УДК 621.396:004.056

ТРАНСКРІБАЦІЯ НА МАЛОПОТУЖНИХ ПРОЦЕСОРАХ: АНАЛІЗ ТЕХНОЛОГІЙ ТА ЗАГРОЗ ІНФОРМАЦІЙНІЙ БЕЗПЕЦІ

Жовтобрюх В.Ю.
e-mail: viktoriia.zovtobriukh@nure.ua
Науковий керівник – ст. викл. Ликова Г.О.
Харківський національний університет радіоелектроніки, каф. ІРТЗІ
м. Харків, Україна

The paper examines technologies for automatic audio transcription on low-power processors and analyzes related information security threats. The study reviews microcontrollers and single-board computers used for local speech recognition and considers TinyML-based models suitable for devices with limited resources. Particular attention is paid to optimization methods that enable efficient transcription on constrained hardware. The research also identifies potential risks related to covert audio data collection and transmission. The results highlight the need for improved monitoring and security measures in systems using autonomous edge devices.

Автоматична транскрібація аудіоінформації займає дедалі вагоміше місце в сучасних системах технічного захисту інформації, забезпечуючи автоматизоване перетворення мовного сигналу у текстовий формат із скороченням обсягу даних у 100–200 разів порівняно з оригінальним аудіозаписом.

Метою роботи є дослідження технологій автономної транскрібації на малопотужних процесорах, аналіз загроз інформаційній безпеці, що виникають при їх використанні.

Проведений аналіз сучасних апаратних платформ виявив принципові відмінності між двома категоріями пристроїв. Мікроконтролери – ESP32, STM32F4, RP2040, Nordic nRF52840 – характеризуються обсягом оперативної пам'яті від 32 до 520 КБ, продуктивністю 100–600 DMIPS та середнім енергоспоживанням нижче 0,2 Вт, що забезпечує автономність до 150–200 годин від акумулятора ємністю 10 000 мАг. Одноплатні комп'ютери Raspberry Pi Zero та Zero 2 W займають проміжне положення: наявність 512 МБ оперативної пам'яті та підтримка повноцінної ОС Linux відкривають доступ до широкого спектру Python-бібліотек та фреймворків машинного навчання, проте за рахунок вищого енергоспоживання і значно більшої поверхні атаки.

Огляд технологій TinyML-транскрібації показав, що для мікроконтролерів практично придатні лише компактні моделі розпізнавання ключових слів: Google Speech Commands з розміром 15–50 КБ і точністю 85–95% на словнику до 50 команд, DS-CNN з вимогами до RAM 20–40 КБ, а також платформа Edge Impulse, яка дозволяє автоматично генерувати оптимізовані моделі під конкретні апаратні обмеження. На Raspberry Pi Zero можливе використання Vosk з моделями до 200 МБ, а Zero 2 W з

чотириядерним Cortex-A53 здатен виконувати транскрібацію Whisper tiny майже в реальному часі. Ключовими методами оптимізації моделей є: постквантизація до 8-бітного представлення, structured pruning та knowledge distillation.

Аналіз загроз безпеці виявив три критичні категорії ризиків. По-перше, збір чутливих аудіоданих: сучасні мікроконтролери дозволяють побудувати пристрій прихованого спостереження розміром менше 15×20×5 мм, який при середньому споживанні близько 34 мВт здатен безперервно функціонувати понад 100 годин від компактного акумулятора. Відсутність помітного теплового сліду та електромагнітного випромінювання при вимкнених радіомодулях суттєво ускладнює виявлення таких пристроїв традиційними засобами технічного контролю. По-друге, поєднання автономної транскрібації з технологіями LPWAN формує надзвичайно ефективний канал витоку інформації: аудіофрагмент тривалістю 3 секунди займає ~96 КБ у форматі PCM, тоді як його текстовий транскрипт – лише 10–30 байт, що цілком вкладається в обмеження навіть Sigfox. Дальність передачі до 15–40 км при споживанні 120 мВт протягом 1–2 секунд робить такий канал практично непомітним для корпоративних систем моніторингу та виявлення.

Сформульовані вимоги до радіомоніторингу діапазонів 868–915 МГц для виявлення несанкціонованих LPWAN-передач, а також застосування нелінійних локаторів і тепловізійного обладнання для фізичного виявлення прихованих пристроїв.

Практична значущість результатів полягає у формуванні комплексного розуміння специфічних ризиків на перетині технологій edge computing, TinyML та систем технічного захисту інформації. Встановлено, що Raspberry Pi Zero 2 W є оптимальним вибором для стаціонарних систем аудіомоніторингу з постійним живленням, тоді як ESP32 або STM32 краще підходять для battery-powered сенсорів та систем. Виявлені загрози вимагають перегляду традиційних підходів до забезпечення інформаційної безпеки та впровадження нових методів контролю, що враховують можливість прихованої локальної обробки аудіоінформації на мініатюрних автономних пристроях з мізерним енергетичним слідом.

Список використаних джерел:

1. Zhang, G., et al. (2024). LaserAdv: Laser Adversarial Attacks on Speech Recognition Systems. USENIX Security Symposium.
2. Zhang, X., et al. (2020). Real-Time, Universal, and Robust Adversarial Attacks Against Speaker Recognition Systems. ICASSP 2020
3. Warden, P., & Situnayake, D. (2019). TinyML: Machine Learning with TensorFlow Lite on Arduino and Ultra-Low-Power Microcontrollers.
O'Reilly Media.