

Проблеми достовірності й експертної оцінки клонованого голосу або голосу, зміненого спотворенням

Ольга Брендель

доктор філософії в галузі права, старший дослідник, завідувач сектору лабораторії інженерно-технічних досліджень, Національний науковий центр «Інститут судових експертиз ім. Засл. проф. М. С. Бокаріуса» Міністерства юстиції України, м. Харків, Україна, ORCID: <https://orcid.org/0000-0002-1723-0203>, e-mail: olgabrenda@ukr.net

Олена Селезньова

головний судовий експерт лабораторії інженерно-технічних досліджень, Національний науковий центр «Інститут судових експертиз ім. Засл. проф. М. С. Бокаріуса» Міністерства юстиції України, м. Харків, Україна, ORCID: <https://orcid.org/0009-0001-9940-6291>, e-mail: seleznova.olena@nncise.org.ua

Висвітлено аналіз сучасних способів фальсифікації аудіоданих, основні особливості проведення судової експертизи відео-, звукозапису клонованого або спотвореного голосу з урахуванням сучасних способів і засобів учинення кримінальних правопорушень.

Ключові слова: експертиза відео-, звукозапису; клонування голосу; нейронні мережі; спотворення голосу; симуляція голосу.

Reliability Issues and Expert Assessment of a Cloned or Distorted Voice

Olha Brendel, Olena Seleznova

Analysis of modern methods of falsifying audio data is presented, as well as the main specifics of conducting forensic examination of a video or sound recording of a cloned or distorted voice taking into account modern methods and means of committing criminal offenses.

Keywords: video and sound recording examination; voice cloning; neural networks; voice distortion; voice simulation.

Стрімкий розвиток інформаційних технологій, зокрема методів штучного інтелекту, зумовив появу програмних засобів, здатних відтворювати (клонувати) й модифікувати голос людини з високим рівнем достовірності. Такі технології широко застосовуються у сфері медіа, телекомунікацій, автоматизованого обслуговування клієнтів та індустрії розваг. Водночас їх застосування створює нові виклики для кримінального судочинства й судової експертизи, зокрема в контексті автентифікації аудіо-записів.

Сучасні системи клонування голосу базуються на нейронних мережах і дають змогу синтезувати мовлення, максимально наближене до голосу конкретної особи. Поміж найбільш відомих можна виокремити такі: *ElevenLabs* (платформа, що забезпечує високоякісне клонування голосу на основі коротких аудіозразків); *Resemble AI* (система створення цифрових голосів із можливістю моделювання емоцій); *Microsoft Azure Speech*

(хмарний сервіс синтезу мовлення з функцією кастомізації голосу); *iSpeech* (платформа тексту мовлення з базовими функціями персоналізації) тощо. Зазначені системи застосовують такі технології: нейронні мережі глибокого навчання (*deep learning*); моделі перетворення тексту на мовлення (*Text-to-Speech, TTS*); векторні представлення голосу (*speaker embeddings*). Це дає змогу відтворювати інтонацію, тембр, ритм мовлення та інші характеристики голосу [1, 2].

Окрім клонування, широко застосовуються засоби зміни (модифікації) голосу, якими можна користуватись як у реальному часі, так і під час постобробки аудіо. До таких належать: *Voicemod* (програма для зміни голосу в режимі реального часу); *MorphVOX* (інструмент для модифікації тембру та висоти голосу); *Clownfish Voice Changer* (простий засіб інтеграції голосових ефектів у комунікаційні платформи) тощо. Основні методи спотворення містять: зміну висоти тону (*pitch shifting*);



зміну формант (*formant shifting*); застосування вокодерів; додавання шумів і ефектів [1, 2].

Отже, програмні засоби клонування та спотворення голосу, засновані на сучасних технологіях штучного інтелекту, значно розширюють можливості створення й модифікації мовлення. Водночас їх застосування створює суттєві ризики для достовірності аудіодоказів у судовому процесі. Спотворення голосового сигналу, незалежно від його походження (фізіологічного, технічного або навмисного), призводить до порушення стабільності параметрів, які в нормальних умовах мають ідентифікаційну цінність. Це так само знижує надійність традиційних методів фоноскопичної ідентифікації та потребує перегляду підходів до оцінювання експертних результатів. У зв'язку з цим виникає об'єктивна необхідність удосконалення методів судово-фоноскопичної експертизи, спрямованих на виявлення ознак штучного походження мовлення й забезпечення надійності доказової бази. Кожен із методів дослідження голосу має власні обмеження, які в умовах спотворення проявляються особливо виразно. Інструментальні методи акустичного аналізу залежать від точності відтворення спектральних і часових характеристик сигналу. Будь-які втручання в частотний спектр або динамічний діапазон сигналу можуть призвести до спотворення або втрати ідентифікаційно значущої інформації. Особливого значення набуває коректне представлення результатів у формі висновків, що відображають ступінь їхньої надійності та можливі альтернативні інтерпретації [3].

Достовірність висновку експерта в умовах спотвореного мовлення визначається не лише наявністю певних акустичних або перцептивних ознак, але й можливістю відтворювати результати аналізу за аналогічних умов. Значний вплив на це мають характер і глибина спотворення, ступінь збереженості індивідуальних характеристик мовця, а також якість і обсяг досліджуваного мовного матеріалу. Короткі фрагменти мовлення, наявність шумів, компресії та цифрових артефактів істотно обмежують можливість формування обґрунтованих висновків. Особливо проблематичні випадки навмисного

маскування голосу, коли мовець свідомо змінює темп мовлення, висоту основного тону, артикуляцію або застосовує технічні засоби для модифікації сигналу. У таких ситуаціях традиційні акустичні показники можуть втрачати свою індивідуальну специфічність або набувати нестабільного характеру, що унеможлиблює їх використання як надійних ідентифікаційних маркерів.

Експертна оцінка в умовах спотвореного голосу має базуватися на принципах наукової обґрунтованості, комплексності й критичного аналізу. Застосування кількох взаємодоповнюючих методів дає змогу частково компенсувати обмеження кожного з них, однак не усуває необхідності чіткого окреслення меж інтерпретації отриманих результатів. Висновки експерта мають супроводжуватися поясненням ступеня їх надійності та зазначенням чинників, що вплинули на результати дослідження. Умови спотворення підвищують ризик помилкової ідентифікації, зумовленої перцептивними особливостями слухового сприйняття. Це обумовлює необхідність мінімізації суб'єктивного компонента шляхом стандартизації процедур, застосування контрольних матеріалів і залучення кількісних критеріїв оцінки.

Експертна практика потребує чіткого усвідомлення меж застосовуваних методів, коректного формулювання висновків та недопущення їх переінтерпретації в судовому процесі. Перспективи розвитку цього напрямку пов'язані з удосконаленням методів аналізу, розробкою критеріїв оцінки ступеня спотворення та інтеграцією новітніх цифрових технологій у судово-експертну діяльність. Сучасні програмні засоби клонування та спотворення голосу, засновані на технологіях штучного інтелекту, дають змогу відтворювати мовлення з високим ступенем подібності до оригіналу або істотно змінювати акустичні характеристики голосу. Це зумовлює необхідність удосконалення методів судово-фоноскопичної експертизи, спрямованих на виявлення ознак штучного походження аудіозаписів і забезпечення достовірності доказів.

Дослідження голосу, зміненого спотворенням, є складним напрямом судово-екс-

пертої діяльності, який потребує високого рівня наукової підготовки й методологічної обережності. Спотворення голосового сигналу істотно обмежують можливості ідентифікації мовця та знижують інформативність традиційних акустичних і перцептивних ознак.

Подальший розвиток цього напрямку пов'язаний із удосконаленням методів аналізу, розробкою критеріїв оцінки ступеня спотворення та інтеграцією досягнень цифрової криміналістики, акустики та лінгвістики в єдину методологічну систему.

Перелік джерел посилання

1. Taylor Paul Text-to-Speech Synthesis. Cambridge : Cambridge University Press, 2009. 626 p.
2. Крак Ю. В., Стеля І. О. Синтез звуків голосу людини на основі фізических моделей голосових зв'язок і речевого тракту. *Искусственный интеллект*. 2009. № 4. С. 542—547. URL: <https://nasplib.isofofs.kiev.ua/server/api/core/bitstreams/e28963f9-c96f-4d49-994c-e76d770e0b64/content> (дата звернення: 15.04.2026).
3. Савкова Т., Опірський І., Сабодашко Д. Дослідження стійкості біометричних систем автентифікації до атак із застосуванням технології клонування голосу на основі глибинних нейронних мереж. *Кібербезпека: освіта, наука, техніка*. 2024. № 2 (26). С. 27—43. URL: <https://doi.org/10.28925/2663-4023.2024.26.670> (дата звернення: 15.04.2026).