https://doi.org/10.31891/2307-5732-2026-365-14
УДК 004.056.53

ПЕТРУШАК ВОЛОДИМИР
Хмельницький національний університет
https://orcid.org/0000-0002-7232-1044
e-mail: petrushak@ukr.net

ЗАСТОСУВАННЯ ПОПЕРЕДНЬО НАВЧЕНОЇ МОДЕЛІ VOSK У СИСТЕМІ КОНТРОЛЮВАННЯ ДОСТУПУ ДО ОБ'ЄКТУ НА ОСНОВІ ГОЛОСОВОГО ПАРОЛЮ

У сучасному цифровому суспільстві розпізнавання голосу стрімко перетворюється на один із ключових інструментів автентифікації та контролю доступу. Швидкий розвиток штучного інтелекту, нейронних мереж і мобільних пристроїв сприяє впровадженню інноваційних методів, які поєднують високий рівень безпеки з максимальною зручністю для користувачів. У статті досліджено характеристики технологій контролювання доступу, в результаті чого з'ясовано, що сучасні системи контролювання доступу на основі голосового паролю мають: вразливості до replay-атаки, середній рівень безпеки і потребують періодичного оновлення. У зв'язку з цим, голосова аутентифікація виступає своєрідним компромісом між легкістю використання та надійністю. Розроблено алгоритм роботи системи контролювання доступу до об'єкту на основі голосового паролю, що чітко розділяє програмні та апаратні обов'язки, робить рішення відмовостійким і масштабованим. Розроблено структурну схему системи контролювання доступу до об'єкту на основі голосового паролю, яка поєднує гнучкість програмного забезпечення, мінімалізм апаратної частини та високу швидкість реакції, створюючи надійний та доступний механізм контролю доступу на основі голосової аутентифікації, який не потребує складної інфраструктури. Отримані результати досліджень можуть бути використані, для розробки системи контролювання доступу до об'єкту на основі голосового паролю.

***Ключові слова**: голосова аутентифікація, голосовий пароль, нейронна мережа, модель vosk, система контролювання доступу.*

PETRUSHAK VOLODYMYR
Khmelnitskyi National University

APPLICATION OF PRE-TRAINED VOSK MODEL IN FACILITY ACCESS CONTROL SYSTEM BASED ON VOICE PASSWORD

In today's digital society, voice recognition is rapidly becoming one of the key tools for authentication and access control. The rapid development of artificial intelligence, neural networks, and mobile devices contributes to the introduction of innovative methods that combine a high level of security with maximum convenience for users. Among them, voice authentication occupies a special place - a technology that allows you to identify a person based on their unique vocal characteristics. The article examines the characteristics of access control technologies, as a result of which it was found that modern access control systems based on voice passwords have: vulnerabilities to replay attacks, an average level of security, and require periodic updates. In this regard, voice authentication is a kind of compromise between ease of use and reliability. An algorithm for the operation of an object access control system based on voice passwords has been developed, which clearly separates software and hardware responsibilities, makes the solution fault-tolerant and scalable. A structural diagram of a voice password-based access control system for an object has been developed, which combines software flexibility, hardware minimalism, and high response speed, creating a reliable and affordable voice authentication-based access control mechanism that does not require complex infrastructure.

This approach does not require a permanent connection to the Internet, since the entire recognition process occurs locally on a PC, and the user's biometric data is not stored in the form of audio recordings or other formats, which greatly simplifies compliance with personal data protection requirements. Architecturally, the system is built so that each element performs its own clearly defined set of functions: the microcontroller is responsible only for receiving and processing commands, and the computer part takes on the entire volume of calculations and working with the Vosk voice model. This allows the solution to be easily scaled - if there is a need to connect several access points, it is enough to organize the appropriate number of Python script instances on different computers or mini-PCs and connect them to the necessary control boards. The obtained research results can be used to develop an access control system to an object based on a voice password.

***Keywords:** voice authentication, voice password, neural network, vosk model, access control system.*

Стаття надійшла до редакції / Received 11.02.2026
Прийнята до друку / Accepted 11.03.2026
Опубліковано / Published 28.05.2026

Постановка проблеми

У сучасному цифровому суспільстві розпізнавання голосу стрімко перетворюється на один із ключових інструментів автентифікації та контролю доступу. Швидкий розвиток штучного інтелекту, нейронних мереж і мобільних пристроїв сприяє впровадженню інноваційних методів, які поєднують високий рівень безпеки з максимальною зручністю для користувачів. Серед них особливе місце посідає голосова автентифікація — технологія, яка дозволяє ідентифікувати особу на основі її унікальних вокальних характеристик[1].

На відміну від традиційних методів аутентифікації — таких як паролі, картки доступу чи PIN-коди — голосові паролі не потребують фізичних носіїв або запам'ятовування складних комбінацій. Достатньо просто вимовити визначену фразу, щоб система здійснила перевірку та дозволила або заборонила доступ. Такий підхід не лише зручний, а й потенційно безпечніший за класичні засоби, оскільки унікальність голосу важко підробити без спеціалізованих засобів. Система голосової аутентифікації аналізує понад сімдесят параметрів мовлення —

зокрема тембр, ритм, інтонацію, паузи, висоту голосу та швидкість вимови — створюючи на цій основі індивідуальний голосовий «відбиток», подібний до відбитка пальця[2].

Однак, як і будь-яка інша технологія, розпізнавання голосу має не лише переваги, а й певні виклики. Однією з головних проблем є вразливість до атак, зокрема так званих replay-атак, коли зловмисники намагаються обдурити систему, відтворюючи попередньо записаний голос користувача. Крім того, стрімкий розвиток технологій синтезу мовлення — зокрема, генеративних нейронних мереж і deepfake-систем — створює загрозу створення підроблених голосів, які практично неможливо відрізнити від справжніх на слух. Це вимагає від розробників нових підходів до верифікації «живості» мовлення (liveness detection), застосування багатофакторної автентифікації та впровадження криптографічного захисту переданих даних[3].

Не менш важливими є питання приватності та етики. Відповідно до законодавства більшості країн, зокрема й України, біометричні дані — серед яких голос — класифікуються як чутливі персональні дані. Їх обробка потребує явної згоди користувача, чіткого регламенту зберігання, шифрування та обмеженого доступу. Ці вимоги накладають додаткові обмеження на розробку голосових систем та зобов'язують інтеграторів враховувати не лише технічні, а й юридичні аспекти[4].

Аналіз останніх джерел

Сучасні системи контролю доступу визначаються як сукупність апаратних та програмних засобів, спрямованих на регулювання доступу користувачів до певних зон або ресурсів. Вони об'єднують фізичні бар'єри, які можуть включати дверні замки, турнікети, автоматичні ворота, а також електронні компоненти — контролери, зчитувачі, сенсори та сервісні сервери. Основна мета таких систем полягає в забезпеченні безпеки об'єкта, що досягається через чітке розмежування прав користувачів і контроль за їхньою активністю. Згідно з Державним стандартом України ДСТУ 4000-2000, що введено в дію з 2001 року, система контролювання доступу (СКД) вважається комплексом технічних, програмних і організаційних заходів, які слугують для управління доступом, обліку подій та своєчасного реагування на спроби несанкціонованого проникнення. Завдяки уніфікації термінології та вимог, встановлених у цьому стандарті, стало можливим синхронне впровадження різних технологічних рішень, забезпечення сумісності обладнання від різних виробників і інтеграція з іншими системами безпеки, такими як відеоспостереження або протипожежні сигналізації[5].

Технічні вимоги до сучасних СКД включають такі аспекти, як швидкість ідентифікації, надійність апаратної частини, стійкість алгоритмів до спроб підробки, можливість масштабування та інтеграції з іншими інформаційними системами. Ідентифікація та верифікація користувача можуть здійснюватися за допомогою різних методів — від традиційних RFID-карток і PIN-кодів до біометричних технологій, серед яких тепер особливу увагу привертає голосова ідентифікація[6]. У таб. 1 подано порівняльну характеристику технологій контролювання доступу.

Таблиця 1

Порівняльна характеристика технологій контролювання доступу

Критерій	**RFID-картки**	**PIN-коди**	**Відбитки пальців**	**Голосовий пароль**	**Розпізнавання обличчя**
Основні вразливості	Клонування карток	Підгляд/забуття коду	Фальшиві відбитки	Replay-атаки, deepfake	Фотографії, маски
Вартість	Низька $5/картка	Дуже низька	Середня – $100/сканер	Середня – $100/мікро	Висока $200–300
Швидкість	~0.2–0.5 с	~1–2 с	~0.5–1 с	~1–2 с	~0.5–1 с
Зручність	Висока	Середня	Висока	Висока	Висока
Рівень безпеки	Стабільний	Низький–Середній	Середній–Високий	Низький–Середній	Середній–Високий
Необхідність обслуговування	Мінімум	Немає	Періодичне калібрування	Періодичне оновлення	Час від часу калібрування
Масштабованість	Висока	Дуже висока	Середня	Середня	Середня

З таб. 1 можна побачити, що сучасні системи контролювання доступу на основі голосового паролю мають: вразливості до replay-атаки, середній рівень безпеки і потребують періодичного оновлення. У зв'язку з цим, голосова аутентифікація виступає своєрідним компромісом між легкістю використання та надійністю.

Формулювання цілей

На відміну від відбитків пальців чи сканування обличчя, для яких потрібні додаткові спецпристрої (оптичні чи ультразвукові сенсори, високоякісні камери), голосовий метод реалізується здебільшого за допомогою звичайного мікрофона та програмного забезпечення. Однак у порівнянні з картками чи PIN-кодами цей підхід дозволяє створити унікальний «голосовий профіль» користувача, що враховує безліч акустичних ознак: тембр, інтонацію, ритміку мовлення, паузи між словами та інші параметри. Такий профіль значно ускладнює підробку: для успішного обходу системи зловмиснику потрібно не лише отримати запис голосу, а й синтезувати його з урахуванням тих самих властивостей, що створює код, максимально наближений до живого

мовлення. До того ж, сучасні алгоритми аналізу дозволяють оцінювати «живість» аудіо, визначаючи, чи воно було записане в реальному часі або відтворене з запису.

Крім суто технічних переваг, голосова аутентифікація забезпечує високу інтегративну гнучкість. Зокрема, у межах однієї організації її можна використовувати не лише для відкриття фізичних дверей або турнікетів, а й для доступу до корпоративних додатків, автоматизованих терміналів тощо. Оскільки алгоритми розпізнавання голосу можуть бути перенесені майже на будь-який пристрій із мікрофоном і достатньою обчислювальною потужністю, це дає змогу побудувати єдину платформу аутентифікації: від смартфонів до стаціонарних терміналів. Такий підхід значно полегшує адміністрування системи і мінімізує витрати, оскільки не потрібно впроваджувати різноманітні методи ідентифікації для кожного виду доступу. Метою цього дослідження є розробка надійної системи контролю доступу, заснованої на голосовому паролі, що дозволяє обмежено допускати користувачів до певного приміщення без застосування складних біометричних технологій.

Виклад основного матеріалу

У запропонованому підході акцент зроблено не на ідентифікації особистості за унікальними характеристиками голосу, а на точному розпізнаванні наперед визначеного голосового паролю, який порівнюється зі збереженим у програмному коді текстовим шаблоном. Тобто під час реєстрації користувач задає певну фразу, а система запам'ятовує її лише у вигляді текстового рядка, без формування голосових відбитків або збереження аудіозаписів. Завдяки цьому реалізація залишається відносно простою й економічною, проте водночас забезпечує достатній базовий рівень захисту — доступ отримають лише ті, хто здатен чітко вимовити саме ту фразу, що лежить в основі ключа.

До складу системи входить мікроконтролер Atmega328P, який відповідає за апаратну частину контролю доступу — зокрема, керує електромеханічним замком або індикаторами стану — та персональний комп'ютер із Python-скриптом, що займається захопленням і обробкою аудіо. Структурна схема такої системи представлена на рис. 1.

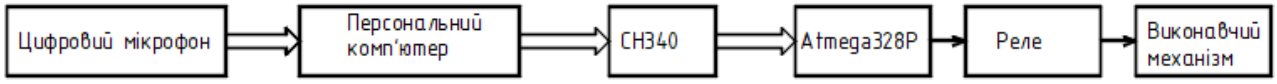

Рис. 1. Структурна схема системи контролювання доступу до об'єкту на основі голосового паролю

Захоплення звуку реалізовано за допомогою цифрового мікрофона з можливістю вибору частоти дискретизації 16 кГц чи 48 кГц і розрядністю 16 біт. Мікрофон підключається до ПК через стандартний аудіоінтерфейс USB-звукової картки, яка виконує оцифрування звукового сигналу та передає його до скрипта на Python. У середовищі Windows або Linux використання бібліотеки PyAudio дозволяє збудувати буферизовану систему отримання аудіопотоку з мінімальною затримкою, а також здійснювати попередню обробку, яка включає нормалізацію рівня сигналу та базове шумоподавлення. Якщо в приміщенні присутні постійні джерела шуму (наприклад, вентиляція чи інше обладнання), алгоритми нормалізації та динамічної компенсації фону знижують кількість хибно негативних розпізнань. Крім того, можна інтегрувати додаткові програми шумозаглушення (наприклад, SpeexDSP), проте для більшості офісних чи лабораторних сценаріїв базової обробки PyAudio достатньо[2].

Практичним доповненням до апаратної платформи є вибір голосового рушія Vosk. Оскільки система покликана працювати в автономному режимі, без постійного доступу до Інтернету, Vosk із рушієм Kaldi став оптимальним варіантом: він може виконувати розпізнавання голосу без з'єднання з віддаленими серверами, що гарантує конфіденційність біометричних даних. Модель Vosk добре масштабована – починаючи від легких лінгвістичних пакетів для вузьких словників і закінчуючи більш об'ємними універсальними модифікаціями, вона дозволяє підібрати оптимальні параметри за швидкістю та точністю розпізнавання. Використання Vosk зменшує затримки обробки аудіопотоку: у типовій конфігурації в умовах офісного ПК з двоядерним процесором та 4 ГБ ОЗП час відгуку на команду не перевищує 200–300 мс, що цілком прийнятно для контролю доступу в реальному часі. Крім того, Vosk підтримує адаптивні словники й можливість донавчання на власних зразках користувача, що особливо актуально при необхідності розгортання системи у багатомовному середовищі чи зі специфічною термінологією. Це дає змогу уникнути додаткових витрат на хмарні сервіси та консультації зовнішніх розробників[7].

Передача команди від Python до Atmega328P відбувається через протокол UART за допомогою бібліотеки pySerial, яка забезпечує конфігурацію серійного порту та передачу одиничного байту (наприклад, «1» для відкриття замка та «0» для скидання стану). Atmega328P з прошивкою на C++ налаштований на швидкість 9600 бод, що дозволяє мінімізувати затримку передачі до кількох мілісекунд. Після отримання байту контролер безпосередньо через транзисторний драйвер керує електромагнітним замком, витримуючи проміжок часу, необхідний для повного відкриття чи закриття, і генерує відповідні сигнали зворотного зв'язку — наприклад, вмикає світлодіод або передає зворотний байт у Python у разі критичних помилок. Такий рівень взаємодії гарантує, що уразливість до мережевих збоїв мінімізована: якщо з якихось причин Python-скрипт припиняє працювати чи переривається USB-канал, Atmega328P лишається в безпечному режимі, не відпускаючи замок до стабільної команди, що підвищує загальну стійкість системи. Алгоритм роботи такої системи представлено на рис.2.

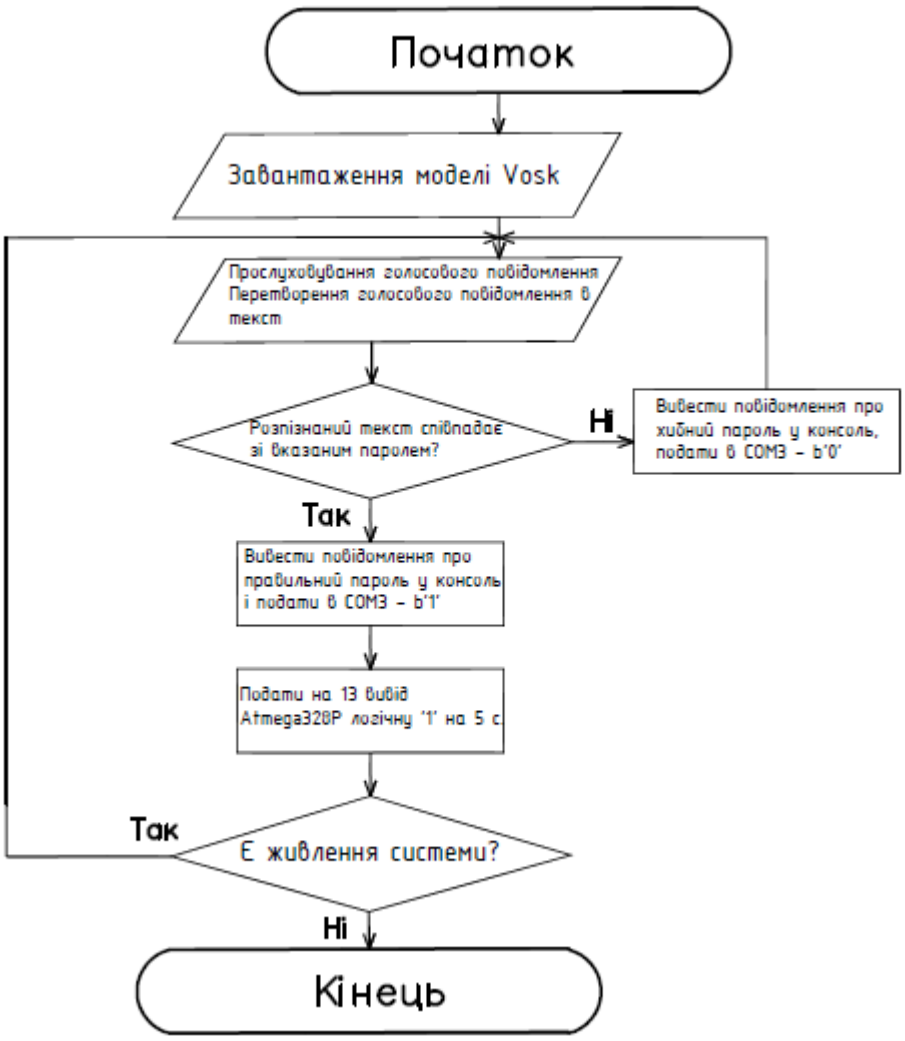

Рис. 2. Алгоритм роботи системи контролювання доступу до об'єкту на основі голосового паролю

Такий підхід не вимагає постійного підключення до мережі Інтернет, оскільки весь процес розпізнавання відбувається локально на ПК, а біометричні дані користувача не зберігаються у вигляді аудіозаписів чи інших форматів, що значно спрощує дотримання вимог захисту персональних даних. Архітектурно система побудована так, щоб кожен елемент виконував свій чітко визначений набір функцій: мікроконтролер відповідає лише за отримання й обробку команд, а комп'ютерна частина бере на себе весь обсяг обчислень та роботу з голосовою моделлю Vosk. Це дозволяє легко масштабувати рішення — якщо виникне необхідність підключити кілька точок доступу, достатньо організувати відповідну кількість екземплярів Python-скрипта на різних комп'ютерах або міні-ПК й з'єднати їх з необхідними платами керування.

Висновки

Досліджено характеристики технологій контролювання доступу, в результаті чого з'ясовано, що сучасні системи контролювання доступу на основі голосового паролю мають: вразливості до replay-атаки, середній рівень безпеки і потребують періодичного оновлення. У зв'язку з цим, голосова аутентифікація виступає своєрідним компромісом між легкістю використання та надійністю.

Розроблено алгоритм роботи системи контролювання доступу до об'єкту на основі голосового паролю, що чітко розділяє програмні та апаратні обов'язки, робить рішення відмовостійким і масштабованим.

Розроблено структурну схему системи контролювання доступу до об'єкту на основі голосового паролю, яка поєднує гнучкість програмного забезпечення, мінімалізм апаратної частини та високу швидкість реакції, створюючи надійний та доступний механізм контролю доступу на основі голосової аутентифікації, який не потребує складної інфраструктури.

Отримані результати досліджень можуть бути використані, для розробки системи контролювання доступу до об'єкту на основі голосового паролю.

Література

1. Biometric Voice Recognition System in the Context of Multiple Languages. Taylor & Francis. [Електронний ресурс]. – Режим доступу https://www.tandfonline.com/doi/full/10.1080/27684520.2024.2362298.

2. Advanced Biometric Voice Verification for Two-Factor Authentication. MDPI. [Електронний ресурс]. – Режим доступу https://www.mdpi.com/2079-9292/12/18/3791.

3. Biometric Authentication—Benefits and Risks. Sumsub. [Електронний ресурс]. – Режим доступу: https://sumsub.com/blog/biometric-authentication-benefits-risks/.

4. Voice Biometrics. ResearchGate. [Електронний ресурс]. – Режим доступу: https://www.researchgate.net/publication/27293606_Voice_Biometrics.

5. Biometric Threats and Exploitation. Identity Management Institute. [Електронний ресурс]. – Режим доступу: https://identitymanagementinstitute.org/biometric-threats-and-exploitation/.

6. Using Biometrics. NCSC.GOV.UK. [Електронний ресурс]. – Режим доступу: https://www.ncsc.gov.uk/collection/device-security-guidance/policies-and-settings/using-biometrics.

7. Voice Biometric Systems for User Identification and Authentication. ResearchGate. [Електронний ресурс]. – Режим доступу: https://www.researchgate.net/publication/360103263_Voice_Biometric_Systems_for_User_Identification_and_Authentication_-_A_Literature_Review .

References

1.Biometric Voice Recognition System in the Context of Multiple Languages. Taylor & Francis. [Electronic resource]. – Access mode: https://www.tandfonline.com/doi/full/10.1080/27684520.2024.2362298.

2.Advanced Biometric Voice Verification for Two-Factor Authentication. MDPI. [Electronic resource]. – Access mode: https://www.mdpi.com/2079-9292/12/18/3791.

3.Biometric Authentication—Benefits and Risks. Sumsub. [Electronic resource]. – Access mode: https://sumsub.com/blog/biometric-authentication-benefits-risks/.

4.Voice Biometrics. ResearchGate. [Electronic resource]. – Access mode: https://www.researchgate.net/publication/27293606_Voice_Biometrics.

5.Biometric Threats and Exploitation. Identity Management Institute. [Electronic resource]. – Access mode: https://identitymanagementinstitute.org/biometric-threats-and-exploitation/.

6.Using Biometrics. NCSC.GOV.UK. [Electronic resource]. – Access mode: https://www.ncsc.gov.uk/collection/device-security-guidance/policies-and-settings/using-biometrics.

7.Voice Biometric Systems for User Identification and Authentication. ResearchGate. [Electronic resource]. – Access mode: https://www.researchgate.net/publication/360103263_Voice_Biometric_Systems_for_User_Identification_and_Authentication_-_A_Literature_Review .