

**Т. В. Педоренко**  
**Г. Б. Язмухамедова**  
**О. Ю. Софина**  
**В. М. Севастьянов**

## **ПАРАМЕТРИЗАЦІЯ МОВНОГО СИГНАЛУ ДЛЯ ІДЕНТИФІКАЦІЇ ЛЮДИНИ ЗА ОСОБЛИВОСТЯМИ ГОЛОСУ**

Вінницький національний технічний університет

### **Анотація**

*В тезах доповіді представлено створену систему автоматичної ідентифікації людини за особливостями голосу. Наведено результати тестування створеної системи.*

**Ключові слова:** ідентифікація людини за особливостями голосу, автентифікація, сегментація, мажоритарне голосування

### **Abstract**

*The created system of automatic human identification by voice characteristics is presented in given abstract. The results of the created system testing are given.*

**Keyword:** identification of a person by voice features, authentication, segmentation, majority voting

### **Вступ**

Ідентифікація людини за особливостями голосу – процес встановлення особи на основі аналізу інформації, отриманої з мовного сигналу. Перевагою даного підходу є, насамперед, невіддільність мовної інформації, природність мовного спілкування для людини. Потреба в ідентифікації людини за особливостями голосу виникає, наприклад, в криміналістиці, де по запису розмови необхідно ототожнити голос мовця з голосом одного з багатьох підозрюваних. В системах безпеки ідентифікація людини за особливостями голосу найчастіше застосовується для контролю за доступом до інформації (наприклад, у банківських телефонних системах для видачі клієнту конфіденційної інформації).

### **Результати дослідження**

В даних тезах представлено створену систему автоматичної ідентифікації людини за особливостями голосу, структурна схема якої приведена на рис. 1.

Розглянемо роботу системи. Система працює в 2 етапи – етап навчання та етап безпосередньо ідентифікації мовця. Фактично, етап навчання системи складається з таких кроків:

1. Визначення кількості дикторів, серед яких система буде проводити процедуру автентифікації.
2. Визначення паролного матеріалу (слово, фраза).
3. Отримання даних для навчання системи (кожен з дикторів промовляє паролний матеріал  $m$  разів).
4. Навчання класифікаторів по отриманому матеріалу.

Після виконання кроку 4 система готова до роботи. Потрібно зауважити, що від кількості навчальних даних залежить точність ідентифікації. Збільшуючи навчальну вибірку ми збільшуємо точність автентифікації. Після завершення етапу навчання система готова до роботи, яка виконується наступним чином. Людина промовляє в мікрофон паролну фразу, і за допомогою стандартної програми “Звукозапис” отримується WAV-файл (запис проводиться в форматі PCM, частота дискретизації складає 8 кГц), тобто проводиться аналогово-цифрове перетворення. На структурній схемі з рис. 1 вищевказану функції виконує блок ПЗІ (попередній збір інформації). Далі в роботу

вступає блок ППС (попередньої обробки сигналу). Його функція наступна – одержаний запис мовного матеріалу сегментується на інформативні і неінформативні ділянки використовуючи методи нульових перетинів та енергії сигналу. Ділянки запису мовного сигналу, які визнаються як інформативні, вирізаються в окремий масив і нормуються по довжині звучання використовуючи метод динамічного програмування. Одержаний масив даних поступає на блок виділення інформативних ознак, тобто проводиться аналіз мовного сигналу на основі амплітудних значень короткочасних спектрів на інтервалах стаціонарності сигналу. В результаті аналізу одержуємо вектори параметрів, що характеризують мовний матеріал, які подаються на входи відповідних класифікаторів (блок прийняття поточних рішень). Для прийняття остаточного рішення використовується блок ПОС (прийняття остаточного рішення), необхідність якого обумовлена наступним – в процесі роботи системи може виникнути ситуація, коли різні класифікатори приймуть не співпадаючі рішення. Для того, щоб уникнути плутаниці в прийнятті рішень в такій ситуації використовується метод мажоритарного голосування.

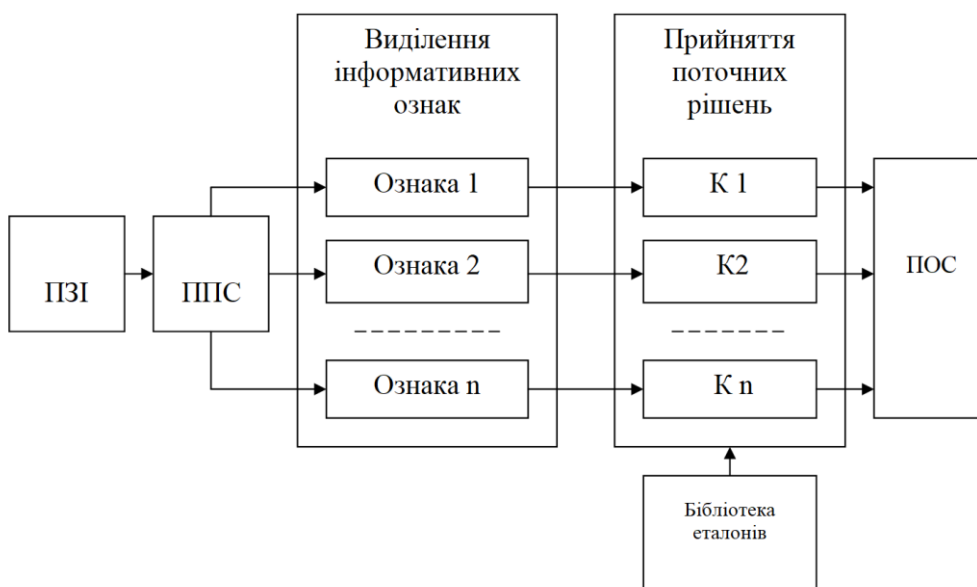


Рисунок 1 – Структурна схема автоматичної системи ідентифікації людини за особливостями голосу

Результати тестування автоматичної системи ідентифікації людини за особливостями голосу представлені у таблиці 1, де  $N_2$  - кількість нейронів на проміжному шарі класифікатора перцептрона,  $t$  – тривалість навчання класифікатора,  $T$  – точність ідентифікації.

Таблиця 1 – Результати тестування

| $N_2$ | Метод навчання                                 | $t$ , с | $T$ , % |
|-------|--|---------|---------|
| 10    | Levenberg-Marquardt backpropogation            | 96      | 63      |
|       | Gradient dencent backpropogation               | 124     | 82      |
|       | Gradient dencent with momentum backpropogation | 114     | 82      |
|       | RPROP backpropogation                          | 103     | 83      |
| 20    | Levenberg-Marquardt backpropogation            | 107     | 64      |
|       | Gradient dencent backpropogation               | 142     | 84      |
|       | Gradient dencent with momentum backpropogation | 128     | 83      |
|       | RPROP backpropogation                          | 126     | 86      |
| 30    | Levenberg-Marquardt backpropogation            | 112     | 63      |
|       | Gradient dencent backpropogation               | 149     | 84      |
|       | Gradient dencent with momentum backpropogation | 132     | 84      |
|       | RPROP backpropogation                          | 126     | 86      |

Як видно з таблиці найкраще співвідношення часу навчання з точністю автентифікації в метода RPROP backpropagation. Середня точність автентифікації за даною ознакою склала 86%.

### Висновки

Отже, результати тестування системи дозволили підтвердити достовірність отриманих практичних результатів. Втім, як виявилось при огляді ринку сучасних систем автоматичної ідентифікації диктора, точність ідентифікації для таких систем становить 94-97%. Як видно з одержаних результатів, розроблена система має комерційну цінність на ринку систем біометричної ідентифікації.

### СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Биков М.М., Ковтун В.В., Раїмі А., Мазур О.В. Метод нормалізації тривалості звучання паролних фраз для системи розпізнавання мовців // Вісник ВПШ. - №4. - 2015. – С.161 – 167;
2. Рамишвили Г.С. Автоматическое опознавание говорящего по голосу. – М.: Радио и связь, 1981. – 224 с.
3. Рабинер Л.Р. Шафер Р.В. Цифровая обработка речевых сигналов: Пер. с англ. – М.: Радио и связь, 1981. – 496 с.
4. Сергиенко А.Б. Цифровая обработка сигналов: Учебное пособие. Второе издание.—СПб.: Питер, 2006.—752 с.

**Педоренко Тетяна Вікторівна** — студентка групи ІАКІТ-176, факультет комп'ютерних систем і автоматики, Вінницький національний технічний університет, с.Агрономічне, Вінницького р-ну, Вінницької обл., e-mail: pedorenkotana@gmail.com

**Язмухамедова Гозел Бабамурадівна** — студентка групи ІАКІТ-176, факультет комп'ютерних систем і автоматики, Вінницький національний технічний університет, м.Вінниця, e-mail: gozelyazmuhamedowa@gmail.com

Науковий керівник: **Софіна Ольга Юрївна** — к.т.н., доцент кафедри автоматизації та інтелектуальних інформаційних технологій, Вінницький національний технічний університет, м.Вінниця

Науковий керівник: **Севастьянов Володимир Миколайович** — к.т.н., доцент кафедри автоматизації та інтелектуальних інформаційних технологій, Вінницький національний технічний університет, м.Вінниця

**Pedorenko Tetiana** — Department of Computer System and Automation, Vinnytsia National Technical University, village Agronomichne, Vinnytsia district, Vinnytsia region, e-mail: pedorenkotana@gmail.com

**Yazmuhamedowa Gozel** — Department of Computer System and Automation, Vinnytsia National Technical University, Vinnytsia, e-mail: gozelyazmuhamedowa@gmail.com

Supervisor: **Sofina Olga** — Ph.D., Associate Professor of Automation and Intellectual Information Technology, Vinnytsia National Technical University, Vinnytsia

Supervisor: **Sevastyanov Volodymyr** — Ph.D., Associate Professor of Automation and Intellectual Information Technology, Vinnytsia National Technical University, Vinnytsia