

## **ВІДГУК**

офіційного опонента на дисертаційну роботу  
Кухарічевої Катерини Андріївни  
на тему «Підвищення робастності систем автоматичного  
розпізнавання мови до дії завад»,  
представлену на здобуття наукового ступеня доктора філософії  
в галузі знань 17 "Електроніка та телекомунікації"  
за спеціальністю 171 "Електроніка"

### **Актуальність теми дисертації.**

Актуальність обраної теми зумовлюється інтенсивним розвитком систем автоматичного розпізнавання мовлення (САРМ), як різновиду систем штучного інтелекту. Системи автоматичного розпізнавання мовлення дозволяють ефективно вирішувати проблеми управління роботами, пришвидшувати пошук інформації, підвищувати продуктивність роботи стенографістів та журналістів. Широке застосування САРМ знайшло в людино-машинних інтерфейсах, інформаційно-пошукових системах, в комунікаційних системах. Інтерактивні платформи голосового обміну інформації включають САРМ і мають все більший попит застосування в сучасному світі.

В автоматичному розпізнаванні мовлення, основною процедурою, що визначає його ефективність, є чутливість до спотворень мовленнєвого сигналу, спричинених завадами, такими як обмеження амплітуди сигналу (кліпування), наявність шуму оточення та реверберація у приміщеннях. Вказані завади призводять до зміни форми сигналу та, як наслідок, суттєвого зниження точності розпізнавання мовлення аж до фактичної втрати достовірності результатів САРМ. Для зменшення впливу завад спотворені сигнали перед подаванням їх на вхід САРМ частково відновлюються. Альтернативним напрямом боротьби із завадами є підвищення стійкості (робастності) самих САРМ до їх дії. Іншим найбільш перспективним шляхом підвищення робастності САРМ є їх налаштування шляхом навчання на сигналах, спотворених завадами. Однак, великий перелік видів та параметрів завад, а також бажаність врахування при навчанні апріорної інформації про умови експлуатації САРМ потребують подальших досліджень методів оцінювання якості та розбірливості мовленнєвих сигналів, розробки нових та

вдосконалення відомих методів навчання систем автоматичного розпізнавання мовлення.

Отже, тема дисертаційної роботи Кухарічевої Катерини Андріївни, що присвячена розробці нових та вдосконалення існуючих методів підвищення точності САРМ в умовах негативної дії завад різної природи і, й разом із тим, відносно простих для практичної реалізації в системах автоматичного розпізнавання мовлення, є актуальною науково-технічною задачею.

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

Наукова новизна результатів дисертаційного дослідження полягає в наступному:

- вперше для реальних мовленнєвих сигналів отримано кількісні оцінки ступеню підвищення точності розпізнавання мовлення, спотвореного шумом різної природи та інтенсивності, шляхом навчання системи автоматичного розпізнавання на спотворених шумом сигналах;
- вперше для реальних мовленнєвих сигналів отримано кількісні оцінки ступеню підвищення точності розпізнавання мовлення, спотвореного реверберацією, шляхом навчання системи автоматичного розпізнавання на спотворених реверберацією сигналах;
- вдосконалено метод оцінювання розбірливості мовлення непрямим методом, із використанням міри якості сигналів у вигляді барківського спектрального спотворення;
- уточнено висновки щодо залежності розбірливості мовлення від щільності відбить звуку та часу реверберації, із використанням імовірнісних моделей імпульсних характеристик приміщень;
- вдосконалено спосіб виявлення ефекту кліпування мовленнєвих сигналів та об'єктивного оцінювання якості мовленнєвих сигналів, спотворених кліпуванням, що базується на використанні коефіцієнта ексцесу як міри спотворення сигналів.

При вирішенні поставлених у дисертаційній роботі задач, створені наукових положень, висновків та рекомендацій здобувачем застосовані дані, які одержані з літературних джерел, з результатів аналізу сучасного стану та перспектив розвитку методів підвищення точності САРМ в умовах негативної дії завад різної природи, тому їх можна вважати обґрунтованими.



Достовірність обраних методів розв'язку підтверджується збігом теоретичних результатів із результатами експериментальних досліджень, коректним застосуванням математичного апарату та сучасних методів наукового моделювання. Це свідчить про науковий внесок здобувача у галузь знань Електроніка та телекомунікації.

Отже, в дисертаційній роботі поставлене наукове завдання 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 Кухарічевої К.А. повністю відповідає Стандарту вищої освіти зі спеціальності 171 Електроніка та напрямкам досліджень відповідно до освітньої програми 171 Електроніка.

Дисертаційна робота є завершеною науковою працею і свідчить про наявність особистого внеску здобувача у науковий напрям Електроніка та телекомунікації.

Розглянувши звіт подібності за результатами перевірки дисертаційної роботи на текстові співпадиння, можна зробити висновок, що дисертаційна робота Кухарічевої Катерини Андріївни 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 відповідне джерело.

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

Дисертаційна робота написана українською мовою.

Дисертаційна робота написана загальноприйнятною науковою мовою із використанням сучасної наукової термінології. Зміст дисертації є достатньо виваженим та цілісним, робота відповідає темі досліджень, дисертаційна робота є завершеною науковою працею.

Дисертація складається зі вступу, чотирьох розділів, висновків, списку використаних джерел із 106 найменувань та 5 додатків, містить 55 рисунків та 33 таблиці. Загальний обсяг дисертаційної роботи становить 148 сторінок.

У вступі обґрунтовано актуальність дисертаційної роботи, сформульовано мету та перераховано задачі дослідження, описано методи дослідження, надано інформацію про наукову новизну та практичне значення

одержаних результатів. Надано інформацію про висвітлення результатів роботи в періодичних наукових виданнях та про апробацію на наукових конференціях та семінарах. Показано зв'язок отриманих результатів з планами застосування у освітньому процесі, а також особистий внесок дисертанта.

У *першому розділі* виконано огляд існуючих підходів до підвищення точності САРМ в умовах дії завад. При цьому надано коротку характеристику двом напрямам підвищення точності розпізнавання: попередня корекція сигналу та адаптація системи до дії завад. Слід зазначити, що в існуючих наукових працях недостатньо висвітлено другий напрям. Оскільки поглиблений аналіз зазначених питань є утрудненим без конкретизації архітектури САРМ, в першому розділі порівняно між собою три різновиди такої архітектури, що базуються на використанні прихованих марковських моделей (ПММ), глибинних нейронних мереж, а також гібридних систем. Для подальших досліджень дисертантом обрано моделі на основі ПММ, оскільки вони краще піддаються аналізу проміжних результатів, а отримані на моделях ПММ рекомендації можуть бути розповсюджені й на інші моделі САРМ.

У *другому розділі* наведено результати дослідження окремих видів лінійних та нелінійних спотворень мовленнєвого сигналу, що можуть істотно впливати на ефективність роботи САРМ, а саме, реверберації та кліпування сигналу. При аналізі характеру та ступеня впливу кліпування на якість (форму) мовленнєвого сигналу, запропоновано міри ступеня спотворення форми сигналу, а саме, коефіцієнт ексцесу та функціонально пов'язані із ним величини. Виконано порівняння запропонованого об'єктивного (інструментального) підходу із конкурентним об'єктивним методом, що базується на евристичному аналізі оцінки закону розподілу сигналу. Показано перевагу запропонованого методу, яка полягає в можливості не тільки виявляти факт кліпування, але й оцінювати його ступінь. Це важливо, оскільки дає змогу більш ретельно, з врахуванням суб'єктивних оцінок якості кліпованих сигналів, регулювати коефіцієнт підсилення тракту передачі інформації.

Що стосується аналізу впливу реверберації на якість мовленнєвих сигналів та на розбірливість мовлення, виконано великий об'єм експериментальних досліджень в аудиторіях малого, середнього та великого розміру. При цьому продемонстровано зростання коефіцієнту кореляції між об'єктивними мірами якості та розбірливості мовлення при зростанні об'єму



приміщення. Отримані результати є практично корисними, оскільки за певних умов дозволяють замість громіздких методів оцінювання розбірливості мовлення використовувати простіші методи оцінювання якості сигналів. Окрім того, представлено результати модельних досліджень впливу раннього багатократного відбиття звуку в приміщенні на розбірливість мовлення. При цьому вказано, за яких умов ранні відбиття можуть бути корисними або шкідливими, з позицій розбірливості мовлення.

У *третьому розділі* представлено короткий огляд мір якості систем автоматичного розпізнавання мовлення й, зокрема, мір якості, що використовуються при оцінюванні точності розпізнавання в програмному комплексі The Hidden Markov Model Toolkit (НТК), який суттєво використовувався дисертантом при експериментальних дослідженнях. Більшу частину розділу присвячено дослідженням можливості підвищення робастності САРМ до дії шумової завади шляхом навчання САРМ на зашумлених сигналах. В дисертації розглянуто чотири комбінації характеристик навчальної та тестової вибірок. В першому випадку САРМ навчається лише на сигналах, що маскуються шумами із певним спектром та певним відношенням сигнал-шум. Показано, що цей окремий випадок є зручним лише за умови, що апіорі відомі шумові умови майбутньої експлуатації САРМ. В другому випадку навчання відбувається на сигналах, що зашумлені завадою із певним спектром, але із різними відношеннями сигнал-шум. В третьому випадку фіксовано відношення сигнал-шум, але варіюються види шумів. В четвертому випадку навчання відбувається для різних видів шумів та в широкому діапазоні можливих значень сигнал-шум. Цікавим з наукової точки зору та корисним для практичних застосувань є той факт, що в останньому випадку точність розпізнавання вдається підтримувати на високому рівні лише починаючи з досить великих відношень сигнал-шум (більше 10 дБ).

Слід відзначити великий обсяг проведених дисертантом експериментальних досліджень. Загальна кількість записів мовлення, створених для досліджень в рамках третього розділу дисертаційної роботи, становила майже 30 тисяч, при цьому використано 14 видів шумів навколишнього середовища.

У *четвертому розділі* дисертаційної роботи наведено результати експериментальних досліджень, спрямованих на підвищення робастності

системи АРМ в умовах переважної дії ревербераційної завади. При цьому визначалась ефективність роботи системи за різних варіантів навчання та тестування для сигналів, спотворених ревербераційною завадою. Імпульсні характеристики, використані в дослідженнях, залучено із відкритої бібліотеки бінауральних імпульсних характеристик бази даних Рейнсько-Вестфальського технічного університету міста Аахен. Загалом використано 62 імпульсних характеристики приміщень із широким діапазоном значень часу реверберації. Час реверберації та тип приміщення обрано таким чином, щоб охопити приміщення, в яких пристрої, оснащені системою АРМ, потенційно будуть використовуватися найчастіше: це навчальна аудиторія, офісне приміщення, кімната для нарад, сходовий майданчик. Усього для експериментальних досліджень в рамках даного розділу було створено і оброблено 13020 звукових файлів.

В дисертації розглянуто чотири випадки навчання САРМ. В першому випадку навчання відбувалося на чистих сигналах, а тестування відбувалося на всіх спотворених реверберацією сигналах. В другому випадку навчання та тестування виконувалися на сигналах, спотворених імпульсними характеристиками з однаковим часом реверберації. В третьому випадку навчання виконувалося на сигналах, спотворених різними імпульсними характеристиками з різним часом реверберації. В четвертому випадку САРМ навчалася на реверберованих сигналах в різних точках одного приміщення. Вказані випадки враховують типові умови експлуатації САРМ в приміщеннях різного призначення.

Анотація дисертації відображає її основні положення. У висновках сформульовано основні результати дисертаційної роботи.

У 5 додатках представлено результати дослідження точності розпізнавання САРМ при методах навчання з використанням зашумлених сигналів, в умовах дії ревербераційної завади та перелік імпульсних характеристик приміщень та часу реверберації T20.

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 достатньо повно висвітлені у 9 наукових публікаціях здобувача. Основні результати дисертації висвітлено у 5 наукових публікаціях здобувача, серед яких: 3 статті у наукових виданнях, включених на дату опублікування до переліку наукових фахових видань України; 1 стаття у періодичних наукових виданнях, проіндексованих у базах даних Scopus, з яких 1 стаття у виданнях, віднесених до третього квартилю (Q3) відповідно до класифікації SCImago Journal and Country Rank або Journal Citation Reports; 1 патент України на корисну модель.

Також результати дисертації були апробовані на 4 наукових фахових конференціях.

Публікація результатів дисертаційних досліджень у рецензованих виданнях, які передбачають попередню перевірку на відсутність запозичень, є одним з елементів підтвердження академічної доброчесності. В цілому у дисертаційній роботі порушень академічної доброчесності не виявлено.

Науковий рівень публікацій здобувача відповідає вимогам, що їх затверджено для подання дисертаційної роботи до захисту. Вимоги Постанови КМУ від 06.03.2019 № 167 "Про проведення експерименту з присудження ступеня доктора філософії", зі змінами згідно Постанови КМУ від 09.06.2021 № 608 до кількості публікацій виконано.

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

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

Загалом дисертаційна робота Кухарічевої Катерини Андріївни за своїм змістом відповідає зазначеній спеціальності 171 «Електроніка», однак стосовно даної роботи можна зробити певні зауваження, а саме:

1. У висновках до Розділу 1 перераховано, що зроблено в даному розділі, але не вказано прогалини в попередніх дослідженнях, наявність яких робить дану дисертаційну роботу актуальною.
2. У підрозділі 1.3 огляду мір оцінювання якості та розбірливості мовлення варто було б включити опис мір якості систем автоматичного розпізнавання мови.

3. У Розділі 2 спочатку представлено результати досліджень впливу нелінійних спотворень (кліпування) на розбірливість мовлення, й лише після цього - впливу лінійних спотворень у вигляді реверберації. Було б більш логічним спочатку надати розгляд лінійних спотворень як більш поширених, оскільки кліпування має місце лише при використанні апаратури звукопідсилення.

4. На початку підрозділу 3.3 наведено тільки англомовні назви способів навчання CAPM на зашумлених сигналах (навчання CAPM за методами Fully-Matched Training (FMT), Noise-Matched Training (NMT, інша назва Spectrum Matched Training - SMT), Signal-to-Noise Ratio Matched Training (SNRMT) та Multistyle Training (MT)), проте не надано посилання на літературні джерела цих способів навчання.

5. У підрозділах 3.3–3.8 для проведення експериментальних досліджень використано 14 шумів навколишнього оточення, які варто було б узагальнити або класифікувати.

6. На початку підрозділу 4.1 наведено тільки англомовні назви (Clean training, Reverb-matched training, All-reverb training, Room training) методів навчання CAPM на реверберованих сигналах, проте незрозуміло, хто є автором цих назв.

7. У розділах 3 і 4 для оцінювання точності розпізнавання використовується програмний комплекс The Hidden Markov Model Toolkit (НТК), однак не наведено опису та загальної структури цього інструментального засобу.

8. В кінцевій частині Висновків до дисертації наведено відомості про практичне значення роботи, які в значній мірі є повторенням тексту, наданого у Вступі до дисертації - було б доречним цей розділ Висновків надати у скороченій формі.

9. Зауваження до оформлення: в тексті дисертаційної роботи наявні граматичні помилки та лінгвістичні неточності, зустрічаються англомовні терміни без перекладу.

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 Кухарічевої Катерини Андріївни на тему «Підвищення робастності систем автоматичного розпізнавання мови до дії завад» 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 галузі знань 17 Електроніка 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 Кухарічева Катерина Андріївна заслуговує на присудження ступеня доктора філософії в галузі знань 17 Електроніка та телекомунікації за спеціальністю 171 Електроніка.

#### Офіційний опонент:

доктор технічних наук, професор  
кафедри автоматизованих систем управління  
Національного університету  
«Львівська політехніка»

 Ігор ПРОЦЬКО

Підпис д.т.н., професора Процька І.О.

«ЗАСВІДЧУЮ»

Вчений секретар Національного Університету  
«Львівська Політехніка»

«15» листопада 2023 р.



 Роман БРИЛИНСЬКИЙ

М.П.

«\_15\_» листопада 2023\_ року