

УДК 004.522

## ОСОБЛИВОСТІ ПОБУДОВИ ІНТЕРФЕЙСУ ЛЮДИНО-МАШИНОЇ ВЗАЄМОДІЇ ДЛЯ ВИКОРИСТАННЯ ЛЮДЬМИ З ДЕФЕКТАМИ МОВЛЕННЯ

Мариморич В.Є.

к.т.н., доцент каф. ІС Годовиченко М.А.

Одеський Національний Політехнічний Університет, УКРАЇНА

**АНОТАЦІЯ.** У роботі розглядається проблема людино-машинної взаємодії за допомогою голосових команд для людей з дефектами мовлення. Дається коротка характеристика основним дефектам мовлення та визначається власний алгоритм роботи з дефектом заїкання.

**Вступ.** На сьогоднішній день існує велика кількість інтерфейсів природної людино-машинної взаємодії. Найбільше практичне використання здобув мовний інтерфейс взаємодії. Він є майже незамінним для людей з вадами зору або опорно-рухової системи, але якщо така людина буде мати також і дефекти мовної активності, використання інтерфейсу стане неможливим. Сучасні сервіси та інтерфейси дуже погано розпізнають слова людей з дефектами мовлення. Тому потрібно покращити сучасні способи розпізнавання мовних команд або створити інші методи.

**Мета роботи.** Покращення правильності розпізнавання команд управління, отриманих з аудіо сигналу, під час розпізнаванні слів людей із дефектами мовної активності з подальшим використанням команд у системах людино-машинної взаємодії.

**Основна частина роботи.** Системи розпізнавання мови на сучасному етапі не можуть забезпечити правильне розпізнавання слів усіх людей через діалектичні особливості, мовні проблеми або відсутність підтримки мови людини подібною системою, але найбільшою проблемою є важкість розпізнавання мови людей з дефектами мовної активності. Тільки у США більше дев'яти мільйонів людей не можуть використовувати мовну взаємодію через подібні проблеми, а саме інтерфейс *Siri*, розроблений компанією *Apple*, або сервіс *Google Now*[1]. Більше того, якщо людина з дефектами мовлення зможе використати подібні інтерфейси та через неправильно розпізнану команду система виконає іншу команду, наслідки можуть бути досить різними, аж до повного форматування пристрою з важливою інформацією.

Дефекти мовлення мають різноманітні проявлення та причини виникнення. Виділяють такі основні види порушень мовлення, як заїкання, шепелявість, нездатність правильно вимовляти деякі звуки, повна або часткова втрата можливості розмовляти, порушення темпу мовлення, тобто його прискорення або уповільнення [2].

Виходячи з цієї інформації, можна сказати, що при повній або майже повній втраті мовленнєвої діяльності користування мовним інтерфейсом неможливо повністю, адже без вимови слів розпізнавання є неможливим. При всіх інших порушеннях мовлення подібні інтерфейси використовувати можливо, але для кожного дефекту потрібно використовувати свій власний метод визначення слова та сприймання системою його особливостей. Системи, яка могла б розпізнавати мовлення з усіма цими дефектами на даний момент не існує

Серед усіх порушень мовленнєвої діяльності заїкання є найбільш поширеним та досить простим до визначення на аудіо сигналі. Заїкання – це порушення мовленнєвої діяльності, яке характеризується частим повторенням або пролонгацією звуків або складів, або слів; або частими зупинками або нерішучістю в промові, що розриває її ритмічність [2].

Приймаючи до уваги всю вищесказану інформацію, можна створити алгоритм обробки аудіо сигналу для дефекту заїкання:

1. Отримання аудіо сигналу та попередня його обробка, тобто зменшення рівню перешкод та видалення ділянок сигналу без звуку.
2. Перетворення аудіо сигналу у спектральну характеристику, залежну від часу.

3. Виділення на всій довжині спектрограми ділянок певного розміру, зі зменшенням значення довжини ділянки, буде зменшуватися чутливість до змін на спектрограмі, що буде значно при розпізнаванні команди.

4. Порівняння деякої кількості ділянок, що розташовані на відстані менше, ніж порогове значення відстані. Якщо такі ділянки були знайдені та їх кількість перевищує порогове значення, що обчислюється за допомогою розміру однієї ділянки та загальної кількості ділянок на аудіо сигналі, тоді дефект заїкання було знайдено.

5. Видалення ділянок, що повторюються задану кількість разів на заданій відстані від початкових ділянок. Наприклад, у команді «перезавантаження» буква е не буде видалена у префіксі для другого випадку зустрічі, бо між ними знаходиться буква р, аудіо сигнал якої не схожий на сигнал букви е.

Отриманий сигнал можна використовувати для визначення команди, використовуючи, наприклад, дослідження, проведене у [3]. Можливе спотворення частини аудіо сигналу через видалення декількох ділянок, що не є дефектом мовлення, не вплине на загальну правильність розпізнавання, завдяки порівнянню коренів команд, що є досить великими за розміром сигналу. Команда являється розпізнаною лише тоді, коли більша частина кореню її буде схожа на еталонну.

Використовуючи записаний вище алгоритм обробки сигналу при дефекті заїканні, було створено систему визначення мовних команд. Перевірка роботи системи проводилася для двадцяти команд, які є найбільш поширеними при використанні СОМ. Промовляння кожної команди відбувалося з різним ступенем спотвореності дефектом заїкання та відбувалося три рази. Таким чином, загальна кількість спроб визначення склала шістьдесят. Загальна правильність розпізнавання команд з дефектом склала 73,66%, при цьому більшість випадків неправильного розпізнавання відводилося до певних слів, звичайно схожих на інші команди. Наприклад, команда «видалити» через спотворення дефектом та подальше видалення цього дефекту була неправильно розпізнана у більшості випадків, адже вона дуже схожа на команду «виділити». Покращення правильності розпізнавання команд можна досягнути за рахунок використання таких слів, що не повторюють склад іншої команди, тобто не мають спільного кореня.

**Висновки.** У дослідженні була розглянута проблема людей з дефектами мовлення при роботі зі звуковими інтерфейсами людино-машинної взаємодії. Були розглянуті найвідоміші сервіси розпізнавання мови та визначена неможливість цих сервісів визначити команди інтерфейсу з необхідною правильністю. Також були розглянуті різноманітні порушення мовлення, а потім визначені особливості розпізнавання команд з дефектом заїкання. Було розроблено алгоритм обробки аудіо сигналу за допомогою визначених особливостей мовлення з заїканням. За розробленим алгоритмом було створено програмну реалізацію системи визначення мовленнєвих команд. Правильність розпізнавання команд з дефектом заїкання складає 73,66% від загальної кількості усіх спроб для вибірки з двадцяти команд, кожна з яких було промовлено по три рази з різною спотвореністю дефектом. Таким чином, створена система потребує подальшого покращення для використання її в інтерфейсах людино-машинної взаємодії, адже правильність розпізнавання не відповідає потребам стійкості до помилок. Подібні системи потребують високої правильності розпізнавання команд, не менше ніж 95% з необхідністю підтвердження тих, що можуть завдати найбільшої шкоди даним користувача.

### СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. SCIENTIFIC AMERICAN [Електронний ресурс]. Режим доступу: URL: <https://www.scientificamerican.com/article/why-siri-won-t-listen-to-millions-of-people-with-disabilities/> - Why Siri Won't Listen to Millions of People with Disabilities

2. Науковий часопис. Актуальні проблеми логопедії [Електронний ресурс]. Режим доступу: URL: <http://enpuir.npu.edu.ua/bitstream/123456789/7821/1/Boryak.pdf> — До питання педагогічної роботи з дітьми, Боряк О.В.

3. Друга міжнародна науково-практична конференція «Project, Program, Portfolio Management, 2017», «Інтерфейс людино-машинної взаємодії за допомогою мовних команд», Мариморич В.С., с. 64 – 66.