DOI 10.34185/1562–9945–5–166–2026–03
УДК 004.7.056

О.В. Супрун, Н.В. Карпенко

СУЧАСНИЙ СТАН ТА ТЕНДЕНЦІЇ БЕЗАГЕНТНОГО ПРОФІЛЮВАННЯ ПОВЕДІНКИ КОРИСТУВАЧІВ НА ОСНОВІ МЕРЕЖЕВОЇ ТЕЛЕМЕТРІЇ

Анотація. У статті наведено дослідження потенціалу застосування даних з мережевого потоку (NetFlow) для аналізу поведінки користувачів інформаційних систем без втручання в роботу кінцевих пристроїв. Метою дослідження є комплексна систематизація наукового доробку у сфері безагентного профілювання на основі мережевої телеметрії без порушення приватності. За методологією B. Kitchenham проаналізовано 121 публікацію за 2010–2025 роки, з яких відібрано репрезентативний пул із 25 ключових робіт.

Результати та ключові висновки. Розроблено оригінальну діаграму еволюції технологій кіберзахисту, яка демонструє логічний інженерний перехід від реактивного сигнатурного контролю відомих загроз на хостах до проактивної контекстно–поведінкової оркестрації на базі інтегрованих платформ UEBA та SOAR. Доведено, що використання NetFlow/IPFIX забезпечує неінвазивний безагентний моніторинг, придатний для масштабного застосування в гетерогенних корпоративних мережах. Водночас аналіз виявив технологічну прогалину між високою спостережуваністю людського фактору в мережевих слідах (підтверджено з високою силою доказів у 48% робіт) та низькою спроможністю суто автономних потокових даних NetFlow для точної ідентифікації користувачів (лише 36% робіт з високою силою доказів).

Ключові слова: кібербезпека, поведінка користувача, NetFlow,IPFIX, user behaviour, user identity, UEBA, Behavioural Quantifying.

Вступ та обґрунтування актуальності дослідження. Цифровізація суспільства вийшла на рівень суттєвого впливу на майбутнє нашої цивілізації, оскільки воно напряму залежить від надійних і безпечних програмних систем. Одним із вирішальних факторів для інформаційної безпеки є розуміння поведінки користувача, як найбільш непередбачуваного елемента системи. Відділи безпеки підприємств та організацій аналізують поведінку користувачів за їхніми «цифровими слідами» в журналах подій. Але робиться це, в більшості випадків, інсталяцією додаткового програмного забезпечення на кінцевому пристрої користувача. Враховуючи важливість приватності людини, важливо знайти методи “неінвазійного” отримання “цифрових слідів”. Тим паче, що до пристроїв IoT взагалі неможливо інсталювати додаткове ПЗ для аналізу.

Проблематика безагентного профілювання поведінки користувачів на основі мережевої телеметрії є одним із перспективних напрямів сучасних досліджень у сфері кібербезпеки. Її актуальність зумовлена необхідністю виявлення аномальної або потен-

© О.В. Супрун, Карпенко Н.В., 2026
Copyright for this paper by its authors. Use permitted under License CC BY 4.0.
ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

ційно шкідливої активності без втручання в роботу кінцевих пристроїв, що особливо важливо в умовах зростаючих вимог до приватності, масштабованості та ресурсної ефективності систем моніторингу [1, 2].

Початкові дослідження показали, що NetFlow дає змогу виявляти індивідуальні особливості мережевої активності користувачів [3]. Роботи, присвячені профілюванню користувачів на основі NetFlow, підтвердили принципову можливість їх ідентифікації за статистичними параметрами мережевих потоків. Водночас у цих дослідженнях визначено обмеження такого підходу, зокрема недостатню деталізацію поведінкових проявів та обмежену інформативність flow–даних за відсутності їх контекстного збагачення [2].

Подальший розвиток теми був пов’язаний із переходом від суто інфраструктурного аналізу мережевого трафіку до побудови поведінкових моделей на основі мережевих метаданих. У низці робіт [3, 4, 5] було доведено, що використання ознак рівня flow і network metadata дозволяє здійснювати як ідентифікацію користувачів, так і їхнє профілювання за характером мережевої активності. Такі підходи є привабливими з огляду на відсутність необхідності встановлення агентного програмного забезпечення на кінцеві пристрої, що знижує інвазивність моніторингу та ризики порушення конфіденційності.

Водночас більшість авторів сходиться на тому, що мережеві потоки самі по собі не забезпечують повного розуміння поведінкового контексту. Отже, все частіше застосовується поєднання NetFlow із журналами автентифікації, часовими характеристиками активності, а також контекстом використання корпоративних сервісів. Саме така інтеграція дозволяє підвищити точність моделей, виявляти відхилення від типової поведінки та зменшувати кількість хибнопозитивних спрацьовувань [6, 7].

Окрему методологічну рамку для таких досліджень формує концепція UEBA (User and Entity Behaviour Analytics), яка орієнтована на виявлення аномалій на основі аналізу поведінки користувачів і сутностей у системі. В UEBA мережевий трафік виступає одним із ключових джерел для формування поведінкового профілю, а машинне навчання використовується для побудови моделей нормальної активності та подальшого виявлення відхилень. Отже, UEBA не лише доповнює, а й концептуально підсилює підходи до профілювання користувачів на основі мережевої телеметрії [1].

Загалом проведений аналіз літератури свідчить про те, що в науковій спільноті сформувався стійкий інтерес до використання мережевого трафіку як джерела для безагентного профілювання поведінки користувачів [2, 4, 7]. Разом із цим зберігається дослідницька ніша, пов’язана з розробкою масштабованих, контекстно–збагачених і менш інвазивних методів аналізу, здатних забезпечити високу точність виявлення аномальної активності в корпоративних інформаційних системах.

Мета дослідження полягає у систематизації та комплексному аналізі науково–технічного доробку у сфері безагентного моделювання цифрової активності для визначення сучасного стану, технологічних можливостей та обмежень використання мережевої телеметрії (зокрема NetFlow/IPFIX) у завданнях профілювання користувачів та виявлення аномальної поведінки без порушення їхньої приватності.

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

Передумови дослідження. Для визначення місця поведінкового аналізу в сучасній системі засобів кіберзахисту доцільно розглянути еволюцію ключових технологій інформаційної безпеки та простежити функціональні й концептуальні зв'язки між ними. З цією метою було систематизовано основні класи технологій за періодами їх появи та сформовано узагальнену схему їх розвитку (рис. 1). Ця схема наочно показує, що відбувся поступовий перехід від базових сигнатурних і периметрових засобів до інтегрованих платформ моніторингу, аналітики та автоматизованого реагування. На ранніх етапах основну роль відігравали антивірусні рішення, міжмережеві екрани, IDS/IPS та засоби захисту кінцевих точок, які переважно працювали з відомими загрозами та забезпечували контроль на окремих рівнях інфраструктури.

Подальше ускладнення атак, розширення корпоративної інфраструктури та зростання обсягів телеметрії сприяли розвитку технологій NTA, NDR, EDR, UBA та UEBA. На відміну від засобів, орієнтованих переважно на виявлення окремих відомих ознак компрометації, ці підходи розширили можливості аналізу мережевих взаємодій, активності кінцевих вузлів та поведінки користувачів і сутностей у часовому та контекстному вимірах. Отже, відбувається поступове зміщення акценту від локального виявлення окремих інцидентів до кореляції подій, аналізу поведінки користувачів і сутностей, а також наскрізного моніторингу безпекових процесів у корпоративному середовищі.

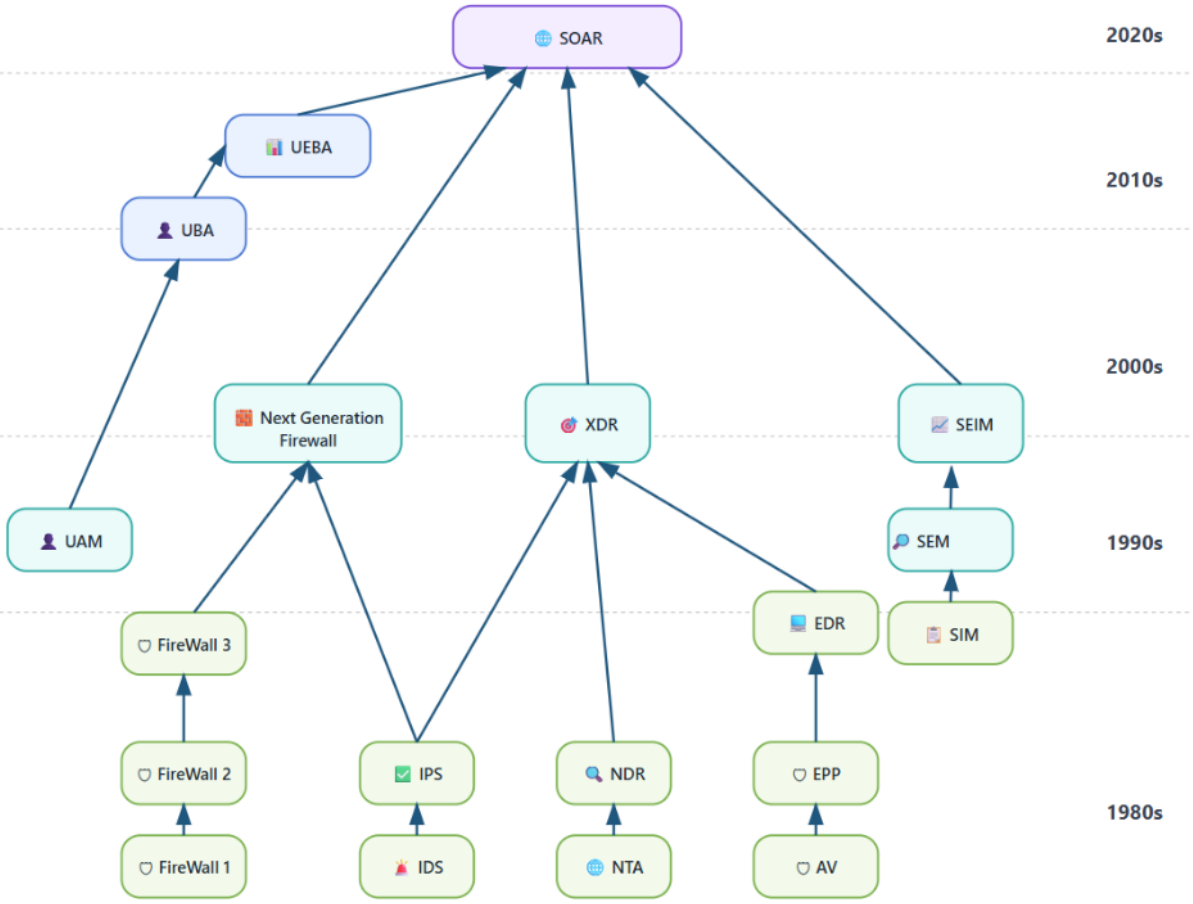


Рисунок 1 – Еволюція технологій кіберзахисту

На сучасному етапі розвиток засобів кіберзахисту характеризується одночасним посиленням аналітичної та автоматизованої складових. Технології “User and Entity Behaviour Analytics” (UEBA) розвивають підходи до контекстного аналізу поведінки користувачів і сутностей на основі різнорідної телеметрії, тоді як платформи “Security Orchestration, Automation, and Response” (SOAR) орієнтовані на оркестрацію процесів безпеки, автоматизацію типових операцій та реагування на інциденти. [8, 9]. У сукупності ці напрями відображають загальну тенденцію переходу від ізольованого виявлення подій до комплексного аналізу контексту та автоматизованого реагування.

Водночас технічні події та мережеві характеристики є лише цифровими слідами активності користувача і не можуть автоматично ототожнюватися з його поведінкою. Перехід від сукупності таких спостережуваних ознак до поведінкової характеристики потребує встановлення стійких взаємозв’язків між технічною активністю, контекстом її виникнення та діями користувача.

Технології поведінкового аналізу посідають важливе місце серед сучасних засобів кіберзахисту. Протягом останніх п’ятнадцяти років цей напрям активно розвивався, що супроводжувалося появою значної кількості методів, моделей та інструментів, орієнтованих на виявлення аномальної активності, побудову поведінкових профілів користувачів і сутностей та ідентифікацію потенційно небезпечних відхилень від нормальної поведінки. Однак, різноманітність джерел телеметрії, способів представлення поведінки та методів її аналізу призвела до суттєвої фрагментації цього напрямку.

Попередній аналіз наукових публікацій показав наявність значної кількості досліджень, присвячених окремим аспектам поведінкової аналітики, UEBA/UBA, аналізу мережевого трафіку та виявленню аномалій. Водночас було виявлено обмежену кількість систематизованих досліджень, у яких ці напрями розглядаються комплексно з позиції можливості безагентного формування поведінкового профілю користувача за мережевими метаданими, зокрема NetFlow/IPFIX. Це визначає необхідність систематизації наявних підходів, оцінювання їхніх можливостей та обмежень, а також визначення невирішених дослідницьких питань у цій предметній області.

Методологія дослідження. Для огляду літератури був використаний спрощений варіант методу «Systematic Literature Reviews» (SLR). Методологічну структуру огляду сформовано з урахуванням рекомендацій Kitchenham та Charters [10] щодо проведення SLR у галузі комп’ютерної інженерії. Принципи Kitchenham використано для організації огляду як послідовного дослідницького процесу.

На початковому етапі огляду було визначено дослідницькі питання, основні джерела пошуку, пошукові терміни та базові критерії релевантності публікацій. Подальша процедура відбору уточнювалася і формалізувалася в процесі аналізу літератури з урахуванням предметної специфіки дослідження.

На етапі повнотекстового аналізу застосовувалося експертне оцінювання предметної релевантності публікацій дослідницьким питанням. Це було зумовлено міждисциплінарним характером предметної області та неоднорідністю термінології, що використ-

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

товується для опису мережевої активності, поведінкового аналізу та профілювання користувачів.

Відповідно до обраного дизайну огляду було сформульовано два дослідницькі питання (RQ), які визначають предметну спрямованість пошуку, критерії відбору публікацій, структуру вилучення даних та подальший синтез результатів.

RQ1: Яким чином людський фактор проявляється в технічних логах інформаційних систем і може бути інтерпретований як поведінкова характеристика користувача? Мета RQ1 полягає у визначенні характеристик технічної активності, які в наявних дослідженнях розглядаються як прояви або непрямі індикатори поведінки користувача.

RQ2: Наскільки ефективно мережеву телеметрію (NetFlow) можна використовувати для побудови поведінкових профілів користувачів без застосування агентного моніторингу? Мета RQ2 полягає в оцінюванні потенціалу та обмежень мережевої телеметрії як джерела даних для безагентного поведінкового профілювання користувачів.

Пошук обмежувався публікаціями за 2010–2025 роки, оскільки цей період охоплює активний розвиток технологій мережевої телеметрії, поведінкової аналітики та UEBA. До подальшого аналізу включалися рецензовані статті, матеріали міжнародних конференцій, оглядові роботи та дисертаційні дослідження, у яких описано використані дані, ознаки, методи аналізу або результати експериментальної оцінки.

Результати дослідження. На етапі ідентифікації пошук публікацій здійснювався у базах IEEE Xplore, ScienceDirect та ResearchGate. Первинний масив було піддано скринінгу за назвами та анотаціями, у результаті чого для подальшого аналізу було відібрано 121 публікацію. Наступне повнотекстове оцінювання дозволило сформувати основний корпус із 22 релевантних досліджень. Для підвищення повноти пошуку додатково застосовано процедуру snowballing, у межах якої було виявлено ще 3 релевантні публікації. Таким чином, остаточний корпус дослідження склав 25 публікацій.

На основі детального аналізу відібраних наукових робіт, дані було систематизовано та згруповано за трьома ключовими напрямами: Research Focus, Behavioural Quantifying / Quantification та Data Scope & Sources.

1. Напрям Research Focus спрямований на класифікацію робіт за їхньою кінцевою прикладною метою. Дослідження чітко розподіляються на три категорії (рис. 2). Більшість робіт [6,12,16,18,19,20,21,22,24,28,29] цієї категорії спрямована на виявлення шкідливої активності, атак, наприклад, impersonation чи masquerade, або статистичних відхилень у мережі без детальної ідентифікації конкретного користувача. Поведінка тут зазвичай маркується «нормальна» або «аномальна».

На створенні сигнатури конкретного користувача за його поведінковими шаблонами в мережі фокусуються 36% дослідницьких робіт [3,4,5,7,13,14,15,23,26]. Головною метою цих робіт є розпізнавання користувача або підтвердження його автентичності на основі поведінкової біометрії без аналізу вмісту пакетів. Останні 5 робіт, що залишились у корпусі [2,11,17,25,27], систематизують наявні методи машинного навчання, описують весь технологічний ланцюг збору даних (наприклад, NetFlow/IPFIX) або пропонують універсальні математичні алгоритми кластеризації без прямої прив’язки до конкретного прикладного завдання. Таким чином, спостерігається домінування дослі-

джень в області безпеки та профілювання, оскільки 80% робіт зосереджені на виявленні аномалій та ідентифікації користувачів.

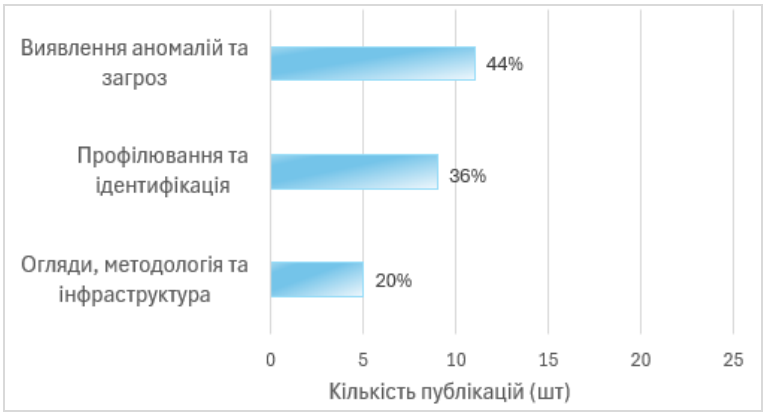


Рисунок 2 – Розподіл досліджень у групі Research Focus

2. Групування робіт за напрямом Behavioural Quantifying / Quantification розподіляється наступним чином:

- 13 з 25 робіт [3,4,5,6,7,13,14,15,16,18,19,23,26] мають пряме моделювання поведінки конкретного користувача. У цих роботах поведінка визначається як індивідуальні патерни активності конкретної людини, наприклад, добові ритми використання додатків, послідовність дій у віртуальних машинах чи унікальні статистичні характеристики сесій.
- у 7 роботах [12,17,21,22,24,28,29] відбувається моделювання не конкретної людини, а активність пристроїв, мережевих вузлів, застосунків чи класів трафіку.
- 20% досліджень [2,11,20,25,27] зосереджені виключно на інфраструктурі моніторингу або загальній класифікації сфер застосування мережевих потоків, без моделювання поведінки користувача.

Таким чином, у цій групі переважає індивідуальний підхід, оскільки 52% досліджень прямо моделюють поведінку конкретного користувача (рис.3).

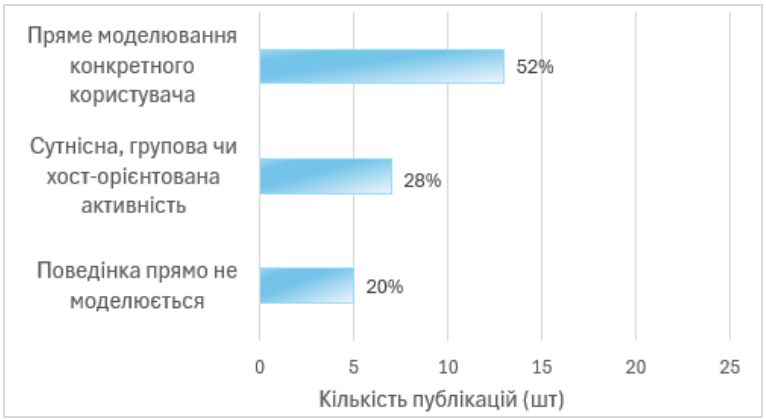


Рисунок 3 – Розподіл досліджень у групі Behavioural Quantifying / Quantification

3. Групування досліджень за Data Scope & Sources дозволило розділити роботи за типом аналізованих даних та джерелами телеметрії, які використовуються для аналізу.

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

Найбільша група досліджень [2,3,11,12,14,15,16,20,23,24,26,28,29] спирається суто на низькорівневі мережеві потоки або метадані трафіку (тривалість сесій, обсяги переданих байтів, порти, IP–адреси). Це підтверджує критичну роль NetFlow як стандарту збору телеметрії, що знижує інвазивність моніторингу та обсяг безпосередньо спостережуваних даних порівняно з агентним моніторингом або аналізом вмісту пакетів (рис. 4).

Понад чверть усіх досліджень [6,17,18,19,21,22,25] спираються на концепцію поведінкової аналітики користувачів та сутностей (UEBA). При цьому дослідники інтегрують мережевий трафік із системними логами, діями у віртуальних машинах, характеристиками пристроїв та іншими джерелами для формування багатовимірного профілю. І лише 5 робіт присвячені профілюванню [4, 5, 7,13,27] , що засноване на взаємодії користувача з конкретними прикладними програмами. Таким чином, у цій групі NetFlow та метадані домінують, оскільки 52% робіт спираються на потоковий аналіз та мережеві метадані (рис. 4).

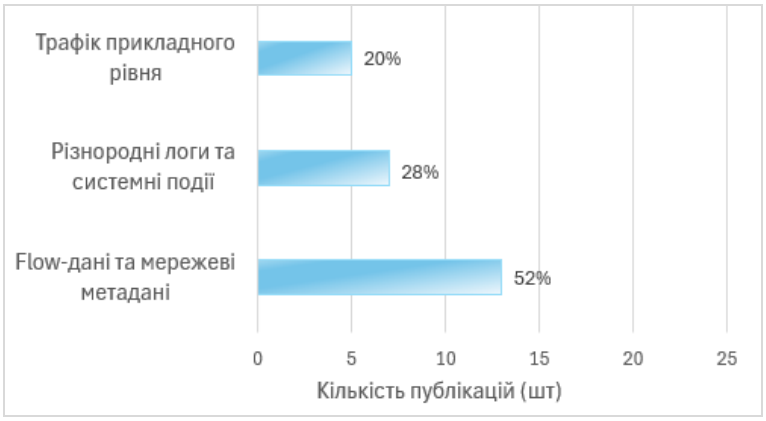


Рисунок 4 – Розподіл досліджень у групі Data Scope & Sources

Аналіз технічних аспектів відібраних наукових праць демонструє чітку спрямованість сучасної академічної спільноти на розробку неінвазивних та масштабованих рішень у сфері поведінкового аналізу. Переважна більшість досліджень (72%) характеризуються виключно високим рівнем безагентності (рис. 5). Вони спираються на збір мережевої телеметрії (такої як NetFlow, IPFIX або загальні метадані сесій) безпосередньо з активного мережевого обладнання або сенсорів, повністю уникаючи необхідності встановлення та підтримки спеціалізованого програмного забезпечення (агентів) на кінцевих пристроях користувачів. Такий підхід спрощує масштабоване впровадження систем аналітики у великих корпоративних мережах, знижує навантаження на кінцеві вузли та усуває ризики несумісності ПЗ на різнорідних операційних системах (зокрема на BYOD–пристроях та системах IoT).

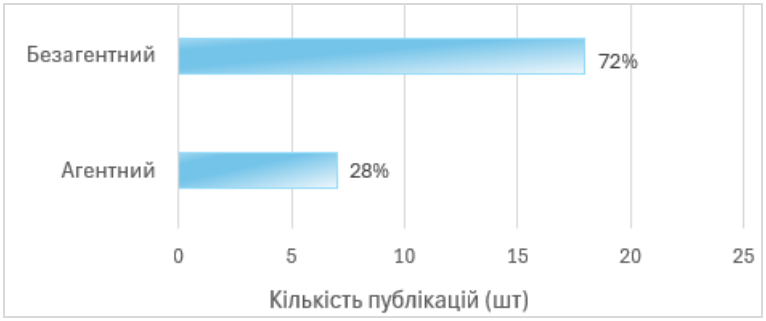


Рисунок 5 – Рівень інтеграції

Водночас дослідження демонструють практично рівний розподіл щодо використання додаткового контексту для збагачення мережевих даних (рис. 6). Майже половина авторів (48%) фокусуються на аналізі «чистих» мережевих характеристик (таких як тривалість сесій, обсяг трафіку чи номери портів) без залучення сторонньої інформації. Проте незначна більшість (52%) успішно інтегрують контекстне збагачення (зокрема логування подій DNS, відображення IP/MAC–адрес, аналіз активності віртуальних машин або системні логи UEBA). Це свідчить про розуміння дослідниками обмеженості чистого NetFlow–аналізу у складних сценаріях кібератак, де для точного розпізнавання легітимного користувача та зниження рівня хибнопозитивних спрацьовувань критично важливо пов’язати фізичну мережеву активність із логічним контекстом дій у системі.

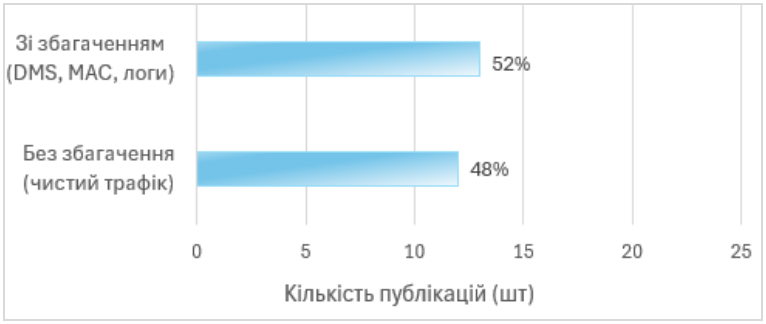


Рисунок 6 – Використання додаткового контексту

Аналіз доказів першого дослідницького питання (RQ1) свідчить про наявність стійкої теоретичної та експериментальної бази для моделювання людського фактору через непрямі ознаки (рис. 7). Автори 48% проаналізованих робіт успішно довели, що поведінка людини у цифровому середовищі матеріалізується у формі стабільних часових ланцюжків (рис. 8) та відповідає патернам комплексів потоків активності, що залежать від конкретного користувача. Людський фактор стає чітко спостережуваним як через агреговані мережеві метадані, так і за допомогою часових профілів користувача, сформованих на основі кількох джерел даних. Лише 4 роботи (16%) показали низьку силу доказів через надмірне спрощення, де індивідуальна поведінка людини фактично прирівнювалася до загальних мережевих аномалій без урахування контексту її ролі чи ідентичності.

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

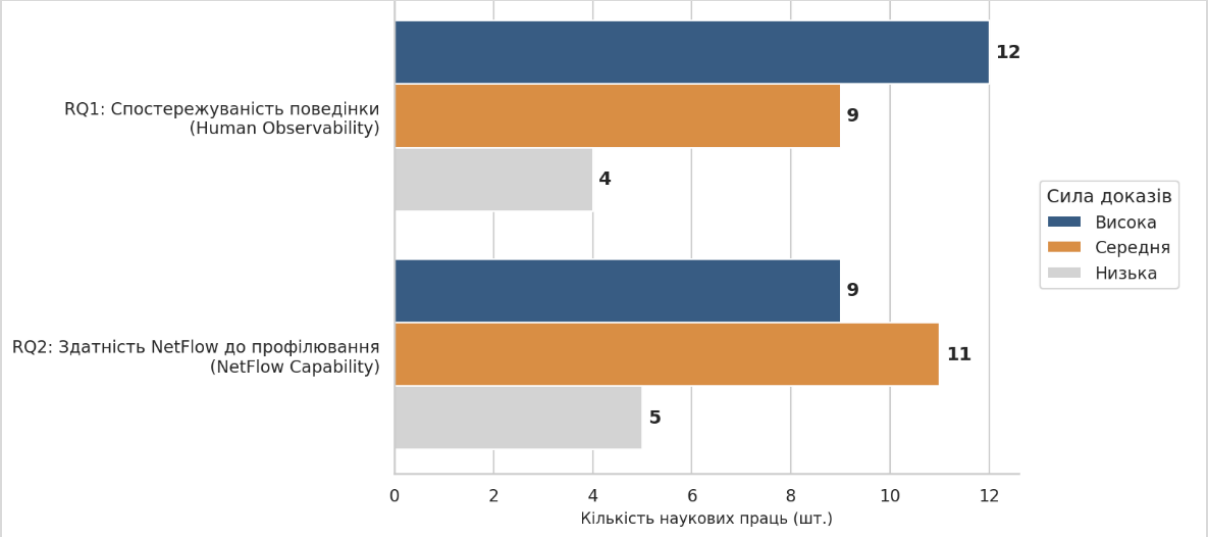


Рисунок 7 – Оцінка сили доказів у мета–аналізі людського фактору проти спроможності NetFlow

З іншого боку, оцінка технологічної спроможності суто безагентних методів на основі NetFlow/IPFIX (RQ2) демонструє більшу стриманість та потребу в додаткових дослідженнях (рис.7). Лише 9 робіт (36%) надали високу силу доказів на користь самодостатності NetFlow для точного профілювання. Більшість досліджень – 11 робіт (44%) з середньою силою та 5 робіт (20%) з низькою – вказують на суттєві обмеження автономного моніторингу потоків.



Рисунок 8 – Прояви дій користувача в цифровому середовищі

5. Обговорення результатів дослідження

Аналіз відібраних наукових джерел засвідчив, що дослідження у сфері поведінкової аналітики користувачів інформаційних систем можна умовно поділити на кілька основних напрямів. Перший напрям пов’язаний із використанням агентних рішень, які встановлюються на кінцевих пристроях і дають змогу збирати детальну інформацію про дії користувача, однак супроводжуються підвищеним навантаженням на систему, ризиками порушення конфіденційності та обмеженнями щодо масштабування.

Другий напрям охоплює застосування мережевої телеметрії, зокрема NetFlow та IPFIX, для аналізу трафіку без втручання в роботу кінцевих станцій. У межах цього підходу дослідники використовують статистичні характеристики потоків даних, такі як джерело і призначення з’єднання, протоколи, обсяг переданої інформації та кількість пакетів, що дозволяє виявляти типові моделі активності користувачів і визначати відхилення від них.

Третій напрям пов’язаний із UEBA/UBA–рішеннями, у яких поведінка користувачів і пристроїв аналізується за допомогою методів машинного навчання та статистич-

ного моделювання. Такі підходи орієнтовані на формування профілів нормальної активності та подальше виявлення аномалій, зокрема внутрішніх загроз, нетипової мережевої активності або ознак компрометації облікових записів.

На рис. 9 представлено класифікацію підходів до аналізу цифрової активності за двома вимірами: локалізацією обробки даних і динамічністю досліджуваних процесів. Горизонтальна вісь відображає перехід від локальних методів, що реалізуються безпосередньо на кінцевому пристрої, до мережевих підходів, які використовують розподілену інфраструктуру або телеметричні дані. Вертикальна вісь характеризує зміну від статичного аналізу, заснованого на заздалегідь сформованих профілях, до динамічного опрацювання потоків даних у режимі, наближеному до реального часу.



Рисунок 9 – Матриця підходів до аналізу цифрової активності за рівнем динамічності та локалізації

Разом із тим аналіз літератури показав, що існує наукова прогалина, пов’язана з недостатнім опрацюванням саме безагентних підходів до профілювання користувачів на основі мережевого трафіку. Більшість наявних рішень або потребує інсталяції додаткового програмного забезпечення на кінцевих пристроях, або не забезпечує достатнього рівня деталізації для побудови поведінкових шаблонів. Це зумовлює актуальність подальших досліджень, спрямованих на поєднання мережевої телеметрії з контекстною інформацією про користувача для підвищення точності виявлення аномальної поведінки.

Підсумовуючи вищезазначене, відповідь на питання, поставлені на початку досліджень, можна сформулювати наступним чином:

✓ **Відповідь на RQ1**: людський фактор може бути інтерпретований із технічних логів як статистично стійка та контекстно зумовлена структура непрямих цифрових слідів. Найсильніші докази у розширеному корпусі стосуються user identity, application

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

interactions і temporal usage profiles. Прямий перехід від цих слідів до наміру або конкретної бізнес–дії користувача залишається недоведеним.

✓ **Відповідь на RQ2**: NetFlow є ефективним і масштабованим безагентним джерелом для нижніх і середніх рівнів поведінкової аналітики – network behaviour, application recognition, user identification та часткового longitudinal user profiling. Однак наявний корпус не доводить, що NetFlow–only є достатнім для реконструкції конкретної робочої або бізнес–дії. Для переходу до higher–level semantics найпереконливіші роботи вже фактично використовують enrichment.

Висновки. Розроблено діаграму еволюції технологій кіберзахисту, яка унаочнює закономірний інженерний перехід індустрії від реактивного периметрово–сигнатурного контролю відомих загроз на локальних хостах та інфраструктурного аналізу телеметрії (NTA, NDR, EDR) до проактивної контекстно–поведінкової оркестрації на базі платформ UEBA та SOAR.

Проведено огляд літератури за період з 2010 по 2025 роки за методологією Kitchenham та Charters щодо потенціалу застосування мережевої телеметрії (NetFlow/IPFIX) для аналізу поведінки користувачів інформаційних систем без втручання в роботу кінцевих пристроїв та порушення їхньої приватності. Наукова репрезентативність та аналітична статистика відібраного масиву публікацій підтверджує високу якість фільтрації матеріалу.

Показано, що агентні підходи забезпечують високий рівень деталізації, однак мають суттєві технічні та організаційні обмеження, пов’язані з навантаженням на кінцеві пристрої, вимогами до ресурсів і ризиками порушення приватності. Натомість використання NetFlow та інших потокових мережевих даних дає змогу реалізувати безагентний підхід до моніторингу, який є менш інвазивним і більш придатним для масштабного застосування в корпоративних мережах.

Результати аналізу виявили фундаментальну технологічну прогалину між двома ключовими дослідницькими питаннями (RQ1 та RQ2). З одного боку, 48% робіт з високою силою доказів підтверджують, що індивідуальні особливості дій людини стабільно та унікально відбиваються в цифрових слідах мережевого трафіку. З іншого боку, лише 36% робіт надають високі докази того, що «чистих» потокових даних NetFlow/IPFIX достатньо для точного профілювання користувачів, що свідчить про те, що комбінований підхід суттєво підвищує точність ідентифікації аномалій і знижує рівень хибнопозитивних спрацьовувань.

Отримані результати обґрунтовують доцільність подальшого розвитку безагентних технологій безпеки. Майбутні дослідження мають концентруватися не на накопиченні сирих мережевих логів, а на вдосконаленні алгоритмів контекстного збагачення NetFlow–даних, глибшій інтеграції з логами аутентифікації в межах систем UEBA/SIEM та побудові адаптивних математичних моделей машинного навчання для виявлення складних внутрішніх загроз у реальному часі.

ЛІТЕРАТУРА / REFERENCES

1. Che cos'è l'analisi del comportamento degli utenti e delle entità (UEBA)? | Microsoft Security. *Microsoft – AI, Cloud, Produktivität, Computing, Gaming und Apps*.
URL: https://www.microsoft.com/it–it/security/business/security–101/what–is–user–entity–behavior–analytics–ueba.

2. Li, B., Springer, J., Bebis, G., & Gunes, M. H. (2013). A survey of network flow applications. *Journal of Network and Computer Applications, 36*(2), 567–581.
https://doi.org/10.1016/j.jnca.2012.12.020.

3. Bakhshandeh A., Eskandari Z. An efficient user identification approach based on Netflow analysis. *2018 15th International ISC (Iranian Society of Cryptology) Conference on Information Security and Cryptology (ISCISC)*, Tehran, 28–29 August 2018. 2018.
URL: https://doi.org/10.1109/iscisc.2018.8546856.

4. Shaman, F., Ghita, B., Clarke, N., & Alruban, A. (2019). User profiling based on application–level using network metadata. *2019 7th International Symposium on Digital Forensics and Security (ISDFS)*, 1–8.
https://doi.org/10.1109/ISDFS.2019.8757503.

5. Clarke N., Li F., Furnell S. A novel privacy preserving user identification approach for network traffic. *Computers & Security*. 2017. Vol. 70. P. 335–350.
URL: https://doi.org/10.1016/j.cose.2017.06.012 (date of access: 24.08.2026).

6. Imbert, C. (2014). *Beyond the cookie: Using network traffic characteristics to enhance confidence in user identity* (GIAC Gold Certification paper). SANS Institute. https://www.giac.org/paper/gcia/9063/cookie–network–traffic–characteristics–enhance–confidence–user–identity/119801.

7. Shaman, F. (2020). *User profiling based on network application traffic monitoring* [Master's thesis, University of Plymouth]. https://doi.org/10.24382/652.

8. Rathore, P., & Kolhe, K. (2025). Implementation of a custom SOAR platform: Architecture, development and evaluation. *International Journal of Environmental Sciences*. https://www.researchgate.net/publication/411330715_Implementation_Of_A_Custom_Soar_Platform_Architecture_Development_And_Evaluation.

9. Hassan, A., Rauf, A., Shafqat, N., Latif, R., & Khan, H. (2025). ZenGuard a machine learning based zero trust framework for context aware threat mitigation using SIEM SOAR and UEBA. *Scientific Reports, 15*, 35871. https://doi.org/10.1038/s41598–025–20998–4.

10. Kitchenham, B., & Charters, S. (2007). Guidelines for performing systematic literature reviews in software engineering (Version 2.3, EBSE Technical Report EBSE–2007–01). Keele University & University of Durham.
https://madeyski.e–informatyka.pl/download/Kitchenham07.pdf.

11. Hofstede, R., Čeleda, P., Trammell, B., Drago, I., Sadre, R., Sperotto, A., & Pras, A. (2014). Flow monitoring explained: From packet capture to data analysis with NetFlow and IPFIX. *IEEE Communications Surveys & Tutorials, 16*(4), 2037–2064.
https://doi.org/10.1109/COMST.2014.2321898.

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

12. Fosić, I., Žagar, D., Grgić, K., & Križanović, V. (2023). Anomaly detection in NetFlow network traffic using supervised machine learning algorithms. *Journal of Industrial Information Integration, 33*, Article 100466. https://doi.org/10.1016/j.jii.2023.100466.
13. Alotibi, G., Clarke, N., Li, F., & Furnell, S. (2016). User profiling from network traffic via novel application–level interactions. In *2016 11th International Conference for Internet Technology and Secured Transactions (ICITST)* (pp. 279–285).
IEEE. https://doi.org/10.1109/ICITST.2016.7856712.
14. Alotibi, G., Clarke, N., Li, F., & Furnell, S. (2016). Identifying users by network traffic metadata. *International Journal of Chaotic Computing, 4*(2), 103–112.
https://doi.org/10.20533/ijcc.2046.3359.2016.0013.
15. Sarmadi, S., Li, M., & Chellappan, S. (2017). On the feasibility of profiling Internet users based on volume and time of usage. In *2017 IEEE Latin–American Conference on Communications (LATINCOM)* (pp. 1–6). IEEE. https://cse.usf.edu/~sriramc/17_Latincom_SLC.pdf
16. Zhao, Y., Chen, J., Wu, D., Teng, J., Sharma, N., Sajjanhar, A., & Blumenstein, M. (2019). Network anomaly detection by using a time–decay closed frequent pattern. *Information, 10*(8), 262. https://doi.org/10.3390/info10080262.
17. Martín, A. G., Fernández–Isabel, A., Martín de Diego, I., & Beltrán, M. (2021). A survey for user behavior analysis based on machine learning techniques: Current models and applications. *Applied Intelligence, 51*(8), 6029–6055. https://doi.org/10.1007/s10489–020–02160–x.
18. Prarthana, T. S., & Gangadhar, N. D. (2017). User behaviour anomaly detection in multi-dimensional data. In *2017 IEEE International Conference on Cloud Computing in Emerging Markets (CCEM)* (pp. 3–10). IEEE. https://doi.org/10.1109/CCEM.2017.19.
19. Kholidy, H. A. (2021). Detecting impersonation attacks in cloud computing environments using a centric user profiling approach. *Future Generation Computer Systems, 117*, 299–320. https://doi.org/10.1016/j.future.2020.12.009.
20. Zhang, X., Marwah, M., Lee, I.–T., Arlitt, M., & Goldwasser, D. (2019). ACE: An anomaly contribution explainer for cyber–security applications. In *2019 IEEE International Conference on Big Data (Big Data)* (pp. 1991–2000).
IEEE. https://doi.org/10.1109/BigData47090.2019.9005989.
21. Shashanka, M., Shen, M.–Y., & Wang, J. (2016). User and entity behavior analytics for enterprise security. In *2016 IEEE International Conference on Big Data (Big Data)* (pp. 1867–1874). IEEE. https://doi.org/10.1109/BigData.2016.7840805.
22. Cui, H., Liu, Y., & Wang, X. (2022). Multi–homed abnormal behavior detection algorithm based on fuzzy particle swarm cluster in UEBA.
https://www.nature.com/articles/s41598–022–26142–
w?utm_source=researchgate.net&utm_medium=article/. Scientific Reports, 12, 22349.
DOI 10.1038/s41598–022–26142–w.
23. Vinupaul, A., & Bhattacharjee, S. (2016). User characterization through network flow analysis. In *2016 IEEE International Conference on Advanced Networks and Telecommunications Systems (ANTS)* (pp. 1–6). IEEE. https://ieeexplore.ieee.org/document/7823965.

24. Krishnamurthy, P., Khorrami, F., Schmidt, S., & Wright, K. (2021). Machine learning for NetFlow anomaly detection with human–readable annotations. *IEEE Transactions on Network and Service Management, 18*(2), 1885–1898.
https://ieeexplore.ieee.org/document/9416281.
25. Liu, S. (2022). User and Entity Behavior Analytics Method Based on Adaptive Mixed–Attribute–Data Density Peaks Clustering. *International Journal of Data Science and Analysis 2022; 8(5): 163–168* http://www.sciencepublishinggroup.com/j/ijdsa
doi: 10.11648/j.ijdsa.20220805.17 ISSN: 2575–1883 (Print); ISSN: 2575–1891 (Online)
26. Melnikov, D., & Schönwälder, M. (2010). Cybermetrics: User identification through network flow analysis. In *Proceedings of the 11th International Conference on Passive and Active Network Measurement (PAM 2010)* (pp. 234–243). Springer.
https://doi.org/10.1007/978–3–642–13986–4_24.
27. Perelman, V., Melnikov, D., & Schönwälder, J. (2011). Flow signatures of popular applications. In *2011 IEEE Network Operations and Management Symposium (NOMS)* (pp. 1068–1071). IEEE. https://doi.org/10.1109/INM.2011.5990668.
28. Piskozub, M. (2020). *Network traffic behaviour profiling* [Doctoral dissertation, University of Oxford]. ORA. https://ora.ox.ac.uk/objects/uuid:f40bf0ff–9adb–4d4f–9f87–2bd69eaf6be4.
29. Aquino, J., & Villanueva, R. (2023). Network anomaly detection using NetFlow and network automation. *International Journal of Computer Networks and Communications, 15*(3), 1–14. https://doi.org/10.1109/ISDFS58141.2023.10131880.

Received 01.06.2026.
Accepted 03.08.2026.
Published 30.10.2026

Current state and trends in agentless profiling of user behavior based on network telemetry

The relevance of this research topic stems from the rapid pace of digitalization and the need to detect anomalous or malicious activity in corporate networks without violating user privacy or overloading the IT infrastructure. The problem lies in the significant methodological fragmentation of existing approaches to behavioural analysis, as well as in the strict technical and organizational limitations of traditional agent–based endpoint monitoring tools.

The objective of the study is to systematize and comprehensively analyse the scientific and technical body of work in the field of agentless modelling of digital activity to determine the current state, technological capabilities, and limitations of using network telemetry (in particular, NetFlow/IPFIX) for user profiling and the detection of anomalous behaviour without infringing on user privacy.

The research methods are based on a simplified systematic literature review (SLR) covering the period from 2010 to 2025, in accordance with the methodological recommendations of B. Kitchenham and S. Charters in the field of computer engineering. Following a step–by–step screening procedure and expert evaluation, a representative pool of 25 key works was selected from the initially identified 121 publications for quantitative and qualitative meta–analysis.

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

Results and Key Findings. An original diagram of the evolution of cybersecurity technologies was developed, demonstrating a logical engineering transition from reactive, signature–based detection of known threats on hosts to proactive, context– and behaviour–based orchestration based on integrated UEBA and SOAR platforms. It has been demonstrated that the use of NetFlow/IPFIX provides non–invasive, agentless monitoring suitable for large–scale deployment in heterogeneous corporate networks. At the same time, the analysis revealed a technological gap between the high observability of the human factor in network traces (confirmed with high evidential strength in 48% of the studies) and the low capability of purely autonomous NetFlow stream data to accurately identify users (only 36% of the studies with high evidential strength). The study substantiates the critical need to transition to combined solutions based on contextual enrichment of NetFlow data with authentication logs, DNS queries, and system logs within UEBA/SIEM platforms using adaptive machine learning.

Keywords: cybersecurity, NetFlow, IPFIX, user behaviour, user identity, UEBA, behavioural quantification.

Супрун Олександр Вікторович – аспірант кафедри електронних обчислювальних машин Дніпровського національного університету імені Олеся Гончара.
ORCID: https://orcid.org/0009–0000–0265–5834.
Карпенко Надія Валеріївна – кандидат фізико–математичних наук, доцент, доцент кафедри електронних обчислювальних машин Дніпровського національного університету імені Олеся Гончара.
ORCID: https://orcid.org/0000–0003–4700–6357.

Suprun Oleksandr – postgraduate student at the department of electronic computing machines, Oles Honchar Dnipro national university.
ORCID: https://orcid.org/0009–0000–0265–5834.
Karpenko Nadiia – Ph.D. in physics and mathematics, associate professor, ass. professor in the department of electronic computing machines Oles Honchar Dnipro national university.
ORCID: https://orcid.org/0000–0003–4700–6357.