

М. Е. Бондаренко, Г. С. Іващенко

Харківський національний університет радіоелектроніки, Харків, Україна

## БАГАТОВЕКТОРНА ГОЛОСОВА ІДЕНТИФІКАЦІЯ КОРИСТУВАЧА З УРАХУВАННЯМ ФОНЕТИЧНИХ ВАРІАЦІЙ ГОЛОСУ

**Анотація. Актуальність.** Системи голосової ідентифікації широко застосовуються в біометричних системах контролю доступу, персоналізованих застосунках та інтелектуальних інтерфейсах, у яких основною вимогою є надійне розпізнавання користувачів з урахуванням індивідуальних варіацій голосу. Традиційні підходи, що ґрунтуються на статичних векторних представленнях фонетичних мовленнєвих ознак, демонструють обмежену здатність адаптації до нових даних зібраних при довготривалій експлуатації системи біометричної ідентифікації. Це зумовлює зростання похибок через поступові зміни акустичних характеристик та індивідуальних параметрів голосу мовця. **Об'єкт дослідження:** процес формування та оновлення векторних профілів користувачів та їх використання у системах голосової ідентифікації. **Мета статті:** розробка адаптивного методу ідентифікації мовця на основі багатовекторного профілю ознак з оновленням векторних представлень. **Результати дослідження.** Запропоновано метод оновлення векторного профілю ознак, що забезпечує адаптивне формування голосового профілю користувача та підтримує його актуальність у процесі подальшої експлуатації, забезпечуючи можливість поступового донавання. Проведені експериментальні дослідження в умовах фізіологічних змін голосу та середовища запису показали, що використання багатовекторного подання дозволяє знизити показник FRR порівняно із використанням статичних профілів та оновленням лише домінантного вектору. **Висновки.** Результати підтверджують доцільність застосування адаптивного багатовекторного профілю ознак для підвищення стійкості та точності систем ідентифікації мовця.

**Ключові слова:** біометричні системи, системи голосової ідентифікації, ідентифікація мовця, голосовий сигнал, вектор ознак, багатовекторний профіль ознак, фонетичні зміни голосу, адаптивне оновлення, FAR, FRR.

### Вступ

**Постановка проблеми.** Голосова біометрія застосовується для ідентифікації та аутентифікації користувачів, контролю доступу, забезпечуючи надійну перевірку особи без необхідності фізичного контакту або використання додаткових носіїв інформації [1]. Однією з основних переваг систем голосової ідентифікації є низька вартість апаратного забезпечення та простота використання. Для здійснення ідентифікації користувачеві достатньо промовити контрольну фразу або довільний вислів, що зумовлює широке застосування таких систем в фінансових установах, колл-центрах та сервісах віддаленого обслуговування клієнтів [2].

Системи голосової ідентифікації мають ряд обмежень, серед яких слід зазначити чутливість до змін акустичного оточення, фізіологічних варіацій голосу у людини (через втому, захворювання, старіння), а також наявність технологій синтезу або підробки аудіо, що можуть бути використані для обходу біометричної перевірки [3].

Для вирішення вказаних обмежень перед безпосередньою ідентифікацією голосовий сигнал проходить декілька етапів попередньої обробки, таких як видалення шумів, нормалізація гучності та частоти, виділення характеристик спектру сигналу та формування векторних представлень, що служать цифровим описом голосу [1, 4]. Така підготовка дозволяє зменшити вплив акустичних змін середовища та технічних особливостей запису, а також полегшує порівняння сигналів між різними сеансами.

Після формування векторних представлень голосу проводиться ідентифікація шляхом зіставлення вхідного сигналу з загальним банком записів, який містить еталонні голосові зразки всіх зареєстрованих користувачів і слугує для оцінки подібності [5].

Зазвичай кожний користувач представлений одним еталонним вектором (домінантним вектором), який отримується шляхом усереднення декількох записів, отриманих під час реєстрації користувача. Таке представлення дозволяє швидко оцінити схожість голосу користувача із збереженими профілями та прийняти рішення про підтвердження або відмову в доступі. Усереднення векторних представлень або використання одного домінантного вектора часто не дозволяє врахувати внутрішньокласову різноманітність голосових характеристик одного користувача, що включає зміни тембру, інтонації та темпу мовлення. Це негативно впливає на точність і надійність системи ідентифікації, особливо при коротких висловлюваннях або наявності фонового шуму.

**Аналіз останніх досліджень і публікацій.** Для компенсації обмежень підходу з одиничним вектором у сучасних дослідженнях [6, 7] пропонуються методи, що використовують декілька векторних представлень мовця та комбіновані схеми оцінки подібності. Один із поширених підходів полягає у використанні кількох векторів ознак, які відображають акустичні характеристики та стилі мовлення, різні умови запису (зокрема тип і якість мікрофона, відстань до джерела звуку, акустичні властивості приміщення та рівень реверберації, наявність і інтенсивність фонового шуму), співвідношення сигнал/шум, характеристики каналу передавання, а також параметри оцифрування сигналу (частота дискретизації та бітова глибина). Крім того, такі векторні представлення можуть бути отримані з різних моделей або сформовані у різних ознакових просторах [8-10]. При порівнянні нового сигналу обчислюється схожість із усіма наявними представленнями, а фінальне рішення формується на основі агрегації результатів багатьох зіставлень шляхом застосування багатокomпонентних схем оцінювання. Такий

підхід дозволяє ефективніше враховувати міжсесійні варіації голосу, каналові викривлення та зміни акустичного середовища, підвищуючи робастність та точність ідентифікації.

Водночас багатовекторні підходи мають суттєві обмеження: вони істотно збільшують обчислювальне навантаження, потребують більше пам'яті для зберігання векторів ознак, є чутливими до вибору метрик подібності та способів агрегації результатів, а також складно інтегруються у системи, де передбачено використання лише одного еталонного вектора на кожного користувача [10-12].

Окремі рішення потребують значних обсягів даних і складного налаштування гіперпараметрів для навчання та адаптації моделей до нових умов [11], а інші залишаються неадаптивними, фіксуючи набір векторів після початкової реєстрації мовця без можливості їх інкрементального оновлення в процесі експлуатації системи [12]. Статистичні моделі, зокрема PLDA (Probabilistic Linear Discriminant Analysis), також дозволяють враховувати внутрішньокласову варіативність, проте потребують додаткових даних і обчислювальних ресурсів [13].

Розглянуті дослідження підтверджують, що ідентифікація користувача за голосом залишається складною задачею через змінність акустичного середовища, умов запису та фонетичних характеристик голосу, таких як інтонація, наголос і тембр. Використання багатокомпонентних систем оцінки і інтеграція кількох векторних ознак дозволяє підвищити точність і стійкість моделей до фонових шумів, викривлень сигналу та міжсесійних варіацій голосу. Водночас ефективність таких підходів значною мірою залежить від параметризації правил агрегації результатів, вибору метрик подібності та топології ознакового простору. Додатково, складні архітектури нейронних мереж або багаторазові порівняння векторів ознак зумовлюють високі обчислювальні витрати, обмежуючи використання систем голосової ідентифікації у реальному часі.

**Метою роботи** є розробка методу голосової ідентифікації мовця на основі багатовекторного профілю ознак із застосуванням адаптивного оновлення, що забезпечує підтримку актуального голосового профілю користувача.

### Постановка задачі

Ідентифікація користувача за голосом представляє собою комплексну задачу, що потребує формування таких векторів ознак, які відображають індивідуальні особливості голосу та дозволяють надійно розрізняти різних мовців. Класичні підходи використовують один вектор для кожного користувача, що дозволяє швидко здійснювати порівняння, проте недостатньо ефективні в умовах міжсесійних варіацій голосу, шумових перешкодах та змін акустичного середовища.

Пропонується багатовекторний підхід, що дозволяє комбінувати декілька представлень голосу мовного профілю користувача.

Реалізація запропонованого методу голосової ідентифікації повинна передбачати багатовекторне

зберігання ознак користувача, що дозволяє ефективно використовувати пам'ять і забезпечувати високу точність та відтворюваність результатів ідентифікації. У роботі необхідно дослідити вплив структури та способу оновлення векторних представлень на результати ідентифікації користувачів за голосом.

### Основний матеріал

Після попередньої обробки сигналу [1] кожен голосовий запис перетворюється на багатовимірний вектор ознак, що відображає спектральні, енергетичні та часові характеристики голосу, такі як MFCC. Векторне представлення ознак дозволяють кількісно порівнювати голосові записи за допомогою косинусної схожості, евклідової відстані або інших метрик.

На етапі ідентифікації новий вхідний вектор ознак порівнюється з вектором користувача, збереженим у базі даних. У класичних системах цей вектор представляє собою статичний голосовий профіль користувача, що інтегрує інформацію про його мовлення з попередніх записів.

Пропонується підхід на основі багатовекторного представлення ознак користувача, де один вектор (домінантний) має найбільший вплив на прийняття рішення, а допоміжні вектори компенсують варіативність характеристик голосу та умов запису. Після етапу попередньої обробки та виділення базових ознак для кожного голосового запису система ідентифікації формує структурований профіль користувача (1), що складається з доміантного вектора та набору допоміжних векторів, готового до використання в процесі ідентифікації:

$$V = \{v_c, v_1, v_2, \dots, v_N\}, \quad (1)$$

де  $v_c$  – доміантний вектор, який інтегрує інформацію про попередні успішні ідентифікації та визначає основну подібність нового запису до профілю користувача,  $v_{i=1..N}$  – допоміжні вектори, що враховують варіації голосу та умов запису. Кількість допоміжних векторів  $N$  визначається залежно від необхідної робастності та наявного обсягу пам'яті.

У запропонованому багатовекторному підході кожен з 10 векторів ознак має власний ваговий коефіцієнт, що визначає його вплив на фінальне рішення системи ідентифікації. Домінантний вектор користувача отримує найбільший ваговий коефіцієнт  $\alpha_c$ , оскільки він відображає найактуальніші характеристики голосу. Допоміжні вектори, що відповідають записам з попередніх сесій перевірки користувача, отримують поступово зменшувані коефіцієнти, що дозволяє компенсувати вплив змін голосу та умов запису:

$$\alpha_c > \alpha_1 \geq \alpha_2 \geq \dots \geq \alpha_N, \quad \sum_{i=0}^N \alpha_i = 1. \quad (2)$$

Здатність системи забезпечувати коректне розрізнення голосів користувачів досягається шляхом підбору значень ваг на наборах записів у складних акустичних умовах, зокрема за наявності шуму, реверберації, варіативності тембру та інших фізіологічних особливостей голосу.

Зміна параметра  $\alpha_c$  визначає чутливість системи до нових записів, тоді як використання допоміжних векторів ( $v_1, v_2, \dots, v_N$ ) дозволяє враховувати природні міжсесійні зміни голосу користувача без збирання застарілих даних та надмірного обсягу даних профілю. Зменшення ваг допоміжних векторів може призвести до втрати робастності системи, а надмірне збільшення – до зниження чутливості до нових записів. Для оцінки того, наскільки вхідний вектор ознак  $v_{xid}$  відповідає вектору профілю  $v_i$ , використовується міра збіжності між ними. Ця міра (ступінь подібності) показує, наскільки схожі два вектори у багатовимірному ознаковому просторі.

Ступінь подібності визначається на основі косинусної міри, нормованої в межах від 0 до 1: 0 відповідає повній відмінності напрямків векторів, а 1 – їхній повній збіжності, тобто максимальної схожості за акустичними характеристиками голосу:

$$\text{sim}_{\cos}(v_i, v_{xid}) = \frac{v_i \cdot v_{xid}}{\|v_i\| \cdot \|v_{xid}\|}. \quad (3)$$

Підсумкова оцінка схожості між вхідним вектором та профілем користувача обчислюється як зважена сума індивідуальних значень подібності:

$$S = \sum_{i=1}^N \alpha_i \cdot \text{sim}(v_i, v_{xid}), \quad (4)$$

де  $\alpha_i$  – ваговий коефіцієнт вектору профілю, причому

$$\sum_{i=1}^N \alpha_i = 1, \quad \alpha_i \geq 0. \quad (5)$$

Найбільший коефіцієнт ( $\alpha_c = 0.4$ ) присвоюється домінуючому вектору, оскільки він репрезентує останній успішний результат і має значущий вплив на остаточне рішення системи.

Коефіцієнт 0.2 присвоюється найновішому допоміжному вектору  $\alpha_1$ , який раніше (у попередній сесії голосової ідентифікації) був домінуючим, що дозволяє зберегти дані про недавні голосові сесії без надмірного впливу на фінальне рішення. Допоміжні вектори поступово зменшують свій вплив:

$$\alpha_2 = 0.12, \alpha_3 = 0.10, \dots, \alpha_9 = 0.02).$$

Актуалізація векторів ознак після кожної спроби ідентифікації забезпечує здатність системи адаптуватися до міжсесійних і фізіологічних змін голосу користувача та виконувати поступове донавчання під час використання. Зображений на рис. 1 процес дозволяє зберігати актуальні характеристики голосу у пріоритетних векторах багатовекторного профілю та одночасно компенсувати варіації минулих сесій.

Коли новий вхідний запис  $x$  успішно ідентифікований, він стає новим домінуючим вектором  $v_c^{new}$ , зсуваючи на наступні позиції попередній домінуючий та допоміжні вектори у багатовекторному профілі ознак. Допоміжний вектор, який перебуває на кінцевій позиції багатовекторного профілю, видаляється, що дозволяє обмежити розмір профілю та покращити використання пам'яті. У разі неуспішної ідентифікації жоден з векторів не оновлюється,

що зберігає попередній стан профілю користувача та забезпечує робастність системи.

У випадках, коли багатовекторний профіль ознак користувача ще не сформований повністю, наповнення його векторів передбачає поступове формування: перший успішно ідентифікований запис стає домінуючим вектором  $v_c$ , а наступні успішні записи послідовно додаються як допоміжні вектори  $v_1, v_2, \dots$ , досягнувши максимальної заданої кількості  $N$ . При цьому вагові коефіцієнти закріплюються за відповідними позиціями векторів у багатовекторному профілі, що забезпечує збереження структури кожного компонента, тоді як самі вектори послідовно оновлюються: домінуючий вектор зберігає найбільший вплив, а нові допоміжні вектори отримують поступово зменшувані значення. Такий підхід забезпечує адаптивне формування профілю користувача після перших ідентифікацій і реакцію системи ідентифікації на можливі майбутні зміни голосу та умов запису.

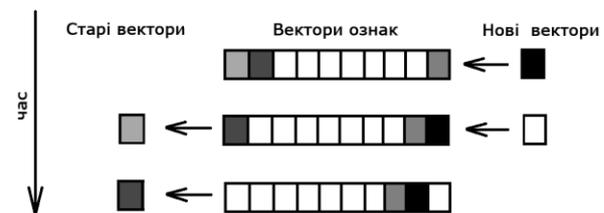


Рис. 1. Процес оновлення векторів ознак

Обчислення схожості виконується за допомогою бібліотеки Faiss, яка призначена для ефективного пошуку найближчих сусідів у великих векторних просторах [14]. Faiss забезпечує як точний, так і наближений пошук у багатовимірних просторах векторних ознак, підтримує апаратне прискорення за допомогою GPU та паралельне виконання обчислень із використанням багатопоточності на CPU.

Для оцінки ефективності системи необхідно проаналізувати роботу при різних обсягах даних, якості записів та наповненості векторів ознак у голосовому профілю користувача.

## Результати експериментальних досліджень

Експериментальна модель системи ідентифікації користувача на основі голосових сигналів розгорнута в середовищі Jupyter Notebook. Модель передбачає використання CPU і GPU для виконання різних етапів обробки та порівняння векторів ознак. Рівні хибного прийняття (FAR) та хибного відхилення (FRR) є основними критеріями оцінки помилок у системах голосової ідентифікації, відображаючи загальну чутливість і точність системи. FAR показує частку випадків, коли система помилково приймає сторонній голос за власного користувача, а FRR – коли система помилково відхиляє власного користувача, навіть якщо його вектори ознак присутні в банку системи.

Результати першого експерименту, наведені в табл. 1, показують вплив різних способів оновлення векторного профілю користувача на точність ідентифікації. Порівнювалися статичний профіль без оновлення, оновлення лише домінуючого вектора

та повне оновлення векторного профілю (запропоноване рішення). Експерименти проводилися на даних у стандартних умовах та за складних акустичних умов, із профілями різної векторної розмірності (одновекторний та багатовекторний підхід). Враховувалися фонові шуми, зокрема вулиці та кондиціонера, а також фонетичні варіації голосу: швидка мова, різна інтонація та зміни артикуляції.

Таблиця 1 – Значення помилок FAR та FRR при різних способах оновлення вектору ознак

Умови експериментів	Спосіб оновлення векторів					
	Не оновлюється		Лише домінуючий вектор		Повне оновлення	
	FAR, %	FRR, %	FAR, %	FRR, %	FAR, %	FRR, %
Профіль з 1 вектору	5.1	4.3	3.7	3.2	1.1	1.2
Профіль з 10 векторів	6.4	7.2	4.1	3.9	1.3	1.4
Наявність шуму	10.1	11.4	9.4	9.2	4.3	3.7
Складні умови запису	11.3	12.5	9.9	11.3	4.8	3.7
Фонетичні зміни голосу	24.4	28.3	12.3	11.9	1.8	1.3

Використаний датасет містить записи, що відображають фізіологічні зміни голосу мовця, зумовлені втому, емоційними чинниками або природними коливаннями. Адаптивне переміщення векторів ознак у просторі ознак дозволило підтримувати стабільну точність ідентифікації. Для тестових сценаріїв із шумами, викривленнями та фонетичними варіаціями голосу користувача використовувався профіль із десяти векторів ознак. Аналіз отриманих результатів показав, що використання неоновлюваного профілю користувача характеризується найгіршими показниками точності, зокрема підвищеним рівнем відмов у розпізнаванні за умов змінних характеристик голосу. Оновлення лише домінуючого вектора частково покращує результати, однак не забезпечує повної компенсації варіативності голосових параметрів мовця. Повне оновлення векторного банку забезпечує найвищу стабільність і точність, суттєво знижуючи FRR навіть у складних акустичних умовах.

В ході другого експерименту досліджувався вплив загальної кількості векторів у профілі ознак користувача на ефективність ідентифікації мовця при використанні запропонованого способу оновлення. Розглядалися три конфігурації: профіль користувача представлений одним вектором; профіль користувача із обмеженою кількістю в десять векторів; та необмежений профіль користувача, у якому всі нові вектори зберігалися без видалення попередніх. В табл. 2 наведені результати експериментів, в ході яких максимальний обсяг профілю одного користувача накопичував 100 векторів.

Таблиця 2 – Значення помилок (FAR, FRR) та часу виконання при різній кількості векторів ознак у профілях користувачів

Кількість векторів ознак у профілі	FAR, %	FRR, %	Час обробки, с
1	7.1	27.1	0.8
10	3.1	3.5	2.9
Необмежене накопичування (100 векторів)	2.5	2.8	12.5

Використання необмеженого профілю векторів ознак дозволяє дещо знизити FRR, проте збільшення числа векторів супроводжується зростанням обчислювальної складності та апаратного навантаження, що робить обмежений профіль ознак більш доцільним для практичного застосування.

## Висновки

Реалізовано адаптивний підхід оновлення векторного профілю ознак у системі ідентифікації мовця, що передбачає використання домінуючого та набору допоміжних векторів із фіксованими ваговими коефіцієнтами. Запропонований метод забезпечує оновлення векторів ознак користувача при надходженні нових голосових сигналів, при цьому домінуючий вектор замінюється новим, а попередній стає допоміжним, тоді як допоміжні вектори зміщуються для компенсації впливу варіацій голосу та умов запису. Такий підхід дозволяє формувати актуальний профіль користувача після перших сесій ідентифікації та підтримувати його актуальність протягом довготривалої роботи системи.

На відміну від PLDA та metric-based підходів, запропонований метод розглядає внутрішньокласову варіативність записів одного користувача не як шум, а як джерело дискримінаційної інформації, що забезпечує ідентифікацію спікера при фізіологічних змінах голосу чи у складних умовах запису. Експериментальні результати показують, що адаптивне оновлення векторів підтримує стабільну точність ідентифікації, зменшуючи вплив варіацій голосових ознак.

Подальшим розвитком запропонованого підходу є впровадження адаптації вагових коефіцієнтів залежно від ступеня схожості нових векторних представлень до поточного профілю.

## Конфлікт інтересів

Автори декларують, що не мають конфлікту інтересів стосовно даного дослідження, в тому числі фінансового, особистісного характеру, авторства чи іншого характеру, що міг би вплинути на дослідження та його результати, представлені в даній статті.

## Використання засобів штучного інтелекту

Автори підтверджують, що не використовували технології штучного інтелекту при створенні представленої роботи.

## СПИСОК ЛІТЕРАТУРИ

1. Бондаренко М. Е., Іващенко Г. С. Використання послідовності методів попередньої обробки в системах голосової ідентифікації. Системи управління, навігації та зв'язку. Полтава: ПНТУ. 2025. № 2 (80). С. 90-96. <https://doi.org/10.26906/SUNZ.2025.2.090>

2. Samonte M. J. C., Callejo J. K., Lumbea D. C. N., Ocaya J. C. B. Mitigating Vishing in Digital Banking Through Caller Authentication and Verification Technologies. 2024 14th International Conference on Software Technology and Engineering ICSTE, Macau, China, 2024, pp. 102-108. <https://doi.org/10.1109/ICSTE63875.2024.00025>
3. Parida S., Mukharjee T., Panda S., Pati B. Deep-Detector: Deepfake Voice Recognition using Machine Learning. 2025 Int. Conf. on ICC (ICoICC), Bhubaneswar, 2025, pp. 1-5. <https://doi.org/10.1109/ICoICC64033.2025.11052136>
4. Бондаренко М. Е., Іващенко Г. С. Організація паралельного виконання методів обробки голосових сигналів на багатоядерних CPU та GPU. Системи управління, навігації та зв'язку. Полтава: ПНТУ. 2025. № 4 (82). С. 39-44. <https://doi.org/10.26906/SUNZ.2025.4.39-44>
5. Feng C., Leer E., Anderson D. V. Identification of Voice Quality Variation Using I-Vectors. 2019 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA), New Paltz, NY, USA, 2019, pp. 105-109. <https://doi.org/10.1109/WASPAA.2019.8937289>
6. Shrestha R. Speaker Recognition using Multiple X-Vector Speaker Representations with Two-Stage Clustering and Outlier Detection Refinement. IEEE Intl Conf on Dependable, Autonomic and Secure Computing, Pervasive Intelligence and Computing, Falerna, Italy, 2022, pp. 1-6. <https://doi.org/10.1109/DASC/PiCom/CBDCCom/Cy55231.2022.9927875>
7. Rahimi A., Afouras T., Zisserman A. Voicevector: Multimodal Enrolment Vectors for Speaker Separation. 2024 IEEE International Conference on Acoustics, Speech, and Signal Processing Workshops (ICASSPW), Seoul, Korea, Republic of, 2024, pp. 785-789. <https://doi.org/10.1109/ICASSPW62465.2024.10627309>
8. Zhang J., Liss J., Jayasuriya S., Berisha V. Robust Vocal Quality Feature Embeddings for Dysphonic Voice Detection. IEEE/ACM Trans. on ASL Proc., vol. 31, pp. 1348-1359. <https://doi.org/10.1109/TASLP.2023.3261753>
9. Yi L., Mak M. W. Disentangled Speaker Embedding for Robust Speaker Verification. 2022 IEEE Int. Conference on Acoustics, Speech and Signal Processing, Singapore, 2022, pp. 7662-7666. <https://doi.org/10.1109/ICASSP43922.2022.9747778>
10. Tiwari, V. K. Speaker identification using multi-modal i-vector approach for varying length speech in voice interactive systems. Cognitive Systems Research, vol. 57, 2019, pp. 66-77. <https://doi.org/10.1016/j.cogsys.2018.09.028>
11. Zhang Y., Bi Z., Xiao F., Yang X., Zhu Q., Guan J. Attacking Voice Anonymization Systems with Augmented Feature and Speaker Identity Difference. 2025 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Hyderabad, India, 2025, pp. 1-2. <https://doi.org/10.1109/ICASSP49660.2025.10889535>
12. Peng S., Sugiyama K., Liu X., Mine T. Balancing Embedding Spectrum for Recommendation. ACM Transactions on Recommender Systems, vol. 3(4), 2025, pp. 1-25. <https://doi.org/10.1145/3718488>
13. Wan Q., Bouchard M. Performance evaluation of mixtures of PLDA and conventional PLDA for a small-set speaker verification system. 2017 IEEE 30th CCECE, Windsor, ON, Canada, 2017, pp. 1-4. <https://doi.org/10.1109/CCECE.2017.7946645>
14. Douze M., Guzhva A., Deng C., Johnson J., Szilvasy G. The Faiss library. IEEE Transactions on Big Data. 2025, pp. 1-17. <https://doi.org/10.1109/TBDATA.2025.3618474>

Received (Надійшла) 19.11.2025

Accepted for publication (Прийнята до друку) 21.01.2026

Publication date (Дата публікації) 27.02.2026

#### ВІДОМОСТІ ПРО АВТОРІВ / ABOUT THE AUTHORS

**Іващенко Георгій Станіславович** – кандидат технічних наук, доцент, доцент кафедри електронних обчислювальних машин, Харківський національний університет радіоелектроніки, Харків, Україна;

**Heorhii Ivashchenko** – Candidate of Technical Sciences, Associate Professor at the Department of Electronic Computers, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine;

e-mail: [heorhii.ivashchenko@nure.ua](mailto:heorhii.ivashchenko@nure.ua); ORCID Author ID: <http://orcid.org/0000-0003-1027-5262>;

Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=57217030807>.

**Бондаренко Максим Едуардович** – аспірант кафедри електронних обчислювальних машин, Харківський національний університет радіоелектроніки, Харків, Україна;

**Maksym Bondarenko** – PhD student at the Department of Electronic Computers, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine;

e-mail: [maksym.bondarenko@nure.ua](mailto:maksym.bondarenko@nure.ua); ORCID Author ID: <http://orcid.org/0000-0002-2500-7626>;

Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=57216159508>.

#### Multidimensional vector voice user identification, including phonetic variations of the voice

Maksym Bondarenko, Heorhii Ivashchenko

**Abstract. Relevance.** Voice identification systems are widely used in biometric access control systems, personalised applications, and intelligent interfaces, where the primary requirement is reliable user recognition, taking into account individual voice variations. Traditional approaches based on static vector representations of phonetic speech features demonstrate limited ability to adapt to new data collected during long-term exploitation of a biometric identification system. It leads to an increase in errors due to gradual changes in the acoustic characteristics and individual parameters of the speaker's voice. **Object of research** is the process of forming and updating vector profiles of users and use in voice identification systems. **Purpose of the article** is to develop an adaptive method for speaker identification based on a multidimensional vector feature bank with updated vector representations. **Research results.** The proposed method of updating the vector bank ensures adaptive formation of the user's voice profile and maintains its relevance during further operation, providing the possibility of gradual retraining. Experimental studies performed under conditions of physiological changes in voice and recording environment variations have shown that the use of multidimensional vector representation allows reducing the FRR compared to the use of static profiles or updating only the dominant vector. **Conclusions.** The results confirm the feasibility of using an adaptive multidimensional vector feature bank to improve the stability and accuracy of speaker identification systems.

**Keywords:** biometric systems, voice identification systems, speaker identification, feature vector, vector comparison, multidimensional vector trait profile, physiological voice changes, adaptive updating, FAR, FRR.