https://doi.org/10.31891/2219-9365-2025-82-41
УДК 004.77: 004.49

МАРЧУК Михайло
Вінницький національний технічний університет
https://orcid.org/0009-0003-4773-6541
smoke22catches@gmail.com
ЛУКІЧОВ Віталій
Вінницький національний технічний університет
https://orcid.org/0000-0002-3423-5436
lukichov.vitalyi@vntu.edu.ua

# МОДЕЛЬ ІНФОРМАЦІЙНОЇ АТАКА НА ПІДПРИЄМСТВА З ВИКОРИСТАННЯМ ПІДРОБЛЕНИХ ЦИФРОВИХ ІДЕНТИЧНОСТЕЙ

*У статті розглядається новий тип кіберзагроз, пов'язаних з використанням технологій deepfake для створення підроблених цифрових ідентичностей у контексті цільових атак на фінансові, корпоративні та державні структури. Метою дослідження є побудова формалізованої моделі інформаційної атаки, що включає етапи збору медіа-даних, налаштування фальшивої цифрової інфраструктури, встановлення довіри з цільовими співробітниками та реалізації зловмисних запитів.*

*Методологія дослідження базується на аналізі інцидентів, зафіксованих у відкритих джерелах, з використанням таксономії The MIT AI Risk Repository. У результаті виявлено, що основними об'єктами імперсонізації є керівники вищої ланки, а самі атаки здебільшого націлені на співробітників, що мають доступ до фінансових ресурсів. Стаття також містить рекомендації щодо проактивного захисту цифрової ідентичності, включаючи використання водяних знаків, багаторівневої верифікації та механізмів типу D-CAPTCHA. Представлені результати можуть бути використані для вдосконалення стратегій кіберзахисту в умовах зростаючої загрози з боку генеративних технологій штучного інтелекту.*

*Ключові слова: deepfake, цифрова ідентичність, соціальна інженерія, кіберзагроза, інформаційна атака, водяний знак, фінансова безпека.*

MARCHUK Mykhailo, LUKICHOV Vitalii
Vinnytsia National Technical University

# MODEL OF INFORMATION ATTACK ON ENTERPRISES USING FAKE DIGITAL IDENTITIES

*The article discusses a new type of cyber threats related to the use of deepfake technologies to create fake digital identities in the context of targeted attacks on financial, corporate and governmental entities. The purpose of the study is to build a formalized model of an information attack, including the stages of collecting media data, setting up a fake digital infrastructure, establishing trust with target employees, and implementing malicious requests.*

*The research methodology is based on the analysis of incidents recorded in open sources using The MIT AI Risk Repository taxonomy. As a result, it was found that the main targets of impersonation are top managers, and the attacks themselves are mostly aimed at employees with access to financial resources. The article also contains recommendations for proactive digital identity protection, including the use of watermarks, multi-level verification, and mechanisms such as D-CAPTCHA. The presented results can be used to improve cybersecurity strategies in the face of the growing threat from generative artificial intelligence technologies.*

*Keywords: deepfake, digital identity, social engineering, cyber threat, information attack, watermark, financial security.*

Стаття надійшла до редакції / Received 22.04.2025
Прийнята до друку / Accepted 11.05.2025

## ПОСТАНОВКА ПРОБЛЕМИ У ЗАГАЛЬНОМУ ВИГЛЯДІ ТА ЇЇ ЗВ'ЯЗОК ІЗ ВАЖЛИВИМИ НАУКОВИМИ ЧИ ПРАКТИЧНИМИ ЗАВДАННЯМИ

З розвитком технологій генеративного штучного інтелекту, зокрема глибоких нейронних мереж, значно зросли можливості створення висоякісного синтетичного мультимедійного контенту, відомого під назвою deepfake. Ця технологія дозволяє у режимі реального часу підмінювати візуальну або голосову ідентичність цільової особи, що відкриває нові вектори для проведення атак соціальної інженерії, спрямованих на корпоративні структури та органи державної влади [6].

На відміну від класичних фішингових кампаній або компрометації паролів, deepfake-атаки експлуатують довіру до зорово або аудіально знайомих ідентичностей. Зловмисники можуть імітувати голос керівника компанії для ініціації фінансових трансакцій або підбурювати державних службовців до небезпечних дій, видаючи себе за високопосадовців. Такі інциденти вже зафіксовано на практиці — наприклад, у 2019 році шахраї використали підроблений голос директора британської компанії для вимагання переказу 220 000 євро, а у 2020-2025 роках відбулося щонайменше п'ять подібних атак на корпоративні та банківські установи з використанням аудіо- та відео-deepfake, що призвело до втрат у сотні тисяч і навіть мільйони доларів.

Особливої уваги заслуговують нещодавні інциденти: у 2024 році співробітник компанії переказав 25 мільйонів доларів на основі відеодзвінка, під час якого шахраї видавали себе за топменеджмент компанії, використовуючи відео та голосові deepfake [7]. У 2025 році фінансовий директор іншої компанії був змушений переказати 670 тисяч доларів після комунікації з особами, які видавали себе за керівництво компанії за допомогою deepfake [8].

Загалом, deepfake дедалі частіше використовується як засіб реалізації високоточних таргетованих атак у сфері фінансів і корпоративного управління [9]. Це створює новий тип загрози, що виходить за межі традиційної моделі кібербезпеки і вимагає розроблення спеціалізованих механізмів виявлення та попередження таких атак.

## МЕТОДОЛОГІЯ ДОСЛІДЖЕННЯ

Метою даного дослідження є формалізація моделі кіберзагроз, пов'язаних із використанням технологій deepfake для створення підроблених цифрових ідентичностей у контексті атак на фінансові та корпоративні структури. Зокрема, акцент зроблено на сценаріях, у яких deepfake-технології використовуються для соціальної інженерії шляхом персоніфікації представників керівного складу організацій з метою введення в оману інших співробітників, отримання доступу до фінансових ресурсів, або отримання конфіденційної інформації.

У рамках методологічного підходу було здійснено систематичний збір і аналіз описів підтверджених випадків атак з використанням deepfake у відкритих джерелах, переважно в медіа. Основу джерельної бази склали публікації зі згадками про інциденти, отримані з AI Incident Database [1] — репозитарію, що акумулює повідомлення в ЗМІ про події, пов'язані з використанням штучного інтелекту. Додатково здійснювався пошук за допомогою пошукової системи Google для виявлення релевантних матеріалів, не охоплених згаданим репозитарієм.

Для оцінки релевантності знайдених інцидентів до предмету дослідження застосовувалась таксономія ризиків зі The MIT AI Risk Repository [2]. Інциденти відбирались за такими критеріями згідно даної таксономії:

- Сутність: людина (ризик виник внаслідок дії людини або групи людей);
- Намір: навмисний (інцидент мав на меті отримання конкретної вигоди або досягнення цілі);
- Домен атаки: шахрайство та цільова маніпуляція.

Додатковими умовами відбору були наявність підтвердженого використання deepfake (аудіо, відео або їх комбінації) та чітко визначена спрямованість атаки на конкретну юридичну особу — фінансову установу, компанію або державну організацію. Обов'язковим також було включення до сценарію атаки персоніфікації співробітника або члена організації з метою маніпулювання іншими представниками цієї ж структури. Результатом такого підходу стало формування якісної вибірки прикладів, що ілюструють механізми, мотивації та наслідки використання deepfake у контексті кіберзлочинної активності.

## АНАЛІЗ АТАК З ВИКОРИСТАННЯМ DEEPFAKE

В додатку в таблиці 1 зібрано інформацію про описані в ЗМІ випадки атак з використанням deepfake на фінансовий, корпоративний та державний сектори. В цілому зібрано інформацію про 8 інцидентів. З них 4 в фінансовому секторі, 3 в корпоративному та 1 в державному (рисунок 1). Найбільше інцидентів припадає на 2024 рік, тоді відбулось 3 інциденти. На 2023 рік припало 2 інцидента. `На всі інші роки припадає по одному інциденту (рисунок 2).

Категорією осіб, яких зловмисники зазвичай імперсоніфікують, є керівники вищої ланки або ж високопосадовці, в той час співробітники, на яких дані атаки спрямовані, мають доступи та повноваження для проведення важливих фінансових операцій та керування критичними системами інфраструктури. Для проведення атаки зловмисники часто просять від таких співробітників зробити дії, які мають залишитись непоміченими для інших спрівробітників компанії та мають ознаки неправомірної діяльності (наприклад, перевести кошти, що належать компанії, на сторонній рахунок).

## МОДЕЛЬ АТАКИ

Використання технології deepfake для щахрайських атак на фінансовий, корпоративний та державний сектори можна розглядати як багатофазову модель атаки. Ця модель окреслює як відбувається кампанія зловмисників від збору даних до реалізації зловмисної мети. Атака, як правило, проходить через наступні етапи:

***1. Збір медіа-даних***

На першому етапі відбувається систематичний збір мультимедійних даних, пов'язаних з особою, яку планується видати за іншу. Зловмисники збирають загальнодоступний або приватний візуальний і аудіоконтент, включаючи фотографії, відеозаписи і зразки голосу.

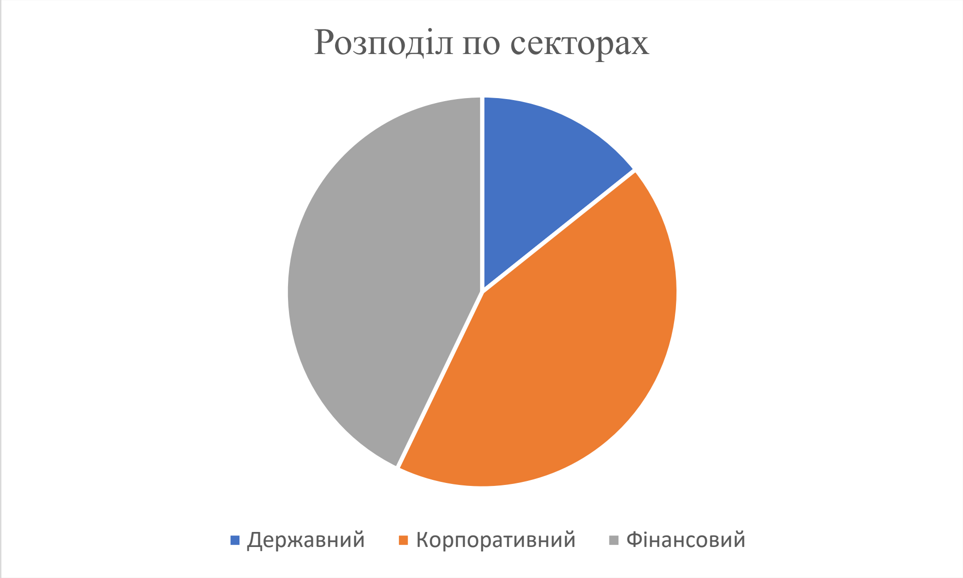


**Рис. 1. Розподіл описаних інцидентів по секторах – державний, корпоративний та фінансовий**

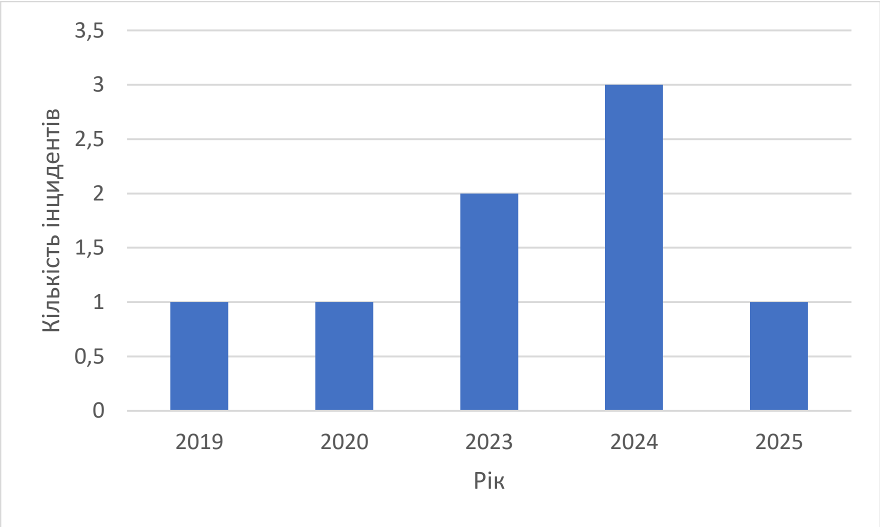


**Рис. 2. Розподіл інцидентів по роках**

Ці матеріали необхідні для навчання моделей глибокого навчання, здатних синтезувати високореалістичні фейкові зображення.

Мета: Отримати достатню кількість високоякісних даних для створення переконливих синтетичних медіа, здатних обходити людський та автоматизований контроль.

Усі загрозу доступу до медіафайлів з зображенням особи, що має бути імперсоніфікована при атаці, виражається множиною:

$$p = \overline{1, P} \tag{1}$$
$$P = P_u \cup P_i \tag{2}$$

Де $P_u$ це множина загроз доступу до публічних каналів (соціальні мережі, ЗМІ, тощо) та $P_i$ це множина загроз доступу до приватних каналів.

***2. Налаштування інфраструктури для імперсоніфікації***

Після того, як зловмисники зібрали медіафайли особи, яку вони хочуть імперсоніфікувати, вони створюють фальшиві канали зв'язку, щоб імітувати цифрову діяльність підробленої особистості. Це, як правило, включає створення фальшивих акаунтів у сервісах електронної пошти, платформах обміну миттєвими повідомленнями та соціальних мережах. Профілі ретельно створюються, щоб відтворити особистість жертви, використовуючи такі елементи, як зображення профілю, імена, посади та стилі спілкування. Для посилення достовірності можуть бути використані підроблені зображення або відео, які безпосередньо використовуються в профілі.

Мета: Створити достовірну цифрову ідентичність, яку можна використовувати для взаємодії з співробітниками компанії чи установи, не викликаючи підозр.

Загрози, пов'язані з підробленням цифрової ідентичності, можна виразити через множину:

$$c = \overline{1, C} \tag{3}$$

***3. Взаємодія з співробітниками цільової організації***

На цьому етапі зловмисники починають взаємодіяти з обраним персоналом організації-жертви. Зазвичай це особи, які мають повноваження затверджувати фінансові операції, розкривати конфіденційну інформацію або надавати доступ до критично важливих систем. Тактика соціальної інженерії використовується для використання довіри та організаційних протоколів, а використання deepfake посилює враження спілкування з справжньою особою. Це може включати відеодзвінки із синтезованими рухами обличчя або аудіоповідомлення, що імітують голос людини, яку видають за іншу особу.

Мета: Встановити контакт і довіру з інсайдерами організації для підготовки до виконання шахрайських інструкцій.

Загрози через контакт зловмисників через підроблені цифрові ідентичності з співробітниками організації описується множиною:

$$s = \overline{1, S} \tag{4}$$

***4. Виконання атаки***

На останньому етапі відбувається передача шкідливих запитів під виглядом особи, яку видають за іншу людину. Ці запити можуть стосуватися термінових фінансових переказів, розкриття конфіденційних даних або надання дозволу на виконання привілейованих дій. Зловмисник покладається на створену на попередніх етапах довіру до себе, щоб змусити жертву виконати вимогу без ретельної перевірки.

Мета: Досягти зловмисної мети, такої як фінансова крадіжка, витік інформації або несанкціонований доступ до системи, за допомогою обману та соціальних маніпуляцій.

Загрози, пов'язані з шкідливими запитами від підроблених цифрових ідентичностей описується множиною:

$$q = \overline{1, Q} \tag{5}$$

Таким чином, модель загроз від використання deepfake для створення підроблених цифрових ідентичностей можна представити у вигляді функціональної залежності:

$$n = f(p, c, s, q) \tag{6}$$

Детальна діаграма атаки показана на рисунку 3.

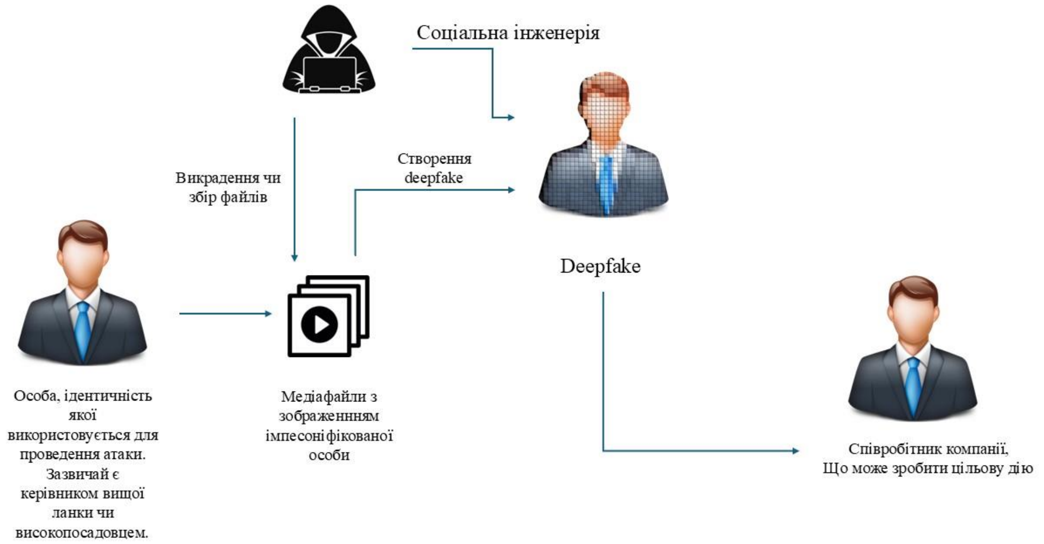


**Рис. 3. Діаграма атаки**

## РЕКОМЕНДАЦІЇ ЩОДО ЗАХИСТУ ВІД АТАК

Ґрунтуючись на вищеописаній моделі можна виділити різноманітні точки вразливості та визначити засоби захисту для кожної з них.

З метою зниження ризику несанкціонованого використання публічно доступних персональних медіафайлів $P_u$, які потенційно можуть бути використані для генерації Deepfake-контенту, доцільно реалізовувати комплексну стратегію обмеження доступу та технічного маркування. В ідеальному випадку, варто мінімізувати обсяг відкритого розміщення таких медіа у відкритих інформаційних середовищах. Однак, у випадках, коли повне виключення медіаконтенту з публічного доступу є неможливим або недоцільним (наприклад, в рамках професійної або комунікаційної активності в соціальних мережах), необхідним є впровадження проактивних заходів захисту.

До таких заходів належать сучасні методи використання стеганографії та водяних знаків, які забезпечують вбудовування прихованої інформації у візуальні або аудіовізуальні дані з метою подальшої ідентифікації їхнього походження та цілісності. Зокрема, ефективним підходом є застосування методів, стійких до активних атак, таких як перетворення, масштабування, компресія або модифікація у процесі створення Deepfake. У цьому контексті особливої уваги заслуговують стратегії маркування водяними знаками, що реалізуються на рівні ознак ідентичності, які, за рахунок вбудовування контрольних сигнатур у простір біометричних ознак, дозволяють достовірно визначити факт маніпуляцій з об'єктом при збереженні високої візуальної подібності до оригіналу [3].

Таким чином, впровадження стеганографічних технологій із підвищеними показниками секретності та стійкості забезпечує стійкий захист персонального медіаконтенту навіть у разі цілеспрямованих спроб його експлуатації зловмисниками для генерації підроблених мультимедійних матеріалів.

Слід зазначити, що вказана вище вразливість, пов’язана з неконтрольованим публікуванням персональних мультимедійних даних у відкритих цифрових середовищах, репрезентує новий клас кіберзагроз, які до останнього часу залишалися поза межами традиційних моделей інформаційної безпеки. На відміну від класичних векторів атак, що орієнтуються на експлуатацію вразливостей програмного забезпечення, мережевих протоколів або політик доступу, даний тип загрози фокусується на експлуатації цифрової ідентичності суб’єкта, зокрема через біометричні ознаки, присутні у відкритих джерелах (наприклад, обличчя, голос, манера мовлення). Таким чином, виникає новий вектор атаки, при якому сам факт присутності цифрових образів особи у відкритому інформаційному просторі стає критичним фактором ризику. Це зумовлює необхідність перегляду існуючих парадигм кіберзахисту з урахуванням нових реалій, у яких контроль над персональними даними набуває значення як засіб превентивної протидії deepfake-технологіям.

Для мінімізації ризиків, пов’язаних із компрометацією цифрових ідентичностей у контексті організаційної кібербезпеки, доцільним є запровадження реєстру дозволених та контрольованих методів автентифікації, а також централізованого переліку верифікованих облікових записів. Такий підхід дозволяє своєчасно виявляти спроби створення підроблених або дубльованих акаунтів, зокрема у внутрішніх системах обміну повідомленнями та сервісах електронної пошти. Стратегія управління цифровими ідентичностями повинна ґрунтуватися на принципах багатофакторної автентифікації, перевірки альтернативних каналів комунікації, ведення обліку привілейованих користувачів та моніторингу відхилень у процедурі реєстрації нових облікових записів або зміни засобів доступу.

Особливу увагу необхідно приділити автентифікаційним сценаріям із залученням людського чинника, для яких рекомендовано застосовувати багаторівневі механізми верифікації: наприклад, підтвердження особи через зареєстрований номер телефону, використання змінних кодів доступу, а також підтвердження запиту через альтернативний незалежний канал [5]. Такі механізми слугують ефективним бар’єром проти атак, що використовують штучно згенеровані зображення чи аудіо задля соціальної інженерії або отримання несанкціонованого доступу

З метою зниження ризиків, пов’язаних із соціально-інженерними атаками, що використовують реалістичні голосові або відео deepfake-імперсонації, доцільним є впровадження механізмів авторизації з активною перевіркою автентичності абонента. Одним із перспективних рішень у цьому контексті є метод D-CAPTCHA, запропонований Yasur L., Frankovits G., Grabovski F. M. та Mirsky Y. [4], що є інноваційною реалізацією CAPTCHA-протоколу, орієнтованого на виявлення фальшивих дзвінків, згенерованих за допомогою deepfake-моделей в режимі реального часу (RT-DF).

На відміну від традиційних CAPTCHA, що вимагають класифікації або інтерпретації контенту, D-CAPTCHA ставить перед зловмисниками завдання генерації контенту, який одночасно повинен відповідати критеріям реалістичності, ідентичності, складності завдання та обмеження в часі. Така постановка виключає можливість використання попередньо підготовлених шаблонів та змушує зловмисника вийти за межі можливостей глибинних нейронних мереж, що лежать в основі RT-DF, спричиняючи появу артефактів або затримок, які легко виявляються як людиною, так і автоматизованими детекторами.

Запровадження D-CAPTCHA в корпоративному середовищі, зокрема на етапах початкової комунікації (вхідні дзвінки, відеозустрічі), дозволяє не лише ефективно протидіяти спробам фішингу чи маніпуляцій з боку зловмисників, але й суттєво знижує ймовірність успішного проникнення в організаційне середовище шляхом компрометації довіри до голосів чи образів. Враховуючи стрімкий розвиток технологій deepfake, активні методи перевірки, подібні до D-CAPTCHA, постають як ключові елементи стратегії стійкого захисту цифрової ідентичності в умовах новітніх кіберзагроз.

Таблиця 1.

**Описи інцидентів, в яких використовувались атаки з deepfake.**

| Опис | Посилання | Тип deepfake | Рік | Сектор |
|---|---|---|---|---|
| Фінансовий директор компанії перевів 670 000 доларів після того як шахраї, що вдавали з себе керівників вищої ланки через deepfake, наказали йому це зробити. | https://www.straitstimes.com/singapore/finance-director-nearly-loses-670k-to-scammers-using-deepfakes-to-pose-as-company-senior-execs | відео, голос | 2025 | Корпоративний |
| Співробітник компанії переказав 25 мільйонів доларів на рахунок шахраїв після того як вони вдали з себе керівників компанії на дзвінку, використавши для цього deepfake. | https://edition.cnn.com/2024/02/04/asia/deepfake-cfo-scam-hong-kong-intl-hnk | відео, голос | 2024 | Корпоративний |
| Шахраї вдавали з себе директора компанії WPP в невдалій спробі переконати працівників компанії переказати кошти через дзвінок | https://incidentdatabase.ai/cite/983/ | голос | 2024 | Корпоративний |
| В американського сенатора намагались вивідати приватну інформацію через deepfake з зображенням Дмитра Кулеби, на той час міністра закордонних справ України. | https://incidentdatabase.ai/cite/805/ | відео, голос | 2024 | Державний |
| Зловмисники намагались вдати з себе клієнта банку та змусити співробітників банку вивести його кошти. | https://incidentdatabase.ai/cite/564/ | голос | 2023 | Фінансовий |
| Зловмисники вдали з себе директора компанії, щоб змусити співробітника банку переказати велику суму | https://incidentdatabase.ai/cite/147/ | голос | 2020 | Фінансовий |
| Шахраї скористались технологією підроблення голосу, щоб вдати з себе директора британської фірми та вимагати переведення 220 000 євро. | https://incidentdatabase.ai/cite/200/ | голос | 2019 | Фінансовий |
| Журналіст використав deepfake для створення аудіозаписів з голосом, щоб обійти банківську систему розпізнавання клієнту по голосу | https://www.vice.com/en/article/how-i-broke-into-a-bank-account-with-an-ai-generated-voice/ | голос | 2023 | Фінансовий |

## ВИСНОВКИ З ДАНОГО ДОСЛІДЖЕННЯ І ПЕРСПЕКТИВИ ПОДАЛЬШИХ РОЗВІДОК У ДАНОМУ НАПРЯМІ

У межах даного дослідження було здійснено аналіз актуальної загрози, пов'язаної з використанням технологій deepfake для створення підроблених цифрових ідентичностей з метою реалізації цільових атак на фінансові установи, корпорації та державні органи. Представлено формалізовану модель загрози, яка враховує особливості таких атак як соціо-технічного феномену, що поєднує високотехнологічні методи генерації мультимедійного контенту з елементами соціальної інженерії.

Аналіз підтверджених інцидентів, зібраних із відкритих джерел, показав, що основними мішенями deepfake-атак є співробітники, які мають доступ до фінансових ресурсів або керують критичними бізнес-процесами, а основними об'єктами імперсонізації є посадові особи вищого рівня. Такі атаки, у більшості випадків, супроводжуються штучно створеним відчуттям терміновості, секретності та авторитетності повідомлення, що істотно підвищує ймовірність успішної маніпуляції.

Виявлено, що існуючі пасивні механізми виявлення deepfake є недостатньо ефективними в умовах зростаючої якості та реалістичності генеративних моделей. У зв'язку з цим обґрунтовано доцільність застосування активних та проактивних методів захисту, зокрема з використанням методів стеганографії для водяних знаків для відслідковування використання публічних медіафайлів.

Отримані результати можуть бути використані для побудови комплексної стратегії захисту цифрової ідентичності в організаціях, зокрема шляхом обмеження публічного доступу до біометричних медіаданих, впровадження стеганографічного маркування, контролю методів автентифікації, а також впровадження процедур додаткової верифікації у критичних комунікаційних сценаріях.

Подальші дослідження доцільно спрямовувати на розробку адаптивних систем виявлення deepfake, здатних до самонавчання в умовах змінного атакуючого середовища, а також на формування нормативноправової бази, що регулює використання технологій генеративного ШІ у високоризикових контекстах.

**Література**

1. AI Incident Database [Electronic resource] // AI Incident Database - https://incidentdatabase.ai/
2. The AI Risk Repository / P. Slattery et al. *AGI - Artificial General Intelligence - Robotics - Safety & Alignment*. 2024. Vol. 1, no. 1. URL: https://doi.org/10.70777/agi.v1i1.10881
3. Proactive Deepfake Defence via Identity Watermarking / Y. Zhao et al. *2023 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, Waikoloa, HI, USA, 2–7 January 2023. 2023. URL: https://doi.org/10.1109/wacv56688.2023.00458
4. Deepfake CAPTCHA: A Method for Preventing Fake Calls / L. Yasur et al. *ASIA CCS '23: ACM ASIA Conference on Computer and Communications Security*, Melbourne VIC Australia. New York, NY, USA, 2023. URL: https://doi.org/10.1145/3579856.3595801
5. Guide for Preparing and Responding to Deepfake Events [Electronic resource] // OWASP GenAI Security Project - https://genai.owasp.org/resource/guide-for-preparing-and-responding-to-deepfake-events/
6. Mirsky Y., Lee W. The Creation and Detection of Deepfakes. *ACM Computing Surveys*. 2021. Vol. 54, no. 1. P. 1–41. URL: https://doi.org/10.1145/3425780
7. Finance worker pays out $25 million after video call with deepfake 'chief financial officer' [Electronic resource] // CNN - https://edition.cnn.com/2024/02/04/asia/deepfake-cfo-scam-hong-kong-intl-hnk
8. Finance director nearly loses $670k to scammers using deepfakes to pose as senior execs [Electronic resource] // The Straits Times - https://www.straitstimes.com/singapore/finance-director-nearly-loses-670k-to-scammers-using-deepfakes-to-pose-as-company-senior-execs
9. Рябокінь М., Котух Є., Папилев О. Вплив технології Deepfake на Fintech сектор: поточний стан в Україні та стратегії запобігання кіберзлочинності. *Problems and prospects of economics and management*. 2024. № 1(37). С. 310–328. URL: https://doi.org/10.25140/2411-5215-2024-1(37)-310-328

**References**

1. AI Incident Database [Electronic resource] // AI Incident Database - https://incidentdatabase.ai/. – (05/27/2025).
2. The AI Risk Repository / P. Slattery et al. *AGI - Artificial General Intelligence - Robotics - Safety & Alignment*. 2024. Vol. 1, no. 1. URL: https://doi.org/10.70777/agi.v1i1.10881 (date of access: 27.05.2025).
3. Proactive Deepfake Defence via Identity Watermarking / Y. Zhao et al. *2023 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, Waikoloa, HI, USA, 2–7 January 2023. 2023. URL: https://doi.org/10.1109/wacv56688.2023.00458
4. Deepfake CAPTCHA: A Method for Preventing Fake Calls / L. Yasur et al. *ASIA CCS '23: ACM ASIA Conference on Computer and Communications Security*, Melbourne VIC Australia. New York, NY, USA, 2023. URL: https://doi.org/10.1145/3579856.3595801
5. Guide for Preparing and Responding to Deepfake Events [Electronic resource] // OWASP GenAI Security Project - https://genai.owasp.org/resource/guide-for-preparing-and-responding-to-deepfake-events/
6. Mirsky Y., Lee W. The Creation and Detection of Deepfakes. *ACM Computing Surveys*. 2021. Vol. 54, no. 1. P. 1–41. URL: https://doi.org/10.1145/3425780
7. Finance worker pays out $25 million after video call with deepfake 'chief financial officer' [Electronic resource] // CNN - https://edition.cnn.com/2024/02/04/asia/deepfake-cfo-scam-hong-kong-intl-hnk
8. Finance director nearly loses $670k to scammers using deepfakes to pose as senior execs [Electronic resource] // The Straits Times - https://www.straitstimes.com/singapore/finance-director-nearly-loses-670k-to-scammers-using-deepfakes-to-pose-as-company-senior-execs
9. Рябокінь М., Котух Є., Папилев О. Вплив технології Deepfake на Fintech сектор: поточний стан в Україні та стратегії запобігання кіберзлочинності. *Problems and prospects of economics and management*. 2024. № 1(37). С. 310–328. URL: https://doi.org/10.25140/2411-5215-2024-1(37)-310-328