

## **РЕЦЕНЗІЯ**

на дисертаційну роботу

**Світловського Євгенія Володимировича**

на тему «Оброблення, запис та передавання засобами IoT мовного

**аудіосигналу з дефектами»,**

представлену на здобуття ступеня **доктора філософії**

в галузі знань 17 «Електроніка та телекомунікації»

за спеціальністю 171 «Електроніка»

### **Актуальність теми дисертації.**

Системи Інтернету речей (IoT) широко використовуються в багатьох сферах життя – промисловій автоматизації, охоронних системах, "розумних" будинках, медицині, освіті. В цій концепції саме аудіосигнали досить часто використовуються для забезпечення взаємодії між людиною та "розумними" пристроями через голосові команди, розпізнавання подій або моніторингу середовища.

Окремої уваги заслуговують мовні аудіосигнали, для яких критично важливим є високий рівень розбірливості. Якщо мовні аудіосигнали містять дефекти (зашумлення, втрати, неповні аудіоданні), це значно ускладнює їх розпізнавання, знижує якість комунікації та може призводити до втрати важливої інформації.

Українська мова характеризується унікальними лінгвістичними властивостями – складною фонетичною структурою, характерною інтонацією та великою кількістю шиплячих та свистячих приголосних, що потребує спеціалізованих підходів до цифрової обробки. Наявні рішення орієнтовані переважно на англomовний контент, створюючи прогалину для україномовних застосувань.

Розроблення алгоритмів, здатних ефективно мінімізувати шумові перешкоди, підвищувати артикуляційну чіткість та забезпечувати стеганографічну передачу даних без втрати якості, є актуальним завданням. Це дозволить покращити розбірливість мовлення, надійність передачі голосової інформації та забезпечити новий рівень інтерактивності в системах людино-машинної взаємодії.

### **Оцінка обґрунтованості наукових результатів дисертації, їх достовірності та новизни.**

Сформульовані в дисертаційній роботі наукові положення, висновки та рекомендації достатньо повно обґрунтовані здобувачем та викладені в доказовій формі.

Достовірність результатів, представлених в роботі, забезпечується коректністю постановки задач дисертаційного дослідження, ґрунтовним та ретельним вивченням існуючих літературних джерел, використанням в роботі широко апробованих методів та методик обробки аудіосигналів, коректним застосуванням сучасного програмного забезпечення.

Наукова новизна результатів дисертаційного дослідження полягає в наступному:

1. Вперше розроблено та обґрунтовано алгоритм обробки аудіофайлів українською мовою в умовах акустичних завад, який структуровано на окремі етапи та адаптовано до використання в середовищі IoT.
2. Вдосконалено підхід до обробки україномовного аудіосигналу шляхом уточнення алгоритму на основі спектрального аналізу із врахуванням особливостей визначення фундаментальної частоти.
3. Вперше запропоновано алгоритм подвійної обробки аудіосигналу з мовленнєвим контентом українською мовою, який забезпечує можливість реалізації методу приховування інформації в структурі аудіофайлу із збереженням якісних характеристик сигналу та без суттєвих змін його енергетичних параметрів.

Наукові дослідження були виконані здобувачем на кафедрі акустичних та мультимедійних електронних систем КПІ ім. Ігоря Сікорського в рамках ініціативної теми "Автоматизована система суб'єктивного та об'єктивного оцінювання якості мовлення" (№ держреєстрації 0116U008959).

Поставлене в дисертаційній роботі наукове завдання, що полягало в розробленні алгоритму подвійного оброблення україномовного аудіосигналу з дефектами, що дозволяє приховувати потрібну інформацію в самій структурі аудіофайлу з одночасним збереженням якості та без суттєвих змін енергетичних параметрів, виконано повністю, здобувач повною мірою оволодів методологією наукової діяльності.

**Оцінка змісту дисертації, її завершеність та дотримання принципів академічної доброчесності.**

За своїм змістом дисертаційна робота здобувача Світловського Євгенія Володимировича повністю відповідає Стандарту вищої освіти зі спеціальності 171 «Електроніка» та напрямкам досліджень відповідно до освітньої програми «Електроніка».

Дисертація складається з вступу, шести розділів, висновків, переліку літературних джерел та додатку. Загальний обсяг дисертації становить 145 сторінок.



У вступі зазначено актуальність дисертаційної роботи, сформульовано мету та перераховано завдання, що сприяють її досягненню, описані обрані методи дослідження, наведено інформацію про наукову новизну та практичне значення отриманих результатів. Представлений перелік публікацій, в яких висвітлено результати дисертаційного дослідження та подана інформація про апробацію на наукових конференціях, вказано особистий внесок дисертанта.

В першому розділі визначено недоліки існуючих методів обробки засобами IoT мовних аудіосигналів в умовах зашумлення. Розглянуто основні чинники, які слід враховувати при проведенні запису мовної інформації. Окреслено фактори, які сприяють покращенню розбірливості мовлення та зниженню впливу фізичних і електронних шумів. Приділено увагу питанням вибору звукової карти.

Другий розділ присвячений особливостям проведення частотного аналізу мовного сигналу. Зокрема, розглянуті особливості розбиття сигналу на сегменти, ідентифікації максимумів, оцінки шумів, а також аналізу формант та фундаментальної частоти.

В третьому розділі дисертації проаналізовано основні етапи розпізнавання мови в сучасних системах автоматичного розпізнавання, описано особливості застосування систем автоматичного розпізнавання мови та визначено їх ключові переваги. Наведено основні принципи текстового кодування за стандартами UTF-8 та ASCII, обґрунтовано використання формату UTF-8 для забезпечення сумісності з сучасними системами та можливості передачі україномовних текстових фрагментів. Розглянуто можливості передачі текстової інформації спільно з мовленнєвими сигналами за допомогою стеганографічного методу найменш значущого біта.

В четвертому розділі було проведено комплексний аналіз апаратного забезпечення, необхідного для запису мовних аудіосигналів з дефектами з врахуванням вимог до якості обробки та передачі сигналів у системах IoT. Зокрема, обґрунтовано вибір аудіоінтерфейсу Scarlett Solo 2nd Gen, розглянуто доцільність використання для запису конденсаторного мікрофона Rode NT2-A. Наведено технічні характеристики обладнання для роботи зі звуком у середовищі IoT, описано канали зв'язку, схему підключення та інтеграцію алгоритмів обробки мовного сигналу до системи.

П'ятий розділ присвячено опису експеримента з підвищення якості та зниження рівня шумового забруднення записаного україномовного сигналу з наявними технічними дефектами. Зокрема, розроблено програмний алгоритм з елементами циклічності, де визначені окремі послідовні етапи обробки сигналу з урахуванням фундаментальної частоти, динамічних та частотних характеристик, рівня шумового забруднення. На основі отриманих



результатів сформульовано підходи з оброблення аудіосигналів, які адаптовано для роботи з українською фонетичною групою.

У шостому розділі представлено алгоритм ідентифікації та кодування текстових даних для інтеграції супутньої прихованої інформації в аудіофайл. Використовуючи відкриту бібліотеку розпізнавання мовлення, здійснено екстракцію текстових даних із записаного сигналу з подальшою корекцією та приведенням до необхідного формату. Розроблено процедуру вбудовування обробленої текстової інформації в аудіосигнал на основі стеганографічного методу найменш значущого біта. Експериментально підтверджено, що модифікований аудіосигнал зберігає характеристики, практично ідентичні до вихідного сигналу, що забезпечує непомітність прихованої інформації та збереження якості звучання.

У висновках сформульовано основні результати дисертаційної роботи.

Дисертаційна робота є завершеною науковою працею і свідчить про наявність особистого внеску здобувача в розроблення електронних засобів оброблення аудіо інформації для пристроїв IoT.

Розглянувши звіт подібності за результатами перевірки дисертаційної роботи на текстові співпадіння, можна зробити висновок, що дисертаційна робота Світловського Євгенія Володимировича є результатом самостійних досліджень здобувача і не містить елементів фальсифікації, компіляції, фабрикації, плагіату та запозичень. Використані ідеї, результати і тексти інших авторів мають належні посилання на відповідні джерела.

### **Мова та стиль викладення результатів.**

Дисертаційна робота написана українською мовою переважно з дотриманням загальноприйнятої наукової термінології в галузі електроніки та прикладної акустики. Матеріал подається послідовно та логічно, що забезпечує доступність викладення основних положень та цілісне розуміння проведених досліджень. Стиль мовлення є аргументованим та науково-обґрунтованим.

Дисертаційна робота оформлена відповідно до вимог наказу МОН України від 12 січня 2017 р. № 40 «Про затвердження вимог до оформлення дисертації».

### **Оприлюднення результатів дисертаційної роботи.**

Наукові результати дисертації висвітлені у 4 наукових публікаціях здобувача, серед яких: 4 статті у наукових виданнях, включених на дату опублікування до переліку наукових фахових видань України. Результати дисертаційного дослідження обговорювались на 2 наукових фахових конференціях.



Таким чином, наукові результати, представлені в дисертаційній роботі, повністю висвітлені у наукових публікаціях здобувача.

### **Недоліки та зауваження до дисертаційної роботи.**

1. В п. 1.5 вказано, що “Оптимальний час реверберації для мовних програм визначається за формулою”, але саме співвідношення не наведено.

2. Вважаю, що в третьому розділі варто було б зробити більш широкий огляд програмних рішень для автоматичного розпізнавання мови (п.3.1) та форматів кодування (п.3.2, 3.3), провести їх порівняльний аналіз, вказати переваги і недоліки, а не обмежуватись одним найбільш поширеним варіантом.

3. В висновках до четвертого розділу два перших пункта.

4. Рисунок 5.3 підписаний “АЧХ та сигналограма до та після застосування обробки”, хоча насправді являє собою часову діаграму (графік сигналу) та спектрограму.

Загалом, в роботі формулювання “Часово амплітудна характеристика сигналу” або “Амплітудно-часова характеристика сигналу” зустрічається досить часто в сенсі часової діаграми або графіка сигналу.

5. В таблиці 5.1 не вказані одиниці вимірювання середньої амплітуди.

6. В шостому розділі роботи у частині, де описується розпізнавання мови з аудіофайлу доцільно було б навести особливості використання відкритої бібліотеки Vosk.

7. Рисунок 6.1 та, власне, опис методу LSB (с.108-111) логічніше було б розмістити в теоретичній частині дисертації, а не в шостому розділі, присвяченому практичній реалізації розробленого алгоритму.

Вважаю, що висловлені зауваження не є визначальними і не зменшують загальну наукову новизну та практичну значимість результатів та не впливають на позитивну оцінку дисертаційної роботи.

### **Висновок про дисертаційну роботу.**

Вважаю, що дисертаційна робота здобувача ступеня доктора філософії Світловського Євгенія Володимировича на тему «Оброблення, запис та передавання засобами IoT мовного аудіосигналу з дефектами» виконана на високому науковому рівні, не порушує принципів академічної доброчесності та є закінченим науковим дослідженням, сукупність теоретичних та практичних результатів якого розв'язує наукове завдання, що має істотне значення для електроніки та телекомунікації. Дисертаційна робота за актуальністю, практичною цінністю та науковою новизною повністю відповідає вимогам чинного законодавства України, що передбачені в п. 6 – 9 «Порядку присудження ступеня доктора філософії та

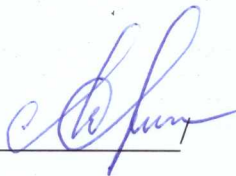
скасування рішення разової спеціалізованої вченої ради закладу вищої освіти, наукової установи про присудження ступеня доктора філософії», затвердженого Постановою Кабінету Міністрів України від 12 січня 2022 р. № 44.

Здобувач Світловський Євгеній Володимирович заслуговує на присудження ступеня доктора філософії в галузі знань 17 – Електроніка та телекомунікації за спеціальністю 171 – Електроніка.

**Рецензент:**

Доцент кафедри акустичних та  
мультимедійних електронних систем  
КПІ ім. Ігоря Сікорського

к.т.н., доцент

/ 

Катерина ДРОЗДЕНКО



М.П. Відділ кадрів  
і архівної  
справи  
№ 6

« 03 » 06 20 25 року

