

**УДК 004.52:004.8**

**Глазок О. М.**

к. т. н., доцент,

доцент кафедри комп'ютеризованих систем управління,

Національний авіаційний університет,

**Бойко В.М.**

студент Факультету кібербезпеки, комп'ютерної та програмної інженерії,

Національний авіаційний університет

### **Проблеми застосування голосових помічників у сфері торгівлі**

Розвиток технологій, що роблять можливим розпізнавання людської мови, привів до появи ряду практичних застосувань, таких як ідентифікація особи за звуком голосу та голосові помічники. Для двостороннього спілкування голосовий помічник повинен бути в змозі проаналізувати почуте і винести судження про сенс сказаного людиною. У найпростіших випадках таке судження зводиться до задачі розпізнавання образів з класифікацією, тобто віднесення висловлювання людини до одного з класів (наперед визначених простих команд, варіантів відповіді на питання, задані голосовим помічником, тощо). Більш складний підхід передбачає можливість визначення сенсу голосових повідомлень довільної форми.

Програмне забезпечення, що використовується для розпізнавання тих чи інших аспектів голосу, використовує технології згорткових нейронних мереж та нейромереж з довгою короткочасною пам'яттю [1, 2].

На даний час голосові помічники впроваджуються у багатьох сферах, таких як керування окремими побутовими приладами, керування застосунками «розумного дому», керування комп'ютером, автомобілем [3], пошук інформації, інтерфейси різноманітних онлайн-сервісів, відправка повідомлень, і навіть вивчення іноземних мов. Незважаючи на зручність та інші очевидні переваги використання голосових помічників для бізнесу і для користувачів,

дослідження показують, що використання голосових помічників у процесі спілкування клієнта з магазином та придбання товарів чи послуг займає одне з останніх місць серед усіх застосувань голосових помічників, і ця ситуація від року у рік не покращується [4]. Небажання споживачів використовувати технологію голосових помічників у процесі добору товарів і здійснення покупок викликане цілим рядом проблем.

Перша з таких проблем – недостатні системні характеристики споживацьких пристроїв (передусім – продуктивність). Алгоритми штучного інтелекту для розпізнавання мови та використовувані ними бази даних мають значний обсяг. Для роботи у реальному часі вони вимагають наявності потужних обчислювальних пристроїв і пристроїв зберігання з великим обсягом і пропускнуою здатністю. Пристрій споживача, такий як смартфон чи ноутбук, не може забезпечити виконання таких вимог, тому програмне забезпечення, що якісно розпізнає голосові повідомлення, має виконуватись у хмарному середовищі, а користувачу для зв'язку з ним необхідно мати швидкісний доступ до Інтернету. Звідси виникають побоювання, пов'язані з проблемами конфіденційності. Голосові помічники повинні постійно слухати звуки, що звучать у приміщенні, очікуючи на команди, надані користувачем. Інформація про почуте передається в інтернет, таким чином голосовий помічник фактично веде постійний запис приватних розмов у повсякденному житті користувача, пізніше отримані таким чином дані можуть бути використані проти користувача.

Ще одне побоювання безпеки пов'язане із можливою вразливістю технології розпізнавання голосу до кібератак. Голосові помічники можуть розпізнавати голосові команди, нечутні для людського вуха. Зловмисники можуть використовувати цю можливість для доступу до приватної інформації користувача або для виконання покупок від його імені. При цьому можуть бути використані записи оригінального голосу користувача або технології симуляції голосу, що унеможливить виявлення підробки чи суттєво знизить її імовірність.

Незважаючи на значну використовувану обчислювальну потужність, якість розпізнавання мови на даний час не ідеальна. Голосові помічники можуть неправильно інтерпретувати певні слова, не розуміти аббревіатури, сленгові слова, специфічний акцент у вимові, тощо. Це може призвести до помилок у роботі системи. Якість розпізнавання голосу також може суттєво падати через наявність фонового шуму, в той час як користувач не завжди може забезпечити відсутність шумів вдома чи на роботі.

Висновки. Актуальними задачами в галузі створення голосових помічників є проведення експериментів з розпізнаванням голосу з урахуванням факторів, які можуть впливати на результати розпізнавання, тобто дослідження моделей розпізнавання мовлення, їх можливих застосувань, та у зв'язку з цим дослідження відповідних архітектур нейронних мереж. В той же час необхідно працювати над вирішенням проблем якості розпізнавання та безпеки інформації, що перешкоджають ширшому впровадженню даної технології у таких сферах, як торгівля.

### **Список літератури**

1. C. H. Taal, R. C. Hendriks, R. Heusdens and J. Jensen. A short-time objective intelligibility measure for time-frequency weighted noisy speech //Proceedings of 2010 IEEE International Conference on Acoustics, Speech and Signal Processing. – 2010. – Pp. 4214-4217. DOI: 10.1109/ICASSP.2010.5495701.

2. A. Torfi, S. M. Iranmanesh, N. Nasrabadi and J. Dawson, 3D Convolutional Neural Networks for Cross Audio-Visual Matching Recognition. //IEEE Access, 2017, vol. 5, pp. 22081-22091. DOI: 10.1109/ACCESS.2017.2761539.

3. M. Jebbar, A. Maizate, R. A. Abdelouahid. Simulation of Car Driving by Voice Commands based on a Deep-Learning Model //Procedia Computer Science 203(7553): August 2022. – Pp. 813-818. DOI: 10.1016/j.procs.2022.07.122.

4 .A. P. K. Muthukumaran, H. Vani. Optimizing the usage of voice assistants for shopping //Indian Journal of Science and Technology. – 2020, No. 13(43). – Pp. 4407-4416. DOI: 10.17485/IJST/v13i43.1911.