

М.С. Пастушенко, К.І. Корягін

Харківський національний університет радіоелектроніки, Харків

АНАЛІЗ НАПРЯМКІВ ПІДВИЩЕННЯ ЯКОСТІ СИСТЕМ ГОЛОСОВОЇ АУТЕНТИФІКАЦІЇ

Сучасні системи доступу будуються на основі паролів, які мають низьку надійність. Підвищення їхньої надійності зв'язували з використанням статичних біометричних ознак, які не виправдали надії розроблювачів через простоту підробки. У теперішній час проводяться дослідження в області використання в системах доступу динамічних ознак користувача й, у першу чергу, його голосового сигналу. Голосові системи будуються на основі добування ознак користувача з амплітудно-частотного спектра, а фазові дані сигналу ігноруються. Через низьку інформативність отриманих ознак доводиться ускладнювати процедури ухвалення рішення. Експериментальні дослідження показали, що фазо-частотний спектр є більше інформативним і дає можливість виділити в півтора рази більше формантних частот стосовно амплітудно-частотного спектра, що істотно впливає на якість процедур аутентифікації користувача.

Ключові слова: амплітуда, аутентифікація, голосовий сигнал, спектр, частота, фаза.

Вступ

Постановка проблеми у загальному вигляді. Діяльність багатьох фахівців у сучасному суспільстві пов'язана з інфокомунікаційним простором, який забезпечується відповідними технологіями, програмно-алгоритмічним забезпеченням і необхідним обчислювальним ресурсом. Це відноситься до сфери управління, фінансів, виробництва, творчості та до інших видів діяльності. Зараз складніше вказати область діяльності, де інфокомунікаційні технології не використовуються.

В даний час доступ до інформаційних, фінансових, обчислювальних ресурсів здійснюється за допомогою паролів системи аутентифікації користувача, яка не відрізняється високою надійністю. Тому дуже часто у інформаційному просторі з'являються повідомлення про фінансові розкрадання, рідше говорять про крадіжку інформації й незаконне використання обчислювальних ресурсів. Але це не означає, що число таких злочинів менше.

У зв'язку з цим, після відомих подій 09.11.2001 року країни G8 прийняли рішення про розробку систем доступу на основі статичних біометричних ознак користувача, які базуються на дактилоскопії, формі і розмірах особи, візерунку райдужної оболонки і/або сітківки очей, які раніше плідно використовувалися в криміналістиці. Однак, як показало практичне застосування таких пристроїв, вони не вирішили задачу надійного доступу до ресурсів інформаційно-комунікаційних систем. Це обумовлено тим, що шаблони для цих біометричних ознак досить просто підробити. Крім цього, зазначені ознаки мають обмежений інформаційний ресурс для аналізу і не дозволяють його оперативно нарощувати в процесі аутентифікації.

Тому останнім часом велика увага приділяється системам аутентифікації на основі поведінкових біометричних ознак. При цьому численні дослідження присвячені розробці голосових систем аутентифікації [1]. Це обумовлено низкою факторів, серед яких виділимо наступні: отримання й перевірка біометричних даних здійснюється без застосування спеціалізованих та дорогих технологій. Досить мати телефон або мікрофон, який підключений до комп'ютера. Таким чином, голосова аутентифікація не вимагає застосування спеціалізованого дорогого устаткування. Все, що потрібно – це звичайний мікрофон та засоби оцифрування голосового сигналу. Голосова аутентифікація може здійснюватися віддалено через телефону мережу, Інтернет, корпоративні мережі. В розглянутих системах аутентифікації існує можливість оперативно змінювати та нарощувати контрольну фразу, а також використовувати сучасні досягнення цифрової обробки сигналів, які широко і ефективно використовуються в радіолокації та радіозв'язку.

Впровадження досягнень цифрової обробки сигналів дозволяє істотно поліпшити якісні характеристики систем голосової аутентифікації, які базуються на аналізі амплітуди та частоти голосового сигналу, а фазу традиційно ігнорують [1]. У той же час, давно відомо, що фазові дані сигналу є більш інформативними [2]. Тому актуальною науковою задачею є розробка напрямків підвищення якості сучасних систем голосової аутентифікації і, в першу чергу, за рахунок вдосконалення процедур цифрової обробки голосового сигналу.

Мета даної статті – виконати аналіз поточного стану голосових систем аутентифікації і на його основі визначити напрями підвищення їх якісних характеристик. Об'єкт дослідження – процес цифрової

обробки голосових сигналів в системах аутентифікації. Предмет дослідження – моделі і методи цифрової обробки голосових сигналів.

Аналіз існуючих голосових систем аутентифікації. Голосові системи аутентифікації в якості апаратної складової використовують мікрофон і аналогово-цифровий перетворювач, які є в більшості комп'ютерних та інших мобільних пристроїв доступу. Завдання аутентифікації користувача системи доступу раніше в науковій літературі трактувалося як завдання розпізнавання диктору [3]. При цьому в роботах [4–5] встановлено, що відмітні ознаки користувача в основному знаходяться за межами стандартного телефонного каналу у діапазонах 0,1–0,3 кГц; 4–5 кГц; 6,5–7,8 кГц. У системах голосової аутентифікації для ідентифікації використовуються переважно спектральні характеристики амплітуди сигналу користувача. Наприклад, в [6] використовувалася огинаюча спектру голосового джерела, в [7] запропонований метод кепстрального перетворення спектру мовних сигналів, в [8] застосовувалася модель, в якій спектрально-часові характеристики мовного сигналу аналізуються гребінкою фільтрів.

Найважливішим елементом успішного вирішення зазначеного завдання є вибір інформативних ознак (мовних параметрів), здатних ефективно представляти інформацію про особливості мови і анатомії мовного тракту конкретного користувача.

До них ставляться такі вимоги: ефективність представлення інформації про особливості мови конкретної людини; простота вимірювання; стабільність в часі; часта і природна поява в мові; несприйнятливості до імітації.

На етапі вилучення ознак контрольний голосовий сигнал піддається попередній обробці, сегментується на короткі ділянки і на кожній ділянці обчислюється набір ознак. На основі цих ознак формується еталонний і контрольний шаблон користувача.

Процедури голосової аутентифікації здійснюються за допомогою програмно-алгоритмічних засобів. Системи голосової аутентифікації базуються на аналізі амплітуди й частоти голосового сигналу, на основі яких формується еталонний і контрольний шаблон користувача.

Процес голосової аутентифікації починається з попередньої обробки даних, що аналізуються, які в даний час зводяться до виділення контрольної фрази. Далі формується еталонний шаблон, який розміщується в базі і використовується в процесі аутентифікації. При використанні системи аутентифікації також формується контрольний шаблон, який порівнюється з шаблонами бази даних і приймається рішення про допуск користувача.

На етапі формування шаблонів використовуються наступні основні процедури:

- формантний аналіз, який базується на частоті

основного тону реєстрованого голосового сигналу;

- параметри у вигляді коефіцієнтів кепстра, які обчислюються по огинаючій спектру сигналу, що аналізується. На додаток до коефіцієнтів кепстра використовуються також їх перші і другі різниці за часом;

- мел-частотні кепстральні коефіцієнти (Mel Frequency Cepstral Coefficient, MFCC) та їх динамічні характеристики;

- розрахунок коефіцієнтів лінійного передбачення (Linear Prediction Coding Coefficients, LPC), який полягає в тому, що голосові дані трактуються як сигнал на виході лінійної системи з повільно змінними параметрами.

Разом з тим, розпізнавальна здатність такого опису голосового сигналу користувача обмежена. Тому значні зусилля дослідників сконцентровані на розробці вирішальних правил.

Найбільш популярними на етапі прийняття рішення про допуск користувача є методи гауссових сумішей (Gaussian Mixture Model, GMM) і опорних векторів (Support Vector Machine, SVM). Використовуються також штучні нейронні мережі і приховані Марківські моделі (Hidden Markov Models, HMM) [3].

Однак роздільне або комплексне використання зазначених ознак, а також перерахованих вирішальних правил не дозволяє побудувати системи голосової аутентифікації з необхідними якісними характеристиками.

Виклад основного матеріалу

Методика формування фазових даних і результати експериментальних досліджень.

Стосовно до голосового сигналу в системах аутентифікації будемо використовувати фазову інформацію, яка до теперішнього часу широко в розглянутих задачах не застосовувалася. Обумовлено це, очевидно, тим, що до останнього часу не існувало необхідного обчислювального ресурсу і необхідних програмно-алгоритмічних засобів для оперативного формування та обліку фазової інформації.

Виділення і оцінка фазової інформації голосових даних, як відомо, пов'язана з моделлю аналітичного сигналу та оцінкою квадратурної складової.

У практичних додатках, як правило, реєструється дійсна складова аналітичного сигналу $U(t)$, а квадратурну складову $K(t)$ необхідно сформувати алгоритмічно або апаратно. Тут t – незалежна змінна, яка має фізичний зміст одиниці часу. Квадратурна складова має відставання (запізнювання) щодо дійсної частини аналізованого сигналу. Тому процедура формування квадратурної складової може являти собою роботу ідеального фазообертача, який здійснює односторонній фазовий зсув на 90° всіх частотних складових аналізованої (дійсної) частини сигналу.

Математично ця процедура реалізується за допомогою перетворення Гілберта [9]:

$$K(t) = \frac{1}{2} \int_{-\infty}^{\infty} \frac{U(\tau)}{\pi(t-\tau)} d\tau, \quad (1)$$

причому інтеграл тут розуміється в сенсі головного значення. Тут τ – змінна інтегрування.

Нажаль, точність перетворення Гілберта істотно залежить від наявності нелінійних і нестационарних спотворень та ширини спектра аналізованого сигналу, що має місце у полігармонійному голосовому сигналі.

Використання квадратурної (уявної) складової аналітичного сигналу дозволяє значно підвищити якість процедур обробки. Як відомо [9], аналітичний сигнал має наступний вигляд:

$$u(t) = A(t) \exp\{j[\omega_0 t + \varphi(t)]\} = \ddot{U}(t) \exp(j\omega_0 t), \quad (2)$$

де функції $A(t) \geq 0$ та $\varphi(t)$ в явному вигляді задають закони амплітудної і фазової модуляції відповідно, а ω_0 – кругова частота реєстрованого коливання.

$$\begin{aligned} \ddot{U}(t) &= A(t) \exp[j\varphi(t)] = \\ &= A(t) \cos \varphi(t) + jA(t) \sin \varphi(t), \end{aligned} \quad (3)$$

комплексна огинаюча наступного реального вузько-смугового сигналу:

$$U(t) = \text{Re}[u(t)] = A(t) \cos[\omega_0 t + \varphi(t)], \quad (4)$$

а складова сигналу:

$$K(t) = \text{Im}[u(t)] = jA(t) \sin[\omega_0 t + \varphi(t)] \quad (5)$$

є квадратурною (уявною) частиною аналітичного сигналу. Знаючи зазначені складові, можна достатньо просто визначити амплітудну огинаючу, як модуль аналітичного сигналу:

$$M(t) = |u(t)| = \sqrt{U^2(t) + K^2(t)}. \quad (6)$$

При цьому повна фаза являє собою аргумент аналітичного сигналу:

$$\begin{aligned} \Psi(t) &= \arg[u(t)] = \arccos \frac{U(t)}{M(t)} = \\ &= \arcsin \frac{K(t)}{M(t)} = \text{arctg} \frac{K(t)}{U(t)}. \end{aligned} \quad (7)$$

Щоб отримати фазову функцію (або початкову фазу) сигналу, потрібно виділити з повної фази лінійний доданок $\omega_0 t$. Для цього, в свою чергу, необхідно знати значення центральної частоти ω_0 . Таким чином,

$$\varphi(t) = \Psi(t) - \omega_0 t. \quad (8)$$

Швидкість зміни несучого коливання можна розрахувати з використанням наступного співвідношення [8]:

$$\omega(t) = \dot{\Psi}(t) = \frac{U(t)\dot{K}(t) - \dot{U}(t)K(t)}{M^2(t)}. \quad (9)$$

Таким чином, маючи квадратурну складову, можна істотно розширити можливості і якісні характеристики цифрової обробки сигналів, що аналізуються.

При дискретному (цифровому) представленні аналітичного сигналу співвідношення (2) перетвориться до виду:

$$u_i = A_i \exp\{j[2\pi f_0(i-1)/f_0 + \varphi_i]\}, \quad (10)$$

де $i = 1, \dots, N$ – номер відліку сигналу, що аналізуються; N – кількість відліків, що аналізуються; f_0 – частота несучого коливання, що вимірюється, в Гц; f_0 – частота дискретизації голосового сигналу.

Ефективність запропонованих процедур покажемо на прикладі формантного методу формування шаблону користувача [10].

Аналізу піддавався експериментальний голосовий сигнал користувача системи аутентифікації, який виголошував цифру “один”. Частота дискретизації дорівнює 64 кГц, співвідношення сигнал/шум більш ніж 20 дБ.

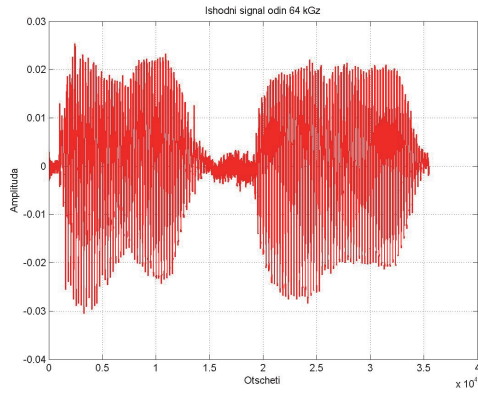
В існуючих голосових системах аутентифікації розраховується за експериментальним голосовим сигналом користувача і в області низьких частот виділяють формантні частоти. У подальшому, як правило, будується їх огинаюча.

У запропонованих процедурах необхідно додатково виконати відновлення квадратурної складової голосового сигналу і розрахувати його фазові дані. Потім необхідно відкоригувати фазові дані та побудувати фазовий спектр. Останній крок – виконати зазначені вище операції (виділення формантних частот і побудова їх огинаючої) з фазо-частотним спектром аналізованого сигналу.

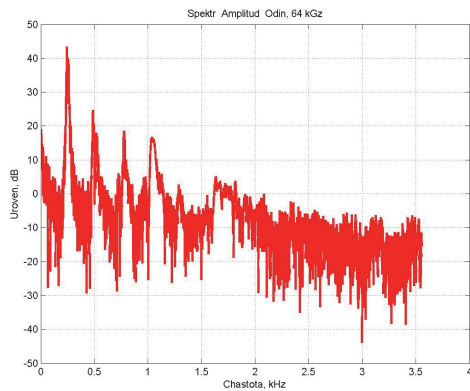
Голосовий сигнал аналізованого сигналу та його амплітудно-частотний спектр (АЧС) представлені на рис. 1. При цьому спектр сигналу обмежимо зверху границею каналу тональної частоти, щоб більш якісно виконати аналіз послідовності основних формант.

Результати обробки формантної інформації цього спектру представлені в табл. 1. Яскраво вираженими в спектрі є чотири форманти, а п'ята має дуже низький рівень спектральної потужності (кілька децибел).

Частота основного тону становить 243 Гц. Друга форманта має частоту в два рази вище. Нажаль, наступні форманти не кратні частоті основного тону.



а



б

Рис. 1. Голосовий сигнал цифри “1” (а) та його короткий спектр (б)

Таблиця 1

Характеристики формант АЧС

Рівень, дБ	43,4	24,6	18,6	14,2
Частота, Гц	243	486	776	1025

Тепер розрахуємо і проаналізуємо фазовий спектр розглянутого сигналу. При цьому треба зробити ряд зауважень.

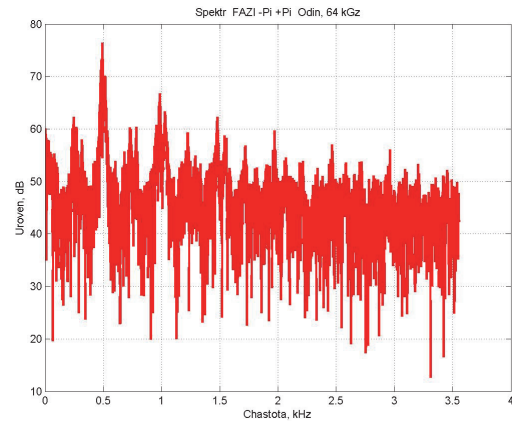
Фаза голосового сигналу не реєструється. Тому її, як правило, розраховують програмно-алгоритмічно.

Нажаль, функція $arctg$ видає значення кута в діапазоні від $-\pi/2$ до $\pi/2$. Тому для визначення правильного значення фазового кута, який у голосового сигналу змінюється в межах від 0 до 2π , необхідно кут $\varphi(t)$ відповідним чином відкоригувати. В іншому випадку фазовий спектр буде некоректним.

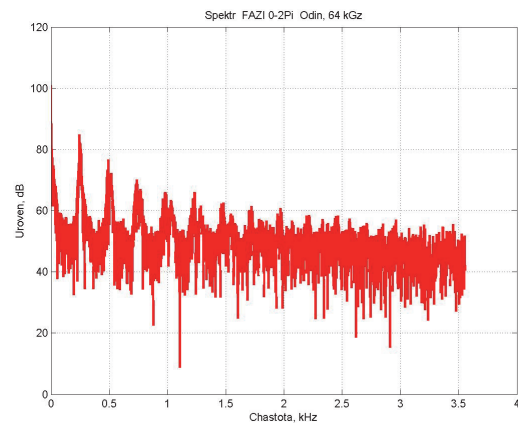
Тепер проаналізуємо фазовий спектр (ФС) сигналу, що аналізується. На рис. 2 представлені два спектри.

Різниця представлених ФС полягає в наступному. На рис. 2, а представлений спектр, який розрахований на основі кутових даних, що отримані за значеннями на виході функції $arctg$. На даному рисунку порушений порядок відображення формант,

характер яких повинен відповідати зображенню формант на рис. 1, б. На рис. 2, б представлений ФС відкоригованого фазового кута, який будемо розглядати нижче.



а



б

Рис. 2. ФС голосового сигналу (а – на виході функції $arctg$; б – відкорегований кут)

Результати обробки формантної інформації ФС представлені в табл. 2. В даному спектрі можна виділити шість формант, а сьома і восьма мають незначну енергетичну відмінність. Частота основного тону, як і в АЧС, становить 243 Гц.

Таблиця 2

Характеристики формант ФС

Рівень, дБ	84,9	76,7	70,3	65	64	62
Частота, Гц	243	492	738	990	1217	1450

Аналіз результатів модельного експерименту свідчить про наступне. Рівень спектральної щільності виділених максимумів в рази перевищує рівень максимумів амплітудного спектру, що істотно спрощує процедуру їх виділення.

Кількість виділених формант у фазового спектра в півтора рази більше. Зазначене вище свідчить про більшу інформативність фазового спектру голосового сигналу.

Висновки

Розглянуто поточний стан сучасних голосових систем аутентифікації користувача. Аналіз цих систем свідчить про те, що при формуванні шаблону в основному використовується амплітудно-частотна інформація, а фазові дані ігноруються.

У дослідженнях запропоновано для підвищення якісних показників голосових систем аутентифікації використовувати фазові дані.

При використанні фазових даних обов'язковою процедурою повинна бути корекція кутових даних на виході функції \arctg до діапазону від 0 до 2π . В іншому випадку отримаємо спотворені результати.

Рівень спектральної потужності формантних

максимумів фазового спектру в кілька разів перевищує максимуми амплітудного спектру, що суттєво спрощує процедури їх виділення.

Кількість виділених формант в фазо-частотному спектрі в півтора рази більше порівняно з амплітудним спектром.

Таким чином, фазовий спектр в області низьких частот (до 4000 Гц) має велику інформативність по відношенню до амплітудного. Тому використання фазової інформації голосового сигналу користувача може істотно поліпшити якісні характеристики систем аутентифікації. Подальші дослідження будуть орієнтовані на пошук характерних ознак користувача по фазовій інформації його голосового сигналу.

Список літератури

1. Beigi H. *Fundamentals of Speaker Recognition* / H. Beigi. – NY: Springer, 2011. – 1029 p.
2. Oppenheim A.V. The Importance of Phase in Signals / A.V. Oppenheim, J.S. Lim // *Proceeding of the IEEE*. – 1981. – Vol. 69(5). – P. 529-541.
3. Сорокин В.Н. Распознавание личности по голосу: аналитический обзор / В.Н. Сорокин, В.В. Вьюгин, А.А. Тананыкин // *Информационные процессы*. – М.: РАН. – 2012. – Т. 12. – № 1. – С. 1-30.
4. Besacier L. Subband architecture for automatic speaker recognition / L. Besacier, J.-F. Bonastre // *Signal Process.* – 2000. – V. 80. – P. 1245-1259.
5. Lu X. An investigation of dependencies between frequency components and speaker characteristics for text-independent speaker identification / X. Lu, J. Dang // *Speech Communication*. – 2007. – Vol. 50. – № 4. – P. 312-322.
6. Sorokin V.N. Speaker verification using the spectral and time parameters of voice signal / V.N. Sorokin, A.I. Tsyplikhin // *Journal of Communications Technology and Electronics*. – 2010. – Vol. 55. – № 12. – P. 1561-1574.
7. Davis S. Comparison of parametric representations for monosyllabic word recognition in continuously spoken sentences / S. Davis, P. Mermelstein // *IEEE Trans. Acoustics, Speech, Signal Process.* – 1980. – Vol. 28. – № 4. – P. 357-366.
8. Patterson R.D. A functional model of neural activity patterns and auditory images / R.D. Patterson, J. Holdsworth // *Advances in Speech, Hearing and Language Processing*. – 1996. – Vol. 3. – P. 547-563.
9. Bendat J.S. *Random data Analysis and Measurement Procedures* / J.S. Bendat, A.G. Piersol. – NY: JOHN WILEY & SONS, 1971. – 541 p.
10. Пастушенко Н.С. Исследование информативности фазовых данных голосового сигнала пользователя системы аутентификации [Электронный ресурс] / Н.С. Пастушенко, В.Г. Педро, О.Н. Файзулаева // *Проблеми телекомунікацій*. – 2018. – № 1 (22). – С. 67-74. – Режим доступу: http://pt.journal.kh.ua/2018/1/1/181_pastushenko_voice.pdf.

References

1. Beigi, H. (2011), *Fundamentals of Speaker Recognition*, Springer, NY, 1029 p.
2. Oppenheim, A.V. and Lim, J.S. (1981), The Importance of Phase in Signals, *Proceeding of the IEEE*, Vol. 69(5), pp. 529-541.
3. Sorokin, V.N., Vyugin, V.V. and Tananyikin, A.A. (2012), “Распознавание личности по голосу: аналитический обзор” [Voice Recognition: An Analytical Review], *Informatsionnyie protsessyi*, Vol. 12, No. 1, pp. 1-30.
4. Besacier, L. and Bonastre, J.-F. (2000), Subband architecture for automatic speaker recognition, *Signal Process*, Vol. 80, pp. 1245-1259.
5. Lu, X. and Dang, J. (2007), An investigation of dependencies between frequency components and speaker characteristics for text-independent speaker identification, *Speech Communication*, Vol. 50(4), pp. 312-322.
6. Sorokin, V.N. and Tsyplikhin, A.I. (2010), Speaker verification using the spectral and time parameters of voice signal, *Journal of Communications Technology and Electronics*, Vol. 55(12), pp. 1561-1574.
7. Davis, S. and Mermelstein, P. (1980), Comparison of parametric representations for monosyllabic word recognition in continuously spoken sentences, *IEEE Trans. Acoustics, Speech, Signal Process*, Vol. 28(4), pp. 357-366.
8. Patterson, R.D. and Holdsworth, J. (1996), A functional model of neural activity patterns and auditory images, *Advances in Speech, Hearing and Language Processing*, Vol. 3, pp. 547-563.
9. Bendat, J.S. and Piersol, A.G. (1971), *Random data Analysis and Measurement Procedures*, JOHN WILEY & SONS, NY, 541 p.
10. Pastushenko, N.S., Pedro, V.G. and Fayzulaeva, O.N. (2018), “Issledovanie informativnosti fazovyih dannyih golosovogo signala polzovatelya sistemyi autentifikatsii” [The study of the information content of the phase data of the voice signal of the user of the authentication system], *Problemi telekomunikatsiy*, Vol. 1 (22), pp. 67-74, available at: www.pt.journal.kh.ua/2018/1/1/181_pastushenko_voice.pdf.

Відомості про авторів:**Пастушенко Микола Савелійович**

кандидат технічних наук професор
професор кафедри Харківського національного
університету радіоелектроніки,
Харків, Україна
<https://orcid.org/0000-0003-2664-1167>

Корягін Костянтин Ігоревич

магістр кафедри Харківського національного
університету радіоелектроніки,
Харків, Україна
<https://orcid.org/0000-0001-5076-3186>

Information about the authors:**Mykola Pastushenko**

Candidate of Technical Sciences Professor
Professor of Department of Kharkiv National
University of Radio Electronics,
Kharkiv, Ukraine
<https://orcid.org/0000-0003-2664-1167>

Kostiantyn Koriahin

Master student of Kharkiv National
University of Radio Electronics,
Kharkiv, Ukraine
<https://orcid.org/0000-0001-5076-3186>

АНАЛИЗ НАПРАВЛЕНИЙ ПОВЫШЕНИЯ КАЧЕСТВА СИСТЕМ ГОЛОСОВОЙ АУТЕНТИФИКАЦИИ

М.С. Пастушенко, К.И. Корягин

Современные системы доступа строятся на основе паролей, которые имеют низкую надежность. Повышения их надежности связывали с использованием статических биометрических признаков, что не оправдало надежды разработчиков из-за простоты подделки. В настоящее время проводятся исследования в области использования в системах доступа динамических признаков пользователя и, в первую очередь, его голосового сигнала. Голосовые системы строятся на основе извлечения признаков пользователя из амплитудно-частотного спектра, а фазовые данные сигнала игнорируются. Из-за низкой информативности полученных признаков приходится усложнять процедуры принятия решения. Экспериментальные исследования показали, что фазочастотный спектр является более информативным и дает возможность выделить в полтора раза больше формантных частот относительно амплитудно-частотного спектра, что существенно влияет на качество процедур аутентификации пользователя.

Ключевые слова: амплитуда, аутентификация, голосовой сигнал, спектр, частота, фаза.

ANALYSIS ON QUALITY ENHANCEMENT OF VOICE AUTHENTICATION SYSTEMS

M. Pastushenko, K. Koriahin

Modern methods of managing financial, informational and other resources are carried out using access systems that are based on passwords. These systems are simple and convenient, but have low reliability. This is confirmed by almost daily reports of financial embezzlement. Improving the reliability of access systems associated with the use of static biometric features, such as fingerprints of a user. Unfortunately, these expectations were not met since the latter can be easily faked. Therefore, intensive research is currently being carried on the use of dynamic (behavioral) features of a user and, above all, his voice signal in access systems. However, voice authentication systems are not very reliable. The aim of this work is to analyze the current state of voice user authentication systems in infocommunication systems and, on its basis, determine the directions for improving the quality characteristics of access systems. The current state of voice systems is analyzed the given paper. Studies have shown that at present, these systems are based on extracting user characteristics from the amplitude-frequency spectrum, and the phase data of the voice signal is ignored. Due to the low informativeness of the features obtained, it is necessary to complicate decision-making procedures. The main attention in the given article is focused on the procedures for the formation and use of the phase information of the voice signal. In the process of experimental studies, it was found that the phase-frequency spectrum is more informative and makes it possible to select one and a half times more formant frequencies with respect to the amplitude-frequency spectrum. The latter will significantly affect the quality of user authentication procedures. Further studies will focus on finding the distinctive features of a user according to the phase information of his voice signal.

Keywords: amplitude, authentication, voice signal, spectrum, frequency, phase.