості векторного представлення при його поступовому доповненні даними, чутливість до міжсесійних варіацій та вплив різних схем пріоритетизації векторів.

У першому експерименті проведено порівняння підходів оновлення профілю користувача (1-4) за однакового обмеження на максимальний розмір багатовекторного банку ознак. Для кожного мовця сформовано початковий профіль на основі базового набору записів, після чого виконано послідовне оновлення профілю новими сесіями запису мовлення. Додані записи містили природні міжсесійні варіації: зміну темпу мовлення, незначні емоційні коливання та помірні акустичні відмінності середовища запису. У межах експерименту забезпечено однакову кількість оновлень для кожного підходу, що дозволило визначити вплив саме механізму оновлення, а не обсягу даних. Розмір банку обмежувався фіксованим числом векторів, тому в окремих схемах застосовувалось витіснення або агрегування.

Таблиця 1

Порівняння підходів оновлення профілю (фіксований розмір банку)

№ підходу	Підхід оновлення профілю	FAR, %	FRR, %	EER, %	Часові витрати, с
1	Інкрементальне додавання	3.9	9.4	6.6	0.42
2	Зсувне оновлення	4.3	8.0	6.1	0.59
3	Агреговане представлення	5.1	7.2	6.1	0.91
4	Адаптивне з вагами	3.4	6.3	4.9	1.23

Отримані результати, наведені в таблиці 1, показали, що просте інкрементальне накопичення (1) без механізму перерозподілу внеску векторів призводить до підвищення FRR. Зсувне оновлення (2) частково втрачає репрезентативність профілю. Отримані результати підтверджують ефективність адаптивного підходу (4), який надалі використовується як базова схема оновлення профілю в наступних експериментах.

У другому експерименті для адаптивного підходу оновлення профілю (4) досліджено вплив різних схем пріоритетизації векторів (5-9) на якість ідентифікації. Оновлення профілю виконувалося на основі багатосесійних записів, що включали змінні акустичні умови, фоновий шум різної інтенсивності та варіації записувальних пристроїв. Для кожної схеми коефіцієнтів формувался окремий варіант профілю, після чого виконувалось тестування на однаковому контрольному наборі. Порівнювались як схеми з фіксованими вагами, так і варіанти з динамічним визначенням внеску окремих векторів.

Аналіз результатів у таблиці 2 свідчить, що рівномірний розподіл (5) внеску між векторами є найменш ефективним у змінних умовах. Часова пріоритетизація покращує адаптивність до поступових змін голосу. Найменші значення помилок отримано у запропонованому рішенні на основі встановлення ваг за мірою подібності (7) до поточного зразка, що дозволяє зменшувати вплив нетипових або зашумлених записів.

Таблиця 2

Порівняння схем пріоритизації векторів

№ підходу	Схема коефіцієнтів	Опис	FAR, %	FRR, %
5	Рівнозначні	Всі вектори мають однакову вагу, без виділення важливіших	4.6	8.6
6	За часом	Новіші записи отримують більшу вагу, старі поступово втрачають значення	3.8	7.0
7	За подібністю	Ваги визначаються рівнем подібності кожного вектора до поточного вхідного зразка	3.3	6.1
8	Нормовані	Вагові коефіцієнти нормуються, їх сума обмежена одиницею	4.0	7.4
9	Незалежні коефіцієнти	Вплив кожного вектора задається без обмеження сумарного значення	3.9	6.7

У третьому експерименті досліджено вплив розміру багатовекторного банку ознак на якість ідентифікації при використанні запропонованої схеми пріоритизації за подібністю (7). Для кожного мовця формувалися профілі різного обсягу: малий (5 векторів), середній (10 векторів) та розширений (20 векторів), із пропорційною кількістю сесійних записів. Крім того, реалізовувався варіант профілю без обмеження на кількість векторів, що дозволяло оцінити вплив максимальної інформаційної ємності на якість ідентифікації. Оновлення виконувалось послідовно з контролем максимально допустимого розміру банку. Додатково фіксувався час зіставлення як непряма характеристика обчислювальної складності.

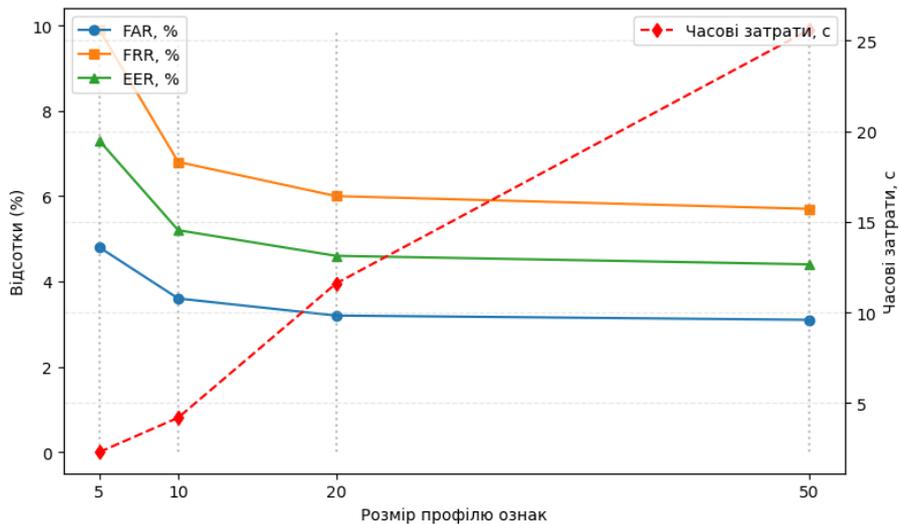


Рис. 3. Графік залежності показників FAR, FRR, EER та часових затрат від розміру банку ознак

Таблиця 3

Залежність показників FAR, FRR, EER та часових затрат від розміру банку ознак

Розмір профіля ознак	FAR, %	FRR, %	EER, %	Часові затрати, с
Малий (5 векторів)	4.8	9.9	7.3	2.3
Середній (10 векторів)	3.6	6.8	5.2	4.2
Розширений (20 векторів)	3.2	6.0	4.6	11.6
Без обмеження	3.1	5.7	4.4	25.6

Згідно з результатами, наведеними в таблиці 3 та графіками, наведеними на рисунку 3, збільшення розміру багатовекторного банку ознак підвищує повноту представлення профілю та зменшує частку хибних рішень, однак призводить до зростання обчислювального навантаження. Банк розмірністю в 10 векторів підвищує точність без істотного зростання часу обробки.

Висновки з даного дослідження і перспективи подальших розвідок у даному напрямі

У роботі запропоновано метод адаптивного оновлення профілю ознак користувача в багатовекторних системах голосової ідентифікації, який передбачає пріоритизацію векторів за їхньою схожістю до поточного сигналу. Проведено аналіз сучасних стратегій оновлення профілю, що слугував підґрунтям для розробки підходу, орієнтованого на підтримку актуальності та стабільності профілю в умовах варіативності голосових характеристик.

Для оцінки ефективності запропонованого методу проведено серію експериментів з різними підходами до оновлення профілю, включно з інкрементальним додаванням, оновленням зі зсувом та агрегованим представленням. Використовувалися профілі різного розміру, що дозволяло порівняти вплив обсягу даних на точність і час обробки. Результати експериментів показали, що адаптивне оновлення за подібністю мінімізує помилки та забезпечує узгоджене зіставлення нових зразків із профілем, одночасно підтримуючи прийнятну точність і обчислювальне навантаження.

Подальший розвиток запропонованого підходу пов'язаний із впровадженням гібридних стратегій оновлення з використанням вагового самоналаштування та автоматичного відбору векторів профілю. Доцільно поєднати багатовекторний підхід із нейромережевою нормалізацією ознак, підвищенням шумостійкості та паралелізацією обчислень для роботи в реальному часі й за обмежених ресурсів.

Література

1. Mou Y., Zhu L. The Effects of Aging on Electroglottographic and Acoustic Parameters of Voices and the Detection of Change Points in Vocal Aging / Y. Mou, L. Zhu // 26th Conference of the Oriental COCODA International Committee for the Co-ordination and Standardisation of Speech Databases and Assessment Techniques (O-COCODA), Delhi, India. – 2023. – P. 1–6. – DOI: 10.1109/O-COCODA60357.2023.10482981.
2. Shofiyah Z., Mahmudah H., Santoso T. B., Puspitorini O., Wijayanti A., Siswandari N. A. Voice Recognition System for Home Security Keys with Mel-Frequency Cepstral Coefficient Method and Backpropagation Artificial Neural Network / Z. Shofiyah, H. Mahmudah, T. B. Santoso, O. Puspitorini, A. Wijayanti, N. A. Siswandari // 2022 International Electronics Symposium (IES), Surabaya, Indonesia. – 2022. – P. 497–501. – DOI: 10.1109/IES55876.2022.9888507.
3. Mishra P., Singh S., Singh S.K., Dixit A. Pre-Processing and Partition of Voice for Semi-Voice Authentication / P. Mishra, S. Singh, S.K. Singh, A. Dixit // Fourth International Conference on Advances in Computing, Communication & Automation (ICACCA), Subang Jaya, Malaysia. – 2018. – P. 1–5. – DOI: 10.1109/ICACCAF.2018.8776849.
4. Бондаренко М. Е., Іващенко Г. С. Використання послідовності методів попередньої обробки в системах голосової ідентифікації / М. Е. Бондаренко, Г. С. Іващенко // Системи управління, навігації та зв'язку. – Полтава: ПНТУ, 2025. – № 2 (80). – С. 90–96. – DOI: 10.26906/SUNZ.2025.2.090.
5. Бондаренко М. Е., Іващенко Г. С. Паралельне виконання методів обробки голосових сигналів на багатоядерних CPU та GPU / М. Е. Бондаренко, Г. С. Іващенко // Системи управління, навігації та зв'язку. – Полтава: ПНТУ, 2025. – № 4 (82). – С. 39–44. – DOI: 10.26906/SUNZ.2025.4.39-44.
6. Shrestha R. Speaker Recognition Using Multiple X-Vector Speaker Representations with Two-Stage Clustering and Outlier Detection Refinement / R. Shrestha // IEEE Intl Conf on Dependable, Autonomic and Secure Computing, Pervasive Intelligence and Computing, Cloud and Big Data Computing, Cyber Science and Technology Congress (DASC/PiCom/CBDCOM/CyberSciTech), Falerna, Italy. – 2022. – P. 1–6. – DOI: 10.1109/DASC/PiCom/CBDCOM/Cy55231.2022.9927875.
7. Rahimi A., Afouras T., Zisserman A. Voicevector: Multimodal Enrolment Vectors for Speaker Separation / A. Rahimi, T. Afouras, A. Zisserman // 2024 IEEE International Conference on Acoustics, Speech, and Signal Processing Workshops (ICASSPW), Seoul, Republic of Korea. – 2024. – P. 785–789. – DOI: 10.1109/ICASSPW62465.2024.10627309.
8. Jaid U.H., AbdulHassan A. End-to-End Speaker Profiling Using 1D CNN Architectures and Filter Bank Initialization / U.H. Jaid, A. AbdulHassan // International Journal of Online and Biomedical Engineering (iJOE). – 2023. – Vol. 19, № 10. – P. 65–81. – DOI: 10.3991/ijoe.v19i10.39061.
9. Mukhamadiyev A., Khujayarov I., Cho J. Voice-Controlled Intelligent Personal Assistant for Call-Center Automation in the Uzbek Language / A. Mukhamadiyev, I. Khujayarov, J. Cho // Electronics. – 2023. – Vol. 12, № 23. – P. 4850. – DOI: 10.3390/electronics12234850.
10. Gopalakrishnan T., Imam S. A., Aggarwal A. Fine Tuning and Comparing Tacotron 2, Deep Voice 3, and FastSpeech 2 TTS Models in a Low Resource Environment / T. Gopalakrishnan, S. A. Imam, A. Aggarwal // 2022 IEEE International Conference on Data Science and Information System (ICDSIS), Hassan, India. – 2022. – P. 1–6. – DOI: 10.1109/ICDSIS55133.2022.9915932.
11. Li T., Wang X., Xie Q., Wang Z., Xie L. Cross-Speaker Emotion Disentangling and Transfer for End-to-End Speech Synthesis / T. Li, X. Wang, Q. Xie, Z. Wang, L. Xie // IEEE/ACM Transactions on Audio, Speech, and Language Processing. – 2022. – Vol. 30. – P. 1448–1460. – DOI: 10.1109/TASLP.2022.3164181.
12. Raharja A.R., Putra V.H.C., Kanugrahan G., Tho C. Indonesian Sign Language (SIBI) Detection: A Landmark-Based Approach with Scikit-Learn Utilizing Random Forest Classifier / A.R. Raharja, V.H.C. Putra, G. Kanugrahan, C. Tho // 2024 Ninth International Conference on Informatics and Computing (ICIC), Medan, Indonesia. – 2024. – P. 1–8. – DOI: 10.1109/ICIC64337.2024.10956917.
13. Yerramreddy D.R., Marasani J., Venkata Gowtham P.S., Harshit G., Anjali. Speech Recognition Paradigms: A Comparative Evaluation of SpeechBrain, Whisper and Wav2Vec2 Models / D.R. Yerramreddy, J. Marasani, P.S. Venkata Gowtham, G. Harshit, Anjali // 2024 IEEE 9th International Conference for Convergence in Technology (I2CT), Pune, India. – 2024. – P. 1–6. – DOI: 10.1109/I2CT61223.2024.10544133.

14. Kacprzak S., Kowalczyk K. Heightceleb – An Enrichment of Voxceleb Dataset With Speaker Height Information / S. Kacprzak, K. Kowalczyk // 2024 IEEE Spoken Language Technology Workshop (SLT), Macao. – 2024. – P. 857–862. – DOI: 10.1109/SLT61566.2024.10832224.

References

1. Mou Y., Zhu L. The Effects of Aging on Electroglottographic and Acoustic Parameters of Voices and the Detection of Change Points in Vocal Aging / Y. Mou, L. Zhu // 26th Conference of the Oriental COCOSDA International Committee for the Co-ordination and Standardisation of Speech Databases and Assessment Techniques (O-COCOSDA), Delhi, India. – 2023. – P. 1–6. – DOI: 10.1109/O-COCOSDA60357.2023.10482981.
2. Shofiyah Z., Mahmudah H., Santoso T. B., Puspitorini O., Wijayanti A., Siswandari N. A. Voice Recognition System for Home Security Keys with Mel-Frequency Cepstral Coefficient Method and Backpropagation Artificial Neural Network / Z. Shofiyah, H. Mahmudah, T. B. Santoso, O. Puspitorini, A. Wijayanti, N. A. Siswandari // 2022 International Electronics Symposium (IES), Surabaya, Indonesia. – 2022. – P. 497–501. – DOI: 10.1109/IES55876.2022.9888507.
3. Mishra P., Singh S., Singh S.K., Dixit A. Pre-Processing and Partition of Voice for Semi-Voice Authentication / P. Mishra, S. Singh, S.K. Singh, A. Dixit // Fourth International Conference on Advances in Computing, Communication & Automation (ICACCA), Subang Jaya, Malaysia. – 2018. – P. 1–5. – DOI: 10.1109/ICACCAF.2018.8776849.
4. Bondarenko M.E., Ivashchenko H.S. Using a Sequence of Pre-Processing Methods in Voice Identification Systems / M.E. Bondarenko, H.S. Ivashchenko // Systems of Control, Navigation and Communication. – Poltava: PNTU, 2025. – № 2 (80). – P. 90–96. – DOI: 10.26906/SUNZ.2025.2.090.
5. Bondarenko M.E., Ivashchenko H.S. Parallel Execution of Voice Signal Processing Methods on Multicore CPUs and GPUs / M.E. Bondarenko, H.S. Ivashchenko // Systems of Control, Navigation and Communication. – Poltava: PNTU, 2025. – № 4 (82). – P. 39–44. – DOI: 10.26906/SUNZ.2025.4.39-44.
6. Shrestha R. Speaker Recognition Using Multiple X-Vector Speaker Representations with Two-Stage Clustering and Outlier Detection Refinement / R. Shrestha // IEEE Intl Conf on Dependable, Autonomic and Secure Computing, Pervasive Intelligence and Computing, Cloud and Big Data Computing, Cyber Science and Technology Congress (DASC/PiCom/CBDCOM/CyberSciTech), Faenza, Italy. – 2022. – P. 1–6. – DOI: 10.1109/DASC/PiCom/CBDCOM/Cy5231.2022.9927875.
7. Rahimi A., Afouras T., Zisserman A. Voicevector: Multimodal Enrolment Vectors for Speaker Separation / A. Rahimi, T. Afouras, A. Zisserman // 2024 IEEE International Conference on Acoustics, Speech, and Signal Processing Workshops (ICASSPW), Seoul, Republic of Korea. – 2024. – P. 785–789. – DOI: 10.1109/ICASSPW62465.2024.10627309.
8. Jaid U.H., AbdulHassan A. End-to-End Speaker Profiling Using 1D CNN Architectures and Filter Bank Initialization / U.H. Jaid, A. AbdulHassan // International Journal of Online and Biomedical Engineering (iJOE). – 2023. – Vol. 19, № 10. – P. 65–81. – DOI: 10.3991/ijoe.v19i10.39061.
9. Mukhamadiyev A., Khujayarov I., Cho J. Voice-Controlled Intelligent Personal Assistant for Call-Center Automation in the Uzbek Language / A. Mukhamadiyev, I. Khujayarov, J. Cho // Electronics. – 2023. – Vol. 12, № 23. – P. 4850. – DOI: 10.3390/electronics12234850.
10. Gopalakrishnan T., Imam S. A., Aggarwal A. Fine Tuning and Comparing Tacotron 2, Deep Voice 3, and FastSpeech 2 TTS Models in a Low Resource Environment / T. Gopalakrishnan, S. A. Imam, A. Aggarwal // 2022 IEEE International Conference on Data Science and Information System (ICDSIS), Hassan, India. – 2022. – P. 1–6. – DOI: 10.1109/ICDSIS55133.2022.9915932.
11. Li T., Wang X., Xie Q., Wang Z., Xie L. Cross-Speaker Emotion Disentangling and Transfer for End-to-End Speech Synthesis / T. Li, X. Wang, Q. Xie, Z. Wang, L. Xie // IEEE/ACM Transactions on Audio, Speech, and Language Processing. – 2022. – Vol. 30. – P. 1448–1460. – DOI: 10.1109/TASLP.2022.3164181.
12. Raharja A.R., Putra V.H.C., Kanugrahan G., Tho C. Indonesian Sign Language (SIBI) Detection: A Landmark-Based Approach with Scikit-Learn Utilizing Random Forest Classifier / A.R. Raharja, V.H.C. Putra, G. Kanugrahan, C. Tho // 2024 Ninth International Conference on Informatics and Computing (ICIC), Medan, Indonesia. – 2024. – P. 1–8. – DOI: 10.1109/ICIC64337.2024.10956917.
13. Yerramreddy D.R., Marasani J., Venkata Gowtham P.S., Harshit G., Anjali. Speech Recognition Paradigms: A Comparative Evaluation of SpeechBrain, Whisper and Wav2Vec2 Models / D.R. Yerramreddy, J. Marasani, P.S. Venkata Gowtham, G. Harshit, Anjali // 2024 IEEE 9th International Conference for Convergence in Technology (I2CT), Pune, India. – 2024. – P. 1–6. – DOI: 10.1109/I2CT61223.2024.10544133.
14. Kacprzak S., Kowalczyk K. Heightceleb – An Enrichment of Voxceleb Dataset With Speaker Height Information / S. Kacprzak, K. Kowalczyk // 2024 IEEE Spoken Language Technology Workshop (SLT), Macao. – 2024. – P. 857–862. – DOI: 10.1109/SLT61566.2024.10832224.