КІБЕРБЕЗПЕКА: освіта, наука, техніка
CYBERSECURITY: EDUCATION, SCIENCE, TECHNIQUE
№ 3 (31), 2025
ISSN 2663 - 4023
DOI 10.28925/2663-4023.2025.31.1041
УДК 004.056.53:004.451

Федынишин Тарас Олегович
аспірант кафедри захисту інформації
Національний університет «Львівська політехніка», Львів, Україна
ORCID: 0009-0006-8233-8057
fedynyshyn.taras@gmail.com

Партика Ольга Олександрівна
к.ф-м.н., доцент кафедри захисту інформації
Національний університет «Львівська політехніка», Львів, Україна
ORCID: 0000-0002-3086-3160
olha.o.mykhailova@lpnu.ua

ЕКОСИСТЕМА ВСТАНОВЛЕНИХ МОБІЛЬНИХ ЗАСТОСУНКІВ ДЛЯ ВИЯВЛЕННЯ ПОВЕДІНКОВИХ ГРУП ТА ОЦІНЮВАННЯ ЦИФРОВОЇ РИЗИКОВОСТІ

Анотація. У статті досліджено екосистеми застосунків мобільних пристроїв як джерело криміналістично значущих цифрових артефактів, що дозволяють виокремлювати поведінкові групи користувачів і оцінювати їхню потенційну цифрову ризиковість. У вступі обґрунтовано актуальність використання встановлених мобільних застосунків як індикаторів поведінки, підкреслено значення цифрових слідів та категоризації застосунків у формуванні профілів користувачів. У постановці проблеми акцентовано на недостатній вивченості структури застосунків як поведінкового маркера та потребі у методах, здатних перетворювати набори програм на інформативні криміналістичні характеристики. Аналіз останніх досліджень охоплює сучасні підходи до поведінкового профілювання, оцінювання ризиків приватності даних і мобільної криміналістики, що сформували теоретичне підґрунтя роботи. У теоретичній частині наведено математичну модель представлення користувача як вектора пропорцій категорій застосунків, описано застосування PCA для візуалізації поведінкових груп і визначено принципи обчислення інтегрального ризикового індексу на основі високо ризикових категорій застосунків. У методиці дослідження детально описано процес формування датасету зі скріншотів установлених застосунків, їх класифікацію та перетворення на структуру “користувач — категорія”, а також використані методи аналізу: кластеризація k-means, обчислення силует-метрики, формування профілів кластерів і розрахунок ризикового індексу. У розділі «Результати дослідження» показано існування двох поведінкових кластерів із різними патернами цифрової активності: один із домінуванням фінансових, державних і криптовалютних застосунків, інший — із ширшим мультимедійно-утилітарним профілем. Продемонстровано, що перший кластер характеризується вищим інтегральним ризиковим індексом, що свідчить про підвищену криміналістичну значущість структури його застосунків. Висновки підкреслюють ефективність аналізу встановлених мобільних застосунків для профілювання користувачів, оцінювання ризиковості та формування контекстних моделей цифрової поведінки. Окреслено перспективи подальших робіт, зокрема розширення вибірки, інтеграцію часових характеристик, використання глибинних моделей для кластеризації та побудови профілів, а також поглиблене врахування дозволів застосунків і мережевої активності.

Ключові слова: мобільна цифрова криміналістика; встановлені мобільні застосунки; кластеризація користувачів; поведінкові профілі; ризиковий індекс; цифрові артефакти.

ВСТУП

Смартфони перетворилися на ключові джерела цифрових доказів, оскільки містять широкий спектр артефактів, що відображають повсякденну діяльність користувача. Попередні дослідження [1] показали, що встановлені мобільні застосунки та їхні цифрові сліди можуть суттєво доповнювати криміналістичний аналіз, оскільки містять інформацію про звички, активності та типові сценарії використання пристрою. Застосунки відіграють важливу роль як індикатори поведінки користувача, а їх сукупність може формувати унікальний цифровий профіль.

Подальші дослідження розширили це бачення, демонструючи, що дані, згенеровані мобільними застосунками, можуть використовуватися для реконструкції детальних моделей поведінки. Наприклад, Huber [2] показав, що навіть дані з ігрових застосунків можуть бути використані для відтворення рухових та тимчасових профілів користувачів. Це підтверджує, що структура встановлених застосунків, а також їх функціональні категорії, можуть служити основою для побудови поведінкових та контекстуальних профілів, релевантних для цифрової криміналістики.

У той самий час дослідження безпеки мобільних екосистем [3] підкреслюють, що різні категорії застосунків пов'язані з різними рівнями ризику, включно з доступом до чутливих даних, функціями приватності, можливостями обходу контролю та фінансовими операціями. Це створює передумови для формування інтегральних показників цифрової ризиковості, які можуть бути використані для криміналістичної оцінки потенційно небезпечної поведінки.

З огляду на важливість встановлених застосунків як цифрових артефактів, а також на потенціал їх структурного аналізу для побудови поведінкових моделей, у даній роботі пропонується підхід до кластеризації користувачів за їх “екосистемою застосунків” та аналізу цифрової ризиковості через категорії застосунків. Робота спрямована на розвиток напряму мобільної криміналістики, де встановлені застосунки використовуються не лише як джерело артефактів, а як поведінкові індикатори, здатні відображати приховані патерни та можливі ризики.

Постановка проблеми. У цифровій криміналістиці мобільних пристроїв основна увага традиційно приділяється відновленню артефактів та подій, тоді як структура встановлених застосунків як джерело поведінкової інформації залишається недостатньо дослідженою. Сукупність застосунків може відображати рівень цифрової активності, технічну обізнаність, використання інструментів приватності чи фінансових сервісів і, відповідно, потенційну ризиковість користувача.

Наразі відсутні методи, що дозволяють системно аналізувати екосистему встановлених застосунків для виявлення поведінкових груп, оцінювання цифрового ризику або формування криміналістичних профілів. Це створює потребу у підходах, які перетворюють набір застосунків на інформативні показники, здатні підтримати криміналістичний аналіз ідентифікації потенційно небезпечних або аномальних користувачів.

Аналіз останніх досліджень і публікацій. Проблематика аналізу поведінки користувачів мобільних пристроїв та оцінювання приватнісних і безпекових ризиків, пов’язаних із використанням мобільних застосунків, активно розвивається у науковій спільноті. Однією з ключових ліній досліджень є профілювання користувачів за шаблонами використання застосунків. У роботі Agrawal та співавт. [4] показано, що статистика взаємодії користувачів із застосунками може бути використана для виявлення поведінкових груп шляхом застосування кластерного аналізу та методів машинного

навчання. Дослідження підтверджує, що типи встановлених або активно використовуваних застосунків можуть виступати індикаторами звичок, інтересів та ризикових патернів.

Інша група робіт зосереджена на формалізації приватнісних ризиків і розробці методів обчислення інтегральних показників небезпеки, пов'язаних із мобільними застосунками. Chih-Chang K. та колеги [5] запропонували фреймворк для обчислення кількісного показника приватнісного ризику на основі дозволів, типів зібраних даних та характеристик розробника. Модель є важливим методологічним підґрунтям для формування інтегрального індексу поведінкового ризику в мобільних екосистемах. У роботі Hamed та El-Kharashi [6] розглянуто підхід до оцінювання приватнісних ризиків на основі аналізу дозволів мобільних застосунків і рівня обізнаності користувачів, що демонструє розрив між технічними загрозами та суб'єктивним сприйняттям безпеки.

Важливим напрямом є дослідження криміналістично релевантних артефактів мобільних систем. Дослідження Belkasoft / Gladyshev [7] показує, що операційна система Android зберігає значний обсяг даних про встановлені застосунки та їх використання, що може бути використано для реконструкції дій користувача. Робота Petraitytė та Dehghantanha [8] пропонує системний фреймворк мобільної телефонної криміналістики, у якому аналіз застосунків розглядається як окремий етап розслідування, що підкреслює важливість структури програмного середовища користувача як джерела доказової інформації.

Новітні дослідження розширюють можливості криміналістичного аналізу застосунків. Spichiger та Adelstein [9] представили підхід Argus — інструмент для комплексного аналізу мобільних застосунків, що поєднує статичний і динамічний аналіз, журналювання та аналіз артефактів. У роботі Bardhan та співавт. [10] демонструється застосування мобільної криміналістики для аналізу фінансового застосунку, що підтверджує роль специфічних доменів (банківських, торгових, криптовалютних сервісів) у формуванні ризикових профілів користувачів. Подальші дослідження, такі як робота Frontiersin team [11], фокусуються на аналізі соціальних медіа для виявлення шахрайства, кібербулінгу та дезінформації, демонструючи значення поведінкових індикаторів у цифровому слідстві.

Важливий внесок у формалізацію приватнісних ризиків зроблено також у дослідженні Par NSF [12], де пропонується оцінювання ризику мобільних застосунків на основі сценаріїв, сформульованих користувачами. Це дозволяє врахувати контекст використання, що є важливим для поведінкових моделей ризику. Нарешті, у роботі Dienlin та співавт. [13] показано, що рішення користувачів щодо розкриття даних суттєво залежать від релевантності приватності конкретної інформації та поведінкового контексту, що підкреслює складність формування інтегральних ризикових моделей.

Узагальнення цих досліджень показує: (1) встановлені мобільні застосунки є джерелом криміналістично значущої інформації, (2) категоризація застосунків може відображати поведінкові та ризикові патерни, (3) існують формалізовані моделі приватнісних ризиків, але бракує робіт, які б поєднували кластеризацію користувачів за складом встановлених застосунків із побудовою інтегрального ризикового індексу. Саме цю прогалину заповнює проведене дослідження.

Мета статті. Метою статті є розроблення та апробація підходу до криміналістичного аналізу екосистеми встановлених мобільних застосунків з метою виявлення поведінкових груп користувачів та оцінювання їхньої цифрової ризиковості.

ТЕОРЕТИЧНІ ОСНОВИ ДОСЛІДЖЕННЯ

Криміналістичний аналіз установлених мобільних застосунків ґрунтується на припущенні, що структура програмного середовища смартфона відображає поведінкові патерни користувача та може бути використана для формування ризикових профілів. У цифровій криміналістиці набір застосунків розглядається як важливий непрямий артефакт, що дозволяє відтворювати звички, інтереси, типи активності та потенційні загрози, пов'язані з використанням інструментів, які ускладнюють слідство (VPN, криптовалютні сервіси, захищені месенджери, технічні утиліти тощо).

У цьому дослідженні кожен користувач розглядається як точка у багатовимірному поведінковому просторі, що описується вектором пропорцій категорій застосунків:

$$\boldsymbol{p_i} = (p_{i1}, p_{i2}, \dots, p_{ik}) \quad (1)$$

де $\boldsymbol{p_{im}}$ — частка застосунків категорії m у профілі користувача. Це дозволяє порівнювати користувачів між собою за структурою їх цифрового середовища.

Для візуального представлення поведінкових груп застосовується метод головних компонент (PCA), який зменшує розмірність початкового простору ознак та дозволяє зобразити користувачів у двовимірній площині. Трансформація описується виразом:

$$\boldsymbol{z_i} = W_2^{\top}(\boldsymbol{p_i} - \boldsymbol{\mu}) \quad (2)$$

, де $\boldsymbol{W_2}$ — матриця двох головних компонент, а $\boldsymbol{\mu}$ — середній вектор ознак. Такий підхід дає змогу зменшити розмірність даних для виявлення природних груп (кластерів) користувачів на основі подібності їхніх векторів.

Характерні риси кожної поведінкової групи визначаються через середні значення пропорцій категорій застосунків у межах кластера:

$$P_{km} = \frac{1}{|S_k|}\sum_{p_i \in S_k} p_{im} \quad (3)$$

, де S_k — множина користувачів у кластері. Ці дані використовуються для побудови теплових карт, що відображають профілі поведінкових груп.

Для оцінювання потенційної цифрової ризиковості користувачів вводиться інтегральний ризиковий індекс, який базується на сукупності додатків з високим криміналістичним значенням. Для множини ризикових категорій R індекс обчислюється так:

$$\text{RiskScore}_i = \sum_{C_m \in R} p_{im} \quad (4)$$

Цей показник використовується для оцінки того, наскільки поведінка користувача може ускладнювати цифрове розслідування.

Виділення категорій VPN-сервісів, криптовалютних застосунків та захищених месенджерів як високоризикових спирається на попередні дослідження цифрової криміналістики та приватнісного ризику. У роботах з поведінкового профілювання встановлено, що конкретні типи застосунків можуть виступати індикаторами атипових поведінкових моделей та підвищених загрозових сценаріїв [4]. Дослідження приватнісних ризиків показують, що VPN-застосунки мають підвищений ризик через здатність змінювати мережеве середовище, обходити моніторинг і використовувати permission-набори з високою чутливістю [5], [6], [12]. Криптовалютні та фінансові застосунки пов'язані зі специфічними артефактами, високою концентрацією конфіденційних даних та значними приватнісними наслідками, що підтверджено дослідженнями у сфері фінансової мобільної криміналістики [10]. Захищені месенджери, відповідно до робіт, присвячених аналізу соціальних платформ, є важливим джерелом цифрових доказів, але водночас створюють значні труднощі для слідства через шифрування та ephemeral-повідомлення [11], а тому класифікуються як high-forensic-

relevance категорія. Крім того, дослідження системних артефактів [7] та форензік-фреймворків [8-9] підтверджують, що перелічені класи застосунків залишають сліди, що суттєво впливають на реконструкцію поведінкових профілів і повинні розглядатись як чинники підвищеної цифрової ризиковості. З урахуванням цих аргументів вибір наведених категорій як високоризикових є методологічно обґрунтованим і узгодженим із сучасними підходами у сфері мобільної цифрової криміналістики.

Таким чином, теоретичні основи дослідження поєднують поведінкову аналітику, методи кластеризації, оцінювання цифрового ризику та підходи мобільної криміналістики. Використання встановлених застосунків як криміналістичного артефакта дозволяє формувати поведінкові моделі, релевантні для ідентифікації потенційно небезпечних осіб у межах мобільного цифрового середовища.

МЕТОДИКА ДОСЛІДЖЕННЯ

Методика проведення дослідження ґрунтується на аналізі встановлених мобільних застосунків як поведінкових цифрових артефактів, що можуть бути використані під час криміналістичної оцінки ризиків та ідентифікації потенційно небезпечних осіб. Експериментальна база сформована на основі анонімного набору скріншотів списків застосунків, отриманих від 100 студентів, які надали від трьох до семи зображень із повним переліком встановленого програмного забезпечення. Приклад скріншотів зображено на рисунку 1. Дані було де персоналізовано за допомогою умовних ідентифікаторів, що унеможливило відновлення особи користувача й забезпечило відповідність етичним вимогам цифрових досліджень.

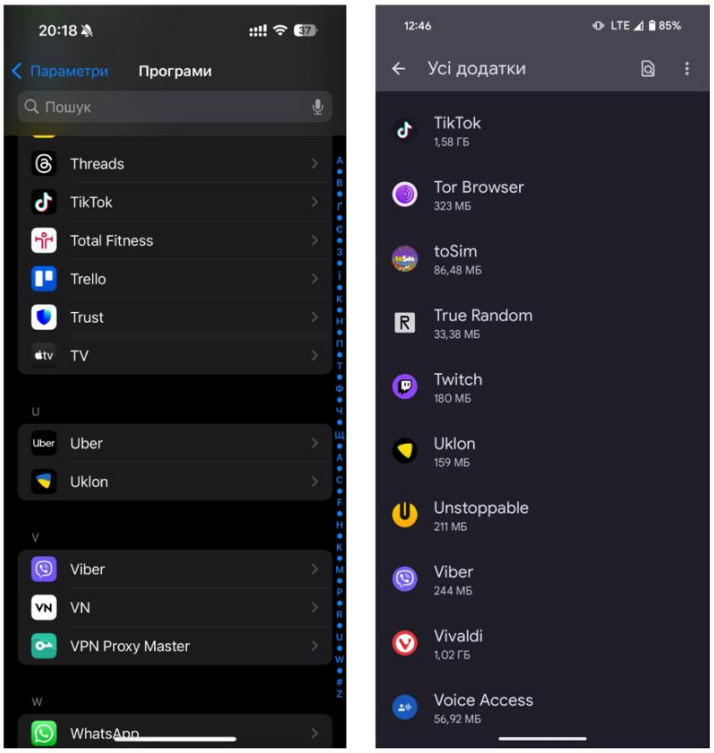

Рис. 1. Приклад скріншотів списків застосунків

Підготовка даних включала автоматизоване зчитування найменувань застосунків зі скріншотів, їх нормалізацію та уніфікацію, а також подальшу ручну та напівавтоматичну перевірку коректності записів. Кожен застосунок було віднесено до однієї з узгоджених категорій, що відображають основні сегменти мобільної екосистеми, включно з інструментами, комунікаціями, фінансовими сервісами, засобами безпеки, криптовалютними платформами, системними компонентами тощо. На основі цієї класифікації було сформовано структуровану таблицю «користувач — застосунок — категорія», яка виступає основою подальшого аналізу.

Методи дослідження охоплюють кілька взаємопов'язаних етапів. На першому етапі виконано кластеризацію користувачів за пропорціями різних категорій застосунків, що дозволило виявити природні групи мобільної поведінки. Кластеризація здійснювалася методом k-means із попередньою нормалізацією даних і подальшою валідацією кількості кластерів за інерцією, силует-коефіцієнтом та стабільністю результатів. На другому етапі дослідження було орієнтовано на виявлення ризикових технологічних індикаторів. Для цього визначено перелік категорій, що у світлі сучасних криміналістичних практик корелюють із підвищеною цифровою ризиковістю, зокрема криптовалютні сервіси, VPN-інструменти, застосунки анонімізації та захищені канали комунікації. Кількісне представлення таких категорій у профілі кожного користувача дозволило сформувати інтегральний показник ризиковості та порівняти його розподіл між кластерними групами.

Наступним етапом стало формування узагальнених поведінкових профілів кластерів, які відображають домінувальні типи мобільної активності, інтенсивність використання тих чи інших типів сервісів та можливі криміналістичні інтерпретації таких структур. Це забезпечило можливість зв'язати статистичні результати з реальними сценаріями поведінки користувачів у цифровому середовищі.

Для оцінювання досліджуваних об'єктів використовувалися показники розподілу категорій застосунків, середні профілі кластерів, різноманіття застосунків на основі ентропії, розподіл ризикових компонентів і характерні цифрові ознаки кожної групи. Обчислення та візуалізація виконувалися засобами Python із використанням бібліотек pandas, scikit-learn, matplotlib і seaborn, що забезпечило відтворюваність та прозорість аналітичних результатів.

Запропонована методика дозволяє оцінювати поведінкові моделі користувачів за їхнім встановленим програмним забезпеченням, забезпечує інтерпретацію отриманих кластерів у криміналістичному контексті та створює основу для подальших досліджень, спрямованих на виявлення потенційно небезпечних осіб за цифровими артефактами мобільних пристроїв.

РЕЗУЛЬТАТИ ДОСЛІДЖЕННЯ

У ході дослідження встановлено, що структура встановлених мобільних застосунків може бути використана як інформативний цифровий артефакт для виокремлення поведінкових груп користувачів та оцінювання їхньої потенційної цифрової ризиковості. Після формування пропорцій категорій застосунків для кожного користувача було проведено кластеризацію, результати якої продемонстрували існування двох чітко розмежованих груп із вираженими відмінностями у цифрових патернах.

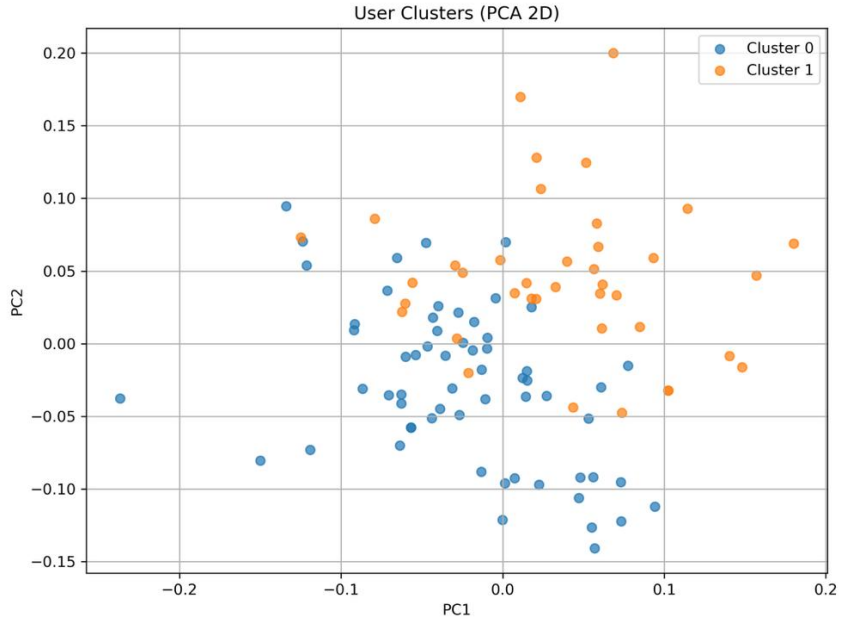

Рис. 2. Кластеризація користувачів у PCA-просторі

У проєкції головних компонент (рис. 2) ці групи утворюють добре відокремлені області: перша група характеризується більш компактною структурою, що свідчить про відносно однорідний набір застосунків утилітарного характеру, тоді як друга демонструє значно ширший спектр поведінкових варіацій, включно з активним використанням мультимедіа, офісних інструментів, утиліт і персоналізаційних сервісів. Така просторово-статистична сепарація підтверджує існування різних стилів цифрової активності, релевантних для криміналістичної інтерпретації.

Таблиця 1

Значення індексу силуету для різної кількості кластерів k

Кількість кластерів	Індекс силуету
2	0.103
3	0.080
4	0.080
5	0.101
6	0.088

Для оцінки оптимальної кількості кластерів було розраховано індекс силуету для k = 2…6 (таблиця 1). Максимальне значення спостерігалося при k = 2 (silhouette = 0.103), що свідчить про найстабільнішу кластерну структуру. Значення для k = 5 (0.101) були близькими, однак подальше збільшення кількості кластерів не призводило до покращення метрики. Значення силуету нижче 0.12 є типовими для поведінкових даних високої варіативності, тому модель із k = 2 була обрана як основна, оскільки забезпечує найбільш інтерпретовану та статистично стійку сегментацію користувачів.

Поглиблений аналіз категорій застосунків показав суттєві відмінності у структурі цифрової поведінки між двома групами.

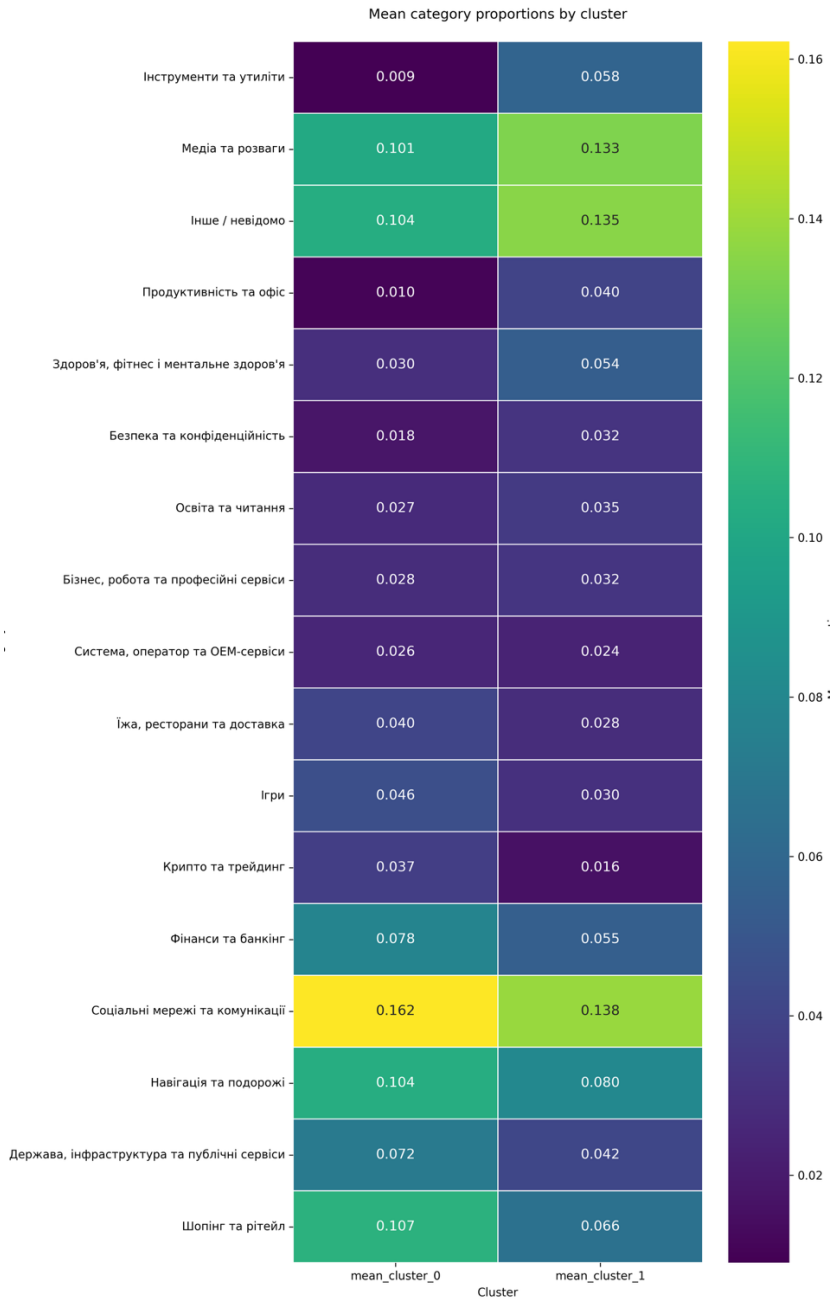

Рис. 3. Середні пропорції категорій застосунків у кластерах

Теплова карта середніх пропорцій категорій (рис. 3) демонструє, що користувачі другої групи мають помітно вищу частку утиліт, медіа, офісних інструментів та інструментів для продуктивності, що відображає їх більш різноманітний і персоналізований стиль використання мобільних пристроїв. Натомість перша група характеризується підвищеною частотою появи застосунків фінансового та державного призначення, а також криптовалютних сервісів. Виявлений контраст свідчить про наявність двох різних типів цифрової поведінки: більш техніко-фінансової, сфокусованої

на операційних сервісах, та більш побутово-утилітарної, орієнтованої на широкий медіа- та інструментальний функціонал.

Окремий етап дослідження був присвячений розрахунку інтегрального ризикового індексу, заснованого на частці встановлених застосунків високого ризику — VPN, криптовалютних сервісів, інструментів анонімізації та безпечного зв'язку.

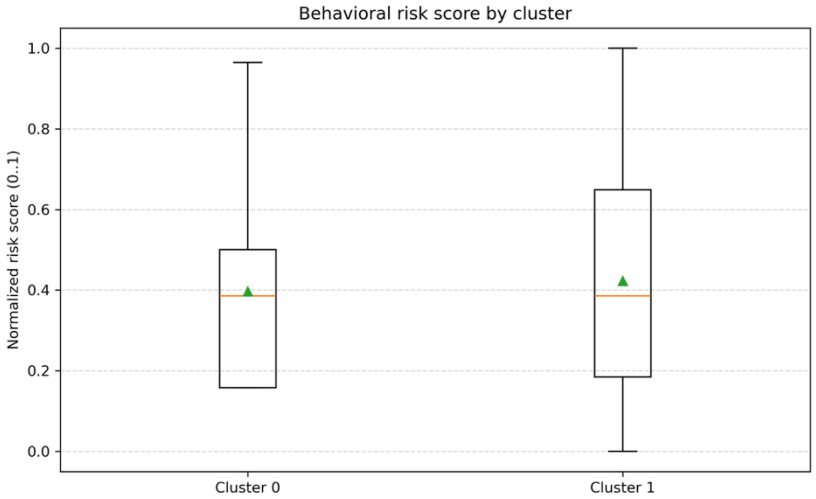

Рис. 4. Розподіл інтегрального ризикового індексу за кластерами

Розподіл значень індексу між кластерами (рис. 4) демонструє, що перша група має загалом вищий рівень потенційної цифрової ризиковості, що пов'язано з більшою поширеністю криптовалютних та захисних сервісів. Друга група, навпаки, переважно концентрується у нижньому та середньому діапазонах ризику, що логічно випливає з її переважно побутового і мультимедійного профілю. Такий контраст підкреслює практичну цінність аналізу застосункової структури як непрямого індикатора можливих ризикових закономірностей у контексті цифрової криміналістики.

Узагальнено отримані результати підтверджують, що аналіз екосистеми встановлених мобільних застосунків є ефективним інструментом для виокремлення поведінкових груп, виявлення латентних характеристик цифрової активності та оцінювання ризиковості користувачів. Це створює підґрунтя для подальшої побудови криміналістично орієнтованих профілів та розширення методів цифрового аналізу мобільних пристроїв у контексті виявлення потенційно небезпечних осіб.

ВИСНОВКИ ТА ПЕРСПЕКТИВИ ПОДАЛЬШИХ ДОСЛІДЖЕНЬ

Проведене дослідження продемонструвало, що структура встановлених мобільних застосунків може слугувати інформативним цифровим артефактом для ідентифікації поведінкових груп користувачів та оцінювання їхньої потенційної цифрової ризиковості. Кластеризація показала існування двох стабільних та чітко відмежованих груп, які відрізняються за типами активності, насиченістю утиліт та мультимедійних сервісів, а також часткою застосунків підвищеного ризику. Побудований інтегральний ризиковий індекс дозволив кількісно оцінити рівень потенційної ризиковості та виявити користувачів, чиї цифрові патерни потребують глибшого криміналістичного аналізу. Отримані результати підтверджують актуальність використання застосункових

екосистем як джерела непрямих поведінкових ознак і показують потенціал такого підходу для аналізу мобільних пристроїв у контексті виявлення потенційно небезпечних осіб. Запропонована методика може бути інтегрована у ширші цифрові криміналістичні фреймворки, зокрема у поєднанні з даними журналів подій, активності у соцмережах, аналізом мережевих слідів та контекстною інформацією про пристрій.

Разом із тим результати дослідження мають низку методологічних обмежень. Запропонований інтегральний ризиковий індекс ґрунтується на рівноважному підході до всіх категорій застосунків і не враховує вагові коефіцієнти, частоту використання чи відмінності у загальній кількості встановлених застосунків у кожного користувача. Індекс не був нормований на розмір вибірки та не проходив статистичної або експертної валідації, що обмежує його інтерпретованість у ширшому криміналістичному контексті.

Крім того, хоча кластеризація виявила дві поведінкові групи, показники якості кластерів (зокрема індекс силуету) залишаються низькими, що вказує на слабко окреслену структуру даних. Це означає, що отримані групи не можна трактувати як чіткі або причинно обумовлені патерни поведінки. Однорідність вибірки, обмеженої студентами одного університетського середовища, також не дозволяє робити узагальнені висновки про ширшу популяцію користувачів. Тому результати слід розглядати як первинну демонстрацію методології, а не як остаточну модель поведінкового ризику.

Перспективи подальших досліджень включають розширення вибірки за рахунок даних з інших регіонів та соціальних груп, інтеграцію часових характеристик (частоти використання, оновлень, видалень застосунків), а також застосування моделей глибинного навчання для автоматизації кластеризації та побудови профілів. Окремим напрямом є розширення ризикового індексу з урахуванням дозволів застосунків, мережевої активності та взаємодії з сервісами, що дозволить формувати більш точні сценарії криміналістичної оцінки. Подальша інтеграція з мультимодальними LLM-моделями відкриває можливості для автоматичного контекстного тлумачення цифрових патернів користувачів, що посилить ефективність аналізу мобільних пристроїв у прикладних задачах безпеки.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Grispos, G., Glisson, W. B., & Storer, T. (2011). Recovering artefacts from a smartphone used as a remote control for a smart TV. Digital Investigation, 8(3–4), 123–134. https://doi.org/10.1016/j.diin.2011.05.004
2. Huber, M. (2019). Forensic determination of movement and usage profiles using data from Pokémon GO. International Journal of Digital Crime and Forensics, 11(4), 1–16. https://doi.org/10.4018/IJDCF.2019100101
3. Alasmary, W., Saeed, M., Alhaidari, F., & Alrasbi, I. (2022). Data usage-based privacy and security issues in mobile app ecosystems. Library Hi Tech, 40(3), 725–747. https://doi.org/10.1108/LHT-05-2021-0164
4. Agrawal, A., Rahate, P., & Jha, S. (2021). User profiling via application usage pattern on digital devices for behavior analysis. Expert Systems with Applications, 168, 114374. https://doi.org/10.1016/j.eswa.2020.114374
5. Chih-Chang, K., et al. (2020). A Framework for Estimating Privacy Risk Scores of Mobile Apps. In J. Zhou et al. (Eds.), Secure IT Systems, NordSec 2020, LNCS 12556, 201–218. https://doi.org/10.1007/978-3-030-62974-8_13
6. Hamed, A., & El-Kharashi, M. W. (2016). Privacy risk assessment and users' awareness for mobile applications. 2016 11th International Conference on Computer Engineering & Systems (ICCES), 225–232. https://doi.org/10.1109/ICCES.2016.7821992
7. Belkasoft, Gladyshev, P. (2020). Android system artifacts: Forensic analysis of application usage. Digital Investigation, 33, 200900. https://doi.org/10.1016/j.diin.2020.200900

8. Petraitytė, M., & Dehghantanha, A. (2017). Mobile phone forensics: An investigative framework based on forensic-by-design principles. Digital Investigation, 21, S96–S105. https://doi.org/10.1016/j.diin.2017.01.012
9. Spichiger, A., & Adelstein, F. (2025). Argus: A new approach for forensic analysis of apps on mobile devices. Forensic Science International: Digital Investigation, 12(2), 100–117. https://doi.org/10.1016/j.fsidi.2025.301938
10. Bardhan, I., Garay, J. L., & Paravisini, D. (2025). Digital Forensics Analysis of a Financial Mobile Application. Forensic Science International: Reports, 34, 221–230. https://doi.org/10.1109/ISNCC62547.2024.10758975
11. Frontiersin team. (2025). Investigating methods for forensic analysis of social media data to detect cyberbullying, fraud, and fake news. Frontiers in Computer Science, 2, 1566513. https://doi.org/10.3389/fcomp.2025.1566513
12. Par. NSF. (2025). Mobile Application Privacy Risk Assessments from User-authored Scenarios. Privacy Risk Assessment, 11(8), 153–167. https://doi.org/10.5281/zenodo.8026501
13. Dienlin, T., Johnson, M., & Jiang, Y. (2025). Privacy Relevance and Disclosure Intention in Mobile Apps. Computers in Human Behavior, 130, 107–119. https://doi.org/10.3390/bs15030324

КІБЕРБЕЗПЕКА: освіта, наука, техніка
CYBERSECURITY: EDUCATION, SCIENCE, TECHNIQUE
№ 3 (31), 2025
ISSN 2663 - 4023

Taras Fedynyshyn
Ph.D. student in the Department of Information Protection
Lviv Polytechnic National University, Lviv, Ukraine
ORCID: 0009-0006-8233-8057
fedynyshyn.taras@gmail.com

Olha Partyka
Ph.D. in Physics and Mathematics,
Associate Professor in the Department of Information Protection
Lviv Polytechnic National University, Lviv, Ukraine
ORCID: 0000-0002-3086-3160
olha.o.mykhailova@lpnu.ua

ECOSYSTEM OF INSTALLED MOBILE APPLICATIONS FOR IDENTIFYING BEHAVIORAL GROUPS AND ASSESSING DIGITAL RISK

Abstract. The article investigates the ecosystem of installed mobile applications as a source of behaviorally relevant digital artifacts that can be used for user profiling and digital risk assessment in mobile device forensics. Smartphones accumulate a wide range of application-level traces that reflect user habits, interests, and activity patterns, making the structure of installed apps a valuable indirect indicator of behavioral characteristics. The study proposes a methodological framework that transforms application lists into a structured representation of user profiles based on proportional distributions of app categories. Using a dataset of anonymized screenshots collected from 100 participants, each mobile application was classified into a unified taxonomy of categories, enabling the construction of "user–category" behavioral vectors. To identify latent groups of digital behavior, k-means clustering was applied with prior normalization and validated using silhouette metrics. Dimensionality reduction through PCA allowed visualizing cluster separation and exploring structural differences between users. The results reveal two distinct behavioral groups with clearly different application ecosystems: one dominated by financial, governmental, security-oriented, and cryptocurrency apps; the other characterized by broader multimedia, utility, productivity, and personalization tools. A digital risk score was introduced to quantify potential forensic relevance based on the presence of high-risk app categories such as VPN services, anonymization tools, secure messengers, and cryptocurrency platforms. Comparative analysis shows that the first cluster demonstrates substantially higher risk levels, indicating potentially more complex forensic reconstruction scenarios. The study demonstrates that analyzing installed mobile applications enables effective behavioral segmentation, uncovering latent usage patterns and supporting forensic evaluation of digital risk. The proposed methodology offers a scalable approach for integrating application-based analysis into mobile forensics workflows and highlights directions for future work, including expanded datasets, incorporation of temporal usage patterns, and advanced machine-learning models for behavioral interpretation.

Keywords: mobile forensics; behavioral profiling; installed mobile applications; clustering; digital risk score; application ecosystem analysis.

REFERENCES (TRANSLATED AND TRANSLITERATED)

1. Grispos, G., Glisson, W. B., & Storer, T. (2011). Recovering artefacts from a smartphone used as a remote control for a smart TV. Digital Investigation, 8(3–4), 123–134. https://doi.org/10.1016/j.diin.2011.05.004
2. Huber, M. (2019). Forensic determination of movement and usage profiles using data from Pokémon GO. International Journal of Digital Crime and Forensics, 11(4), 1–16. https://doi.org/10.4018/IJDCF.2019100101
3. Alasmary, W., Saeed, M., Alhaidari, F., & Alrasbi, I. (2022). Data usage-based privacy and security issues in mobile app ecosystems. Library Hi Tech, 40(3), 725–747. https://doi.org/10.1108/LHT-05-2021-0164

4. Agrawal, A., Rahate, P., & Jha, S. (2021). User profiling via application usage pattern on digital devices for behavior analysis. Expert Systems with Applications, 168, 114374. https://doi.org/10.1016/j.eswa.2020.114374
5. Chih-Chang, K., et al. (2020). A Framework for Estimating Privacy Risk Scores of Mobile Apps. In J. Zhou et al. (Eds.), Secure IT Systems, NordSec 2020, LNCS 12556, 201–218. https://doi.org/10.1007/978-3-030-62974-8_13
6. Hamed, A., & El-Kharashi, M. W. (2016). Privacy risk assessment and users' awareness for mobile applications. 2016 11th International Conference on Computer Engineering & Systems (ICCES), 225–232. https://doi.org/10.1109/ICCES.2016.7821992
7. Belkasoft, Gladyshev, P. (2020). Android system artifacts: Forensic analysis of application usage. Digital Investigation, 33, 200900. https://doi.org/10.1016/j.diin.2020.200900
8. Petraitytė, M., & Dehghantanha, A. (2017). Mobile phone forensics: An investigative framework based on forensic-by-design principles. Digital Investigation, 21, S96–S105. https://doi.org/10.1016/j.diin.2017.01.012
9. Spichiger, A., & Adelstein, F. (2025). Argus: A new approach for forensic analysis of apps on mobile devices. Forensic Science International: Digital Investigation, 12(2), 100–117. https://doi.org/10.1016/j.fsidi.2025.301938
10. Bardhan, I., Garay, J. L., & Paravisini, D. (2025). Digital Forensics Analysis of a Financial Mobile Application. Forensic Science International: Reports, 34, 221–230. https://doi.org/10.1109/ISNCC62547.2024.10758975
11. Frontiersin team. (2025). Investigating methods for forensic analysis of social media data to detect cyberbullying, fraud, and fake news. Frontiers in Computer Science, 2, 1566513. https://doi.org/10.3389/fcomp.2025.1566513
12. Par. NSF. (2025). Mobile Application Privacy Risk Assessments from User-authored Scenarios. Privacy Risk Assessment, 11(8), 153–167. https://doi.org/10.5281/zenodo.8026501
13. Dienlin, T., Johnson, M., & Jiang, Y. (2025). Privacy Relevance and Disclosure Intention in Mobile Apps. Computers in Human Behavior, 130, 107–119. https://doi.org/10.3390/bs15030324