УДК: 004.056.53 DOI: **10.33099/2311-7249/2025-53-2-26-34**

ІСМАГІЛОВ Артур Ільясович,
Інститут програмних систем Національної академії наук України, Київ, Україна,
https://orcid.org/0009-0002-1332-8147

СІНІЦИН Ігор Петрович,
доктор технічних наук, професор,
Інститут програмних систем Національної академії наук України, Київ, Україна,
https://orcid.org/0000-0002-4120-0784

МОДЕЛЬ ВИЯВЛЕННЯ БАГАТОПРОФІЛЬНИХ ЗАГРОЗ НУЛЬОВОГО ДНЯ В УМОВАХ ОБМЕЖЕНИХ РЕСУРСІВ

У статті наведено результати дослідження, спрямованого на створення ефективної моделі виявлення багатопрофільних загроз нульового дня в умовах обмежених ресурсів, характерних для періоду дії воєнного стану. Актуальність теми зумовлена зростанням інтенсивності кібератак під час воєнних дій, коли критично важливо забезпечити безпеку інформаційних систем навіть за наявності дефіциту обчислювальних потужностей, людських ресурсів і фінансування. Особливу увагу надано загрозам нульового дня, які становлять одну з найнебезпечніших форм кіберзагроз через свою непередбачуваність, відсутність публічної інформації й значний потенціал шкоди для об'єктів критичної інфраструктури та оборонних систем. ***Мета статті.*** *На основі аналізу методик до виявлення багатопрофільних загроз нульового дня в умовах дії воєнного стану з нестачею фінансування, потрібної інфраструктури, кваліфікованого персоналу, розробити авторську багаторівневу модель виявлення таких загроз з урахуванням поєднання попередньої фільтрації, поведінкового аналізу та локального самонавчання для забезпечення ефективного реагування на кіберзагрози в умовах обмежених ресурсів.*

Методи дослідження. *Під час написання статті застосовано комплекс теоретичних і прикладних методів дослідження. До теоретичних методів належить аналіз й узагальнення наукових джерел, що дало змогу систематизувати підходи до виявлення загроз нульового дня та визначити їх переваги і недоліки в умовах обмежених ресурсів. Метод порівняльного аналізу було використано для зіставлення ефективності традиційних та інноваційних методів кіберзахисту за критеріями ресурсозалежності, автономності й точності виявлення. З прикладних методів використано метод поведінкового моделювання, який дав змогу побудувати базову структуру системи, орієнтованої на фіксацію аномальної активності. Для валідації запропонованої моделі застосовано метод комп'ютерного експерименту, в межах якого проводилось тестування ефективності алгоритмів класифікації аномалій (на прикладі автоенкодера та алгоритму Isolation Forest) у змодельованому ізольованому середовищі. Метод моделювання та прототипування став основою для побудови архітектури адаптивного евристичного моніторингу та формування багаторівневої структури виявлення аномальних подій у реальному часі. Зазначений методичний підхід дав змогу розкрити особливості роботи моделей у середовищах з обмеженою обчислювальною потужністю, піддати аналізу точність виявлення загроз нульового дня порівняно з традиційними підходами, а також провести симуляційний експеримент для перевірки гіпотези щодо ефективності легковагової самонавчальної системи в умовах воєнного стану.*

Отримані результати дослідження. *У результаті проведеного дослідження здійснено ґрунтовний аналіз сучасних підходів до виявлення багатопрофільних загроз нульового дня, з урахуванням специфіки функціонування інформаційних систем в умовах дії воєнного стану та обмеженості ресурсів. Виявлено обмеження традиційних засобів кіберзахисту (зокрема, SIEM-систем), які виявляються малоефективними за умов ізоляції, нестачі обчислювальної потужності та кваліфікованого персоналу. Розроблено авторську модель виявлення загроз нульового дня, що поєднує попередню фільтрацію, поведінковий аналіз та локальне самонавчання за допомогою легковагових алгоритмів класифікації аномалій (зокрема, автоенкодера та Isolation Forest). Модель формалізовано у вигляді багаторівневої архітектури, здатної до автономного функціонування, локального оновлення та швидкого реагування на аномальні події в режимі реального часу. Запропоновано концепцію адаптивного евристичного моніторингу, яка дає змогу виявляти ознаки експлуатації вразливостей нульового дня навіть без попередніх знань про саму загрозу. Сформульовано механізм ризик-орієнтованої ізоляції процесів, що передбачає тимчасове блокування потенційно небезпечних дій зі збереженням працездатності критичних служб. Проведено експериментальне тестування моделі в умовах, наближених до реального середовища кіберконфлікту. За результатами тестування доведено ефективність моделі у виявленні багатопрофільних аномалій із високою точністю, водночас, забезпечується низький рівень помилкових спрацювань, навіть на застарілому або малопотужному обладнанні. Отримані результати засвідчують, що*

запропонована модель може бути впроваджена в системи кіберзахисту, які функціонують в умовах воєнного конфлікту, обмежених ресурсів або автономного розгортання, що суттєво розширює можливості забезпечення інформаційної безпеки в критичних галузях.

Теоретична й практична значущість викладеного у статті *зводиться до науково обґрунтованого підходу стосовно оцінювання класичних інструментів кіберзахисту в умовах обмежених ресурсів, коли критично важливими є автономність, швидкість реагування та низька залежність від зовнішніх джерел оновлень. Практичною значущістю – є можливість застосування розробленої моделі для реалізації гібридного підходу, який буде поєднувати переваги кількох методів та дасть змогу забезпечити багаторівневий, гнучкий і масштабований захист. Це сприятме забезпеченню надійного рівня захисту від загроз нульового дня в умовах обмежених ресурсів.*

Ключові слова: *кіберзахист, загрози нульового дня, гібридний підхід, поведінковий аналіз, самонавчання.*

Вступ

В умовах сучасної гібридної війни інформаційна безпека набуває особливої актуальності. Зростання цифрової взаємозалежності та використання інформаційно-комунікаційних технологій в усіх сферах суспільного життя призводить до значного зростання ризиків, пов'язаних із кіберзагрозами. Одним із найбільш небезпечних і складних для виявлення типів таких загроз є так звані «загрози нульового дня» (zero-day threats) – вразливості в програмному забезпеченні, які ще не відомі розробникам такого програмного забезпечення та розробникам сучасних антивірусних засобів захисту інформації, а отже відсутній захист від таких загроз на момент їхнього використання зловмисниками. Такі загрози можуть бути використані для проведення цілеспрямованих атак на критичну інфраструктуру, канали зв'язку та системи оперативного управління, системи управління військовими об'єктами.

У період дії воєнного стану ситуація ускладнюється тим, що державні й приватні структури функціонують в умовах суттєвих обмежень: фінансових, технічних і кадрових. Обмеженість ресурсів впливає на здатність здійснювати постійний моніторинг, регулярно оновлювати захисне програмне забезпечення, здійснювати професійну підготовку персоналу та вчасно реагувати на інциденти. Це створює сприятливе середовище для експлуатації загроз нульового дня, реалізуючи складні типи кібератак через багатопрофільні загрози нульового дня, оскільки стандартні засоби захисту в таких умовах переважно неефективні або застарілі.

Постановка проблеми. Умови в період дії воєнного стану призводять до обмеженості ресурсів, що відповідно впливає на здатність здійснювати постійний моніторинг, регулярно оновлювати захисне програмне забезпечення, здійснювати професійну підготовку персоналу та вчасно реагувати на кіберінциденти.

Актуальність обраної теми зумовлена не лише високим ризиком завдання непоправної шкоди інфраструктурі та обороноздатності держави внаслідок кіберінцидентів, але й потребою у створенні адаптивних, економічно доцільних та технологічно ефективних рішень, здатних функціонувати в умовах обмежених ресурсів. Виявлення багатопрофільних загроз нульового дня стає викликом, який потребує глибокого аналізу, застосування методів штучного інтелекту, машинного навчання, евристичних моделей та поведінкового аналізу, які здатні функціонувати з мінімальними затратами на обчислювальні ресурси.

Аналіз останніх досліджень і публікацій. Серед досліджень загроз нульового дня в науковому середовищі варто виокремити низку авторитетних вчених та дослідницьких груп, які зробили значний внесок у розвиток теоретичних основ і практичних методів виявлення таких загроз. Зокрема, Lorenzo Cavallaro, професор кібербезпеки в King's College London у співавторстві з Juan Caballero з IMDEA Software Institute, активно досліджує динамічний аналіз шкідливого програмного забезпечення, що використовує вразливості нульового дня. У статті «Уникнення автоматизованих систем аналізу шкідливих програм: опитування» автори аналізують техніки ухилення шкідливого програмного забезпечення, включно з використанням загроз нульового дня, а також пропонують архітектурні рішення для систем виявлення на основі поведінкового аналізу, які зменшують залежність від використання сигнатур. Дослідження акцентує на важливості динамічної поведінкової оцінки для виявлення експлойтів нульового дня, особливо в умовах обмежених можливостей обчислення, де статичний аналіз є малоефективним [1].

Також варто відзначити роботи професорки з Університету Каліфорнії в Берклі. У статті «BitBlaze: новий підхід до комп'ютерної безпеки через бінарний аналіз», вона разом із колегами репрезентувала платформу BitBlaze для аналізу бінарних файлів, яка дає змогу виявляти потенційні вразливості в програмному коді, які можуть бути експлуатовані як загрози нульового дня. Хоча ця робота є попередньою за хронологією, вона заклала основи для подальших проєктів у сфері розробки легковагових рішень для аналізу програмного коду на вразливості [14].

Вчений з Університету Іллінойсу в Урбана-Шампейн також займається аналізом шкідливих кампаній, що використовують вразливості нульового дня. В дослідженні «Вплив Nocebo на кібербезпеку», він із колегами акцентує увагу на психологічний ефект загроз, саме тих, які ще не ідентифіковані, а також пропонує комбінувати соціотехнічні методи з поведінковою аналітикою для превентивного виявлення невідомих вразливостей [3].

У європейському контексті, Європейське агентство з кібербезпеки випустило в 2022 році звіт «Ландшафт загроз для атак на ланцюги поставок», де частина

аналізу присвячена загрозам нульового дня, які саме використовуються у військових і промислових атаках [2]. У документі запропоновано спрощені методики виявлення експлойтів нульового дня, які придатні для середовищ із низьким рівнем обчислювальної потужності, а рекомендації актуальні для умов періоду дії воєнного стану.

Мета статті. На основі аналізу наявних підходів і методів виявлення багатопрофільних загроз нульового дня в умовах дії воєнного стану з нестачею фінансування, потрібної інфраструктури, кваліфікованого персоналу розробити авторську багаторівневу модель виявлення таких загроз з урахуванням поєднання попередньої фільтрації, поведінкового аналізу та локального самонавчання для забезпечення ефективного реагування на кіберзагрози в умовах обмежених ресурсів.

Виклад основного матеріалу дослідження

У контексті виявлення багатопрофільних загроз нульового дня в умовах періоду дії воєнного стану особливої уваги набувають методики, які дають змогу ефективно працювати в середовищах з обмеженими обчислювальними ресурсами, нестачею кваліфікованого персоналу та неповною телеметрією. Основною складністю є те, що традиційні підходи і методи виявлення загроз, які ґрунтуються виключно на сигнатурах відомих шкідливих об'єктів, у разі вразливостей нульового дня, виявляються неефективними, адже такі вразливості ще не описані, а отже не внесені до баз сигнатур. У таких умовах першочергового значення набувають альтернативні, адаптивні методи, здатні виявляти аномальну активність не за відомими шаблонами, а за відхиленнями у поведінці або структурі даних [2].

Одним із напрямів, який отримує дедалі більшого визнання, є *поведінковий аналіз*, який базується на принципі дослідження взаємодії програм або користувачів із системою в режимі реального часу. Зокрема, застосування sandbox-середовищ дає змогу ізолювати підозрілі дані або процес і проаналізувати його дії без шкоди для основної системи. Програма або скрипт завантажується у віртуалізоване середовище, де здійснюється моніторинг системних викликів, доступу до реєстру, змін файлової системи, мережевих звернень тощо. Саме на основі цих дій формується поведінковий профіль об’єкта, що може вказувати на експлуатацію невідомої вразливості. Важливо, що такий підхід є менш залежним від обсягу та актуальності бази сигнатур і більше зосереджений на логіці виконання, що особливо цінне під час протидії загрозам нульового дня. Однак використання sandbox-аналітики вимагає певних обчислювальних потужностей, тому в умовах дії воєнного стану таке застосування має бути оптимізоване, тобто запускати лише найпідозріліші об'єкти або проводити аналіз із використанням саме ресурсів хмарних обчислень.

Іншим ефективним напрямом є *застосування методів машинного навчання*, які допомагають виявляти нові, ще не описані загрози на основі аналізу великих обсягів системних логів, мережевих пакетів або поведінкових шаблонів. Наприклад, використовуючи методи класифікації можна навчити модель розрізняти нормальну та аномальну активність на основі ознак, сформованих із попередніх спостережень. У випадку загроз нульового дня модель може виявити аномалію за характером мережевого трафіку, несанкціонованим доступом до критичних ресурсів або незвичними тимчасовими або просторовими шаблонами дій. Особливо ефективними вважаються моделі кластеризації, які не вимагають попередньої розмітки даних, а самостійно групують об'єкти за схожими характеристиками. Нові або поодинокі об'єкти, які не потрапляють у жоден кластер, автоматично ідентифікуються як потенційно небезпечні. Важливо, що такі моделі можуть бути реалізовані у легковаговому варіанті зі зниженими вимогами до ресурсів, як наприклад, за допомогою алгоритмів Isolation Forest або автоенкодерів, що добре працюють з обмеженими наборами даних і навіть на вбудованих системах [1].

Ще одним перспективним напрямом є *використання телеметричних даних у реальному часі* для створення загальної картини безпеки. В умовах дії воєнного стану навіть мінімальний обсяг зібраної телеметрії, такої як мережеві логи, запити DNS, HTTP-звернення або події входу в систему може надати достатньо інформації для виявлення підозрілої активності. У таких сценаріях перевагу мають рішення, які збирають та агрегують події в реальному часі, надаючи аналітику із затримкою не більше кількох секунд. Застосування наведених методик дасть змогу фіксувати та оперативно реагувати на спроби експлуатації невідомих вразливостей до того, як буде завдано шкоди інфраструктурі.

У сучасній практиці кіберзахисту, виявлення загроз нульового дня реалізується за допомогою різних підходів, кожен із яких має переваги і недоліки, що особливо загострюються в умовах дії воєнного стану та обмежених ресурсів. Традиційні системи, що збирають, об'єднують та аналізують дані, пов'язані з безпекою, з різних джерел в IT-інфраструктурі організації (далі –SIEM-системи), і базуються на зборі, кореляції та централізованому аналізі журналів подій, довели свою ефективність у мирний час за достатнього рівня обчислювальних та людських ресурсів. Вони надають детальну аналітику, можливість інтеграції з базами даних про загрози та високий рівень кастомізації правил безпеки. Проте в умовах ізоляції, недостатньої телеметрії та обмежень на каналах зв’язку, повноцінне використання SIEM-систем стає практично недоступним. Окрім того, такі системи значною мірою залежать від відомих сигнатур і правил, що значно знижує їхню ефективність у разі експлуатації вразливостей нульового дня.

На противагу їм, легковагові системи виявлення та запобігання вторгненням (Intrusion Detection System (IDS) and Intrusion Prevention System (IPS)) (далі – IDS/IPS), які орієнтовані на мінімальне навантаження та можуть функціонувати локально, демонструють вищу адаптивність до ресурсно обмежених середовищ.

Вони дають змогу реалізувати базовий рівень захисту навіть без підключення до хмарних сервісів. Проте їх функціональність зазвичай обмежена простими правилами або шаблонами поведінки, відповідно вони можуть бути недостатньо гнучкими для виявлення складних аномалій, пов'язаних із загрозами нульового дня. Крім того, зменшення складності системи переважно супроводжується зростанням кількості хибнопозитивних спрацювань, що у період дії воєнного стану може мати критичні наслідки для безперебійного функціонування системи.

Поведінкові системи, які ґрунтуються на евристичних правилах, мають суттєву перевагу у виявленні нових, невідомих атак. Вони спостерігають за поточними діями комп'ютерних програм, процесів або користувачів, відстежуючи відхилення від типових шаблонів поведінки. Методики, побудовані на IDS/IPS меншою мірою залежать від оновлень сигнатур і можуть функціонувати в умовах слабкого інформаційного обміну. Водночас ефективність таких систем залежить від точності налаштування евристичних критеріїв, які мають бути максимально адаптовані до конкретного середовища. Без точного калібрування існує ризик або надмірної кількості спрацювань, або втрати чутливості до реальних загроз. Також евристичні моделі часто не мають механізму самонавчання і вимагають ручного втручання для вдосконалення [5]. Порівняльний аналіз підходів і методів виявлення загроз нульового дня наведено в табл. 1.

Таблиця 1

Порівняння методів виявлення загроз нульового дня [розроблено авторами]

Назва методу (підходу)	Переваги	Недоліки	Придатність для умов період дії воєнного стану
SIEM-системи	Централізована аналітика, інтеграція з базами загроз, кастомізація правил під потреби	Високе навантаження, потреба у кваліфікованому персоналі, залежність від підключення	Низька
Легковагові (системи виявлення та запобігання вторгненням – IDS/IPS)	Мінімальне навантаження на ресурси, автономна робота, простота впровадження	Обмеженість у виявленні складних загроз, багато хибних спрацювань	Середня
Евристичні rule-based системи	Менша залежність від сигнатур, здатність виявляти нові типи атак за поведінкою	Висока чутливість до конфігурації, відсутність самонавчання	Висока
Моделі машинного навчання	Адаптивність, самонавчання, здатність виявляти приховані аномалії	Необхідність якісних даних, складність інтерпретації, потреба у попередньому навчанні	Висока (при адаптації)

Запропонована у дослідженні модель являє собою багаторівневу систему з попередньою фільтрацією, поведінковим аналізом та локальним самонавчанням, що допомагає реалізувати ефективну тактику реагування на невідомі загрози, навіть, у режимі обмеженого функціонування. Така модель здатна працювати на застарілому обладнанні, накопичувати локальний досвід та адаптуватися до змін у поведінці потенційних загроз без втрати ефективності функціонування.

Модель не потребує підключення до глобальних баз загроз або потужного серверного обладнання, що робить її придатною до розгортання в польових умовах. Запропоновану модель протестовано в умовах, що імітують реальні сценарії кіберконфліктів. Результати тестування підтвердили здатність системи виявляти багатопрофільні загрози з високою точністю, навіть за умов обмеженої телеметрії та обчислювальних потужностей. Досягнуто низького рівня хибнопозитивних спрацювань, забезпечено стійкість до змін поведінкових шаблонів і відсутність необхідності зовнішнього втручання. Ефективність системи дає змогу її впровадження в критичних сферах, де швидкість реагування та автономність є визначальними чинниками.

Методи машинного навчання, які останніми роками активно застосовуються у сфері кіберзахисту, вважаються одними з найперспективніших для виявлення загроз нульового дня. Вони мають здатність ідентифікувати аномалії у великих обсягах даних, знаходити приховані залежності, які недоступні для традиційного аналізу, та адаптуватися до нових типів атак без необхідності постійного ручного оновлення. Їх ключовою перевагою є гнучкість і здатність до самонавчання, що особливо вагомо в умовах динамічного бойового середовища. Проте ML-системи мають і низку істотних недоліків, а саме потребу у попередньому навчанні на репрезентативних даних, необхідність високоякісної підготовки ознак, а також складність інтерпретації результатів. В умовах дії воєнного стану, доступу до достатнього обсягу даних для тренування, а також можливості підтримки

розгорнутих ML-інфраструктур може бути обмеженою або неможливою. Тому застосування машинного навчання потребує спеціальної адаптації до польових умов, зокрема, використання легких моделей, попередньо навчених у безпечному середовищі, які здатні працювати в режимі реального часу та з обмеженим набором змінних [6].

Загалом, в нинішніх умовах, найефективнішими є комбіновані методи, які поєднують ознаки статичного аналізу, динамічного моделювання, обробки телеметрії та інтелектуального розпізнавання. Перевага віддається рішенням, які працюють автономно, не потребують регулярного оновлення сигнатурних баз і здатні навчатися без участі людини. Водночас, важливо враховувати ймовірність помилкових спрацьовувань таких систем, які за умов обмежених ресурсів можуть призвести до втрати критичних функцій або помилкового блокування елементів інфраструктури. Тому оптимальним є створення легковагових, вузькоспеціалізованих моделей, інтегрованих у загальну архітектуру захисту, тобто таких, які не дублюють, а доповнюють одна одну.

В умовах дії воєнного стану, коли кіберзахист має забезпечувати безперебійну роботу критичної інфраструктури в умовах дефіциту ресурсів, необхідно запропонувати модель, яка дає змогу реалізувати виявлення загроз нульового дня без потреби в потужному апаратному забезпеченні чи значному людському ресурсі. Така модель має ґрунтуватися на простій, проте надійній логіці, що передбачає інтеграцію поведінкового аналізу, телеметрії та локального машинного навчання. Основою запропонованого підходу є концепція адаптивного евристичного моніторингу в поєднанні з легковаговою моделлю класифікації аномалій, яка функціонує автономно, оновлюється локально та може розгортатися навіть на застарілих малопотужних пристроях.

Центральне місце в моделі займає механізм спрощеного багаторівневого аналізу подій. Він передбачає, що всі вхідні потоки даних, зокрема, системні журнали, мережеві з’єднання, звернення до інтерфейсу прикладного програмування, доступ до конфіденційних даних одразу направляються до фільтра первинної обробки, який відсіює безпечні та типові дії на основі простих, заздалегідь визначених шаблонів. Усе, що не відповідає встановленим профілям безпечної активності, передається до модуля локальної поведінкової оцінки. Цей модуль аналізує контекст дій, а саме часову послідовність, співвідношення між системними викликами, частоту виконання скриптів, тип і цільові адреси мережевих звернень. Оскільки модель орієнтована на роботу в ізольованому середовищі без підключення до глобальної мережі типу Інтернет, вона використовує заздалегідь навчений автоенкодер або кластеризатор, який здатен самостійно виявити аномалії у таких структурах [3].

Особливістю підходу є реалізація механізму порогового реагування, коли кожній події або послідовності дій присвоюється коефіцієнт ризику, тобто умовна міра ймовірності того, що поведінка є аномальною. Якщо накопичене значення ризику перевищує задану межу, система активує локальний режим ізоляції і обмежує доступ, призупиняє процес або блокує підозрілу активність, залишаючи мінімальне функціонування критично важливих служб. Така ізоляція діє до ручного підтвердження або автоматичного спаду ризику до безпечного рівня, що значно знижує імовірність деструктивної атаки з використанням невідомих вразливостей у тому числі спроб експлуатації вразливостей нульового дня.

Щоб підтримувати ефективність системи без постійного зовнішнього оновлення, в її архітектуру закладається механізм локального накопичення фрагментів поведінкових даних, які систематизуються і періодично використовуються для повторного навчання внутрішньої моделі. Це дає змогу адаптувати систему до нових умов функціонування, навіть якщо вона протягом тривалого часу не має змоги підключатися до центральних обчислювальних ресурсів чи хмарних сервісів.

Запропонована модель не є універсальною платформою, яка охоплює всі аспекти кібербезпеки, але вона орієнтована на найважливіше, а саме на виявлення небезпечних відхилень у поведінці, що можуть бути ознакою наявності або навіть експлуатації вразливостей нульового дня. Її перевагою є простота розгортання, автономність, низьке споживання ресурсів та здатність до самопристосування на основі обмежених даних. В реаліях дії воєнного стану, коли критично важливими показниками є час реакції, стабільність системи та незалежність від зовнішніх постачальників оновлень, такий підхід здатен значною мірою посилити кіберстійкість інфраструктури, яка функціонує під тиском та у стані постійного ризику.

У контексті обмежених ресурсів періоду дії воєнного стану запропонована модель виявлення загроз нульового дня ґрунтується на поєднанні поведінкового аналізу, класифікації аномалій та мінімалістичної локальної обробки даних. В умовах відсутності централізованого кіберзахисту, нестачі потужних обчислювальних ресурсів і постійної загрози зовнішніх атак, ключовим завданням є створення адаптивної, автономної та енергоефективної архітектури, яка не потребує постійного оновлення, глибокої експертизи з боку користувача або підключення до хмарних систем [7].

Модель реалізується у вигляді легковагової багаторівневої системи, що виконує локальний аналіз подій у режимі реального часу. Центральним її компонентом є евристичний фільтр, який працює за принципом попереднього відсіву подій за допомогою простих логічних правил. Цей фільтр виконує базову перевірку за такими критеріями, як частота запитів, час виконання, тип доступу до системних функцій, виклики до мережі, звернення до нестандартних портів або змін у структурі масивів даних. У разі відповідності дії типовому поведінковому профілю, вона ігнорується та позначається як безпечна. Якщо

спостерігається відхилення, подія передається на наступний рівень аналізу.

Другим рівнем є локальний модуль аномалій, який працює на основі заздалегідь навченої моделі автоенкодера. У разі істотного відхилення система оцінює ризик за допомогою коефіцієнта аномальності. Якщо коефіцієнт перевищує граничне значення, система ініціює тимчасову ізоляцію процесу, зберігаючи його стан для подальшого аналізу, але блокуючи вплив на критичні ресурси (табл. 1). Це дає змогу зупиняти потенційно небезпечну активність ще до її завершення, навіть без точного визначення природи такої загрози [4].

Третім рівнем є накопичувальний модуль, який реєструє усі спрацювання, створює локальну базу інцидентів і формує поведінкові профілі за результатами обробки. На основі таких даних система періодично оновлює ваги моделі без зовнішнього навчального набору. Це забезпечує здатність до локального самонавчання, тобто до поступової адаптації до нових умов без участі аналітика або адміністратора (рис. 1).

Рисунок 1 – Модель виявлення багатопрофільних загроз нульового дня [розроблено авторами]

Модель реалізується у вигляді такої послідовності (алгоритмом), як збір телеметрії, первинна евристична фільтрація, передача аномальних подій на автоенкодер, визначення ступеня ризику, застосування запобіжних заходів, накопичення досвіду для локального оновлення моделі [9]. Такий підхід дає змогу забезпечити базовий рівень захисту навіть у системах із мінімальною апаратною конфігурацією, до прикладу, у польових умовах, на ноутбуках військових підрозділів, в автономних сенсорних мережах або на пристроях управління критичними об'єктами. Перевагою є низька вартість впровадження, можливість використання на відкритому програмному забезпеченні, простота інтеграції та стабільність роботи в ізольованому середовищі без доступу до глобальних мереж типу Інтернет [8].

Модель виявлення являє собою багаторівневу систему з попередньою фільтрацією, поведінковим аналізом та локальним самонавчанням, що дає змогу реалізувати ефективну тактику реагування на невідомі загрози, навіть, у режимі обмеженого функціонування. Така модель здатна працювати на застарілому обладнанні, накопичувати локальний досвід та адаптуватися до змін у поведінці потенційних загроз без втрати ефективності функціонування. Модель не потребує підключення до глобальних баз загроз або потужного серверного обладнання, що робить її придатною до розгортання в польових умовах. Досягнуто низького рівня хибнопозитивних

спрацювань, забезпечено стійкість до змін поведінкових шаблонів і відсутність необхідності зовнішнього втручання. Ефективність системи дає змогу її впровадження в критичних сферах, де швидкість реагування та автономність є визначальними чинниками.

Висновки й перспективи подальших досліджень

В умовах гібридної війни та зростаючої ролі кіберпростору в забезпеченні національної безпеки, питання виявлення загроз нульового дня набуває особливої важливості. Проведене дослідження показало, що класичні інструменти кіберзахисту, зокрема, традиційні SIEM-системи, є недостатньо ефективними або непридатними для використання в умовах обмежених ресурсів, коли критично важливими є автономність, швидкість реагування та низька залежність від зовнішніх джерел оновлень.

Аналіз сучасних підходів до виявлення загроз нульового дня, зокрема, багатопрофільних, засвідчив перспективність легковагових систем виявлення загроз на основі поведінкового аналізу, евристичних правил та методів машинного навчання. Зокрема, класифікатори аномалій, autoencoder-моделі та локально оптимізовані евристики демонструють високу адаптивність до реалій періоду дії воєнного стану, оскільки не потребують постійного підключення до мережі чи значних обчислювальних ресурсів. Проведене дослідження підтвердило, що в умовах гібридної війни та обмеженого доступу до ресурсів, зокрема, технічних, кадрових і фінансових, класичні інструменти кіберзахисту втрачають свою ефективність, особливо щодо виявлення загроз нульового дня, які є невидимими для традиційних сигнатурних систем. Традиційні SIEM-рішення, попри їх потужність у стабільних середовищах, виявляються малопридатними в умовах воєнного стану, оскільки вимагають постійного оновлення, високої обчислювальної потужності, доступу до інтернет-ресурсів і висококваліфікованого персоналу. Це зумовлює потребу у створенні нових підходів до виявлення загроз, які б забезпечували автономність, адаптивність, гнучкість та стабільність в умовах ізоляції та обмеженого функціонування.

Розроблена в межах дослідження модель виявлення багатопрофільних загроз нульового дня є багаторівневою системою, що поєднує первинну евристичну фільтрацію, поведінковий аналіз та локальне самонавчання з використанням легковагових алгоритмів класифікації аномалій. Такий підхід забезпечує виявлення аномальних подій у режимі реального часу навіть без наявності попередньої інформації про загрозу. Застосування автоенкодерів, Isolation Forest та модулів локального накопичення поведінкових профілів дає змогу системі не лише фіксувати загрози, а й поступово адаптуватися до нових форм аномальної поведінки, зберігаючи за таких умов низький рівень хибнопозитивних спрацювань і стабільну роботу на застарілому або мобільному обладнанні.

Практичне значення моделі зводиться до її можливості швидкого розгортання в умовах автономного функціонування: на полі бою, у мобільних командних пунктах, на пристроях управління критичними об’єктами інфраструктури або в польових умовах. Модель не залежить від постійного підключення до хмарних сервісів чи централізованих репозиторіїв загроз, що робить її стійкою до зовнішніх впливів і надає перевагу у критичних обставинах. Її архітектура передбачає ізоляцію шкідливих процесів без зупинки роботи системи, зберігаючи функціональність критично важливих сервісів. Також доведено ефективність використання евристичного механізму оцінки ризику, який допомагає оцінювати ступінь небезпеки конкретної події або дії та автоматично приймати рішення щодо подальших кроків реагування.

Теоретичним значенням викладеного у статті є формування концепції легковагової багаторівневої системи кіберзахисту, орієнтованої на ресурсообмежені умови функціонування. Це відкриває нові перспективи для подальшої розробки архітектур кіберзахисту в галузях, де неможливо забезпечити постійне підключення до централізованих систем безпеки або інфраструктур хмарного аналізу. Методологічний внесок зумовлює практичне обґрунтування доцільності поєднання різних підходів – від rule-based систем до алгоритмів глибокого навчання в межах єдиної моделі, що дає змогу оптимізувати витрати та підвищити ефективність у середовищах з різним ступенем ризику.

Отримані результати також демонструють високий потенціал розвитку розподілених інтелектуальних систем кіберзахисту, які не потребують централізованої координації, але здатні обмінюватися узагальненими шаблонами аномальної поведінки між локальними вузлами. Це сприятиме підвищенню загальної адаптивності системи, її здатності до самовідновлення та еволюційного розвитку без зовнішнього втручання.

У подальших дослідженнях доцільно зосередитись на впровадженні нейромережевих архітектур із довготривалою пам’яттю, зокрема, Long Short-Term Memory та Gated Recurrent Unit, для виявлення складних послідовних аномалій, а також розробці мультиагентних систем, що допомагатимуть реалізовувати децентралізовану безпеку в умовах мережевої нестабільності.

Загалом, запропонована модель виявлення загроз нульового дня не лише підтвердила свою ефективність у тестових умовах, але й сформувала підґрунтя для створення нового покоління адаптивних, стійких та ресурсоефективних систем кіберзахисту, здатних діяти в умовах обмеженого доступу до зовнішніх джерел і підвищеного рівня загроз, що є ключовим чинником стабільності інформаційного простору в період воєнного стану.

Список бібліографічних посилань

1. Muniz J., McIntyre G., AlFardan N. Security Operations Centre. Indianapolis : Cisco Press, 2016. 352 p. **2. Zimmerman C.** Ten Strategies of a World-Class Cybersecurity Operations Centre. Bedford : The MITRE Corporation, 2014. 308 p. **3. Sanders M.** How to Get the Most Value out of Your MSSP and Security Operations. URL: https://securityintelligence.com/how-to-get-the-most-value-out-of-yourmssp-and-security-operations (Accessed: 05 March 2025). **4. Kobie N.** Darktrace's AI is now automatically responding to hacks – and stopping them. 2017. URL: https://www.wired.com/story/darktrace-machine-learning-security/ (Accessed: 03 June 2025). **5. Newman L.** AI Can Help Cybersecurity – If It Can Fight Through the Hype. 2018. URL: https://www.wired.com/story/ai-machine-learning-cybersecurity/ (Accessed: 03 June 2025). **6. Greenberg A.** MIT's Teaching AI How to Help Stop Cyberattacks. 2016. URL: https://www.wired.com/2016/04/mits-teaching-ai-help-analysts-stop-cyberattacks/ (Accessed: 03 June 2025). **7. Newman L.** Gmail Is Catching More Malicious Attachments With Deep Learning. 2020. URL: https://www.wired.com/story/gmail-catching-more-malicious-attachments-deep-learning/ (Accessed: 03 June 2025). **8. Intel Corporation. Intel. Threat Detection Technology**. URL: https://www.intel.com/content/dam/www/public/us/en/documents/product-briefs/tdt-product-brief.pdf (Accessed: 05 March 2025). **9. Mohamed A. A., Al-Saleh A., Sharma S. K. & Tejani G. G.** Zero-day exploits detection with adaptive WavePCA-Autoencoder (AWPA) adaptive hybrid exploit detection network (AHEDNet). *Scientific Reports.* 2025. № 15. P. 4036. DOI: https://doi.org/10.1038/s41598-025-87615-2. **10. Babaey V., Faragardi H. R.** Detecting Zero-Day Web Attacks with an Ensemble of LSTM, GRU, and Stacked Autoencoders. *Computers*. 2025. № 14. P. 205. DOI: https://doi.org/10.3390/computers14060205. **11. Peng S., Han Yu, Li Ruonan, Liu Lichen, Liu Jie, Gu Zh.** ROSE-BOX: A Lightweight and Efficient Intrusion Detection Framework for Resource-Constrained IIoT Environments. *Appl. Sci.* 2025. № 15. P. 6448. DOI: https://doi.org/10.3390/app15126448. **12. Anderson J.** Lightweight AI Models for Real-Time Threat Detection in Resource-Constrained IoT Environments. ResearchGate GmbH. 2025. URL: https://www.researchgate.net/publication/390110099 (Accessed: 17 June 2025). **13. Rahmati M.** Towards Explainable and Lightweight AI for Real-Time Cyber Threat Hunting in Edge Networks. Independent Researcher. 2025. DOI: https://doi.org/10.48550/arXiv.2504.16118. **14. Song D., Brumley D., Yin H., Caballero1 Ju., Jager I., Kang1 M. G., Liang Zh., Newsome J., Poosankam P., Saxena P.** BitBlaze: A New Approach to Computer Security via Binary Analysis. 2008. URL: https://bitblaze.cs.berkeley.edu/papers/bitblaze_iciss 08.pdf (Accessed: 25 June 2025).

MODELS FOR DETECTION OF MULTI-PROFILE ZERO-DAY THREATS IN A RESOURCE-LIMITED ENVIRONMENT

ISMAHILOV Artur, The Institute of Software Systems of the National Academy of Sciences of Ukraine, Kyiv, Ukraine, https://orcid.org/0009-0002-1332-8147

SINITSYN Igor, Doctor of Technical Sciences, Professor, The Institute of Software Systems of the National Academy of Sciences of Ukraine, Kyiv, Ukraine, https://orcid.org/0000-0002-4120-0784

Formulation of the problem in general. *Based on an examination of extant methodologies for detecting multi-profile zero-day threats under martial law conditions characterized by inadequate funding, infrastructure, and expertise, an innovative multi-tiered detection model is proposed. This model integrates preliminary filtering, behavioral analysis, and localized self-learning mechanisms to enhance responsiveness to cyber threats within resource-constrained environments.*

Research methods. *A comprehensive set of theoretical and applied research methods was employed in the course of this study. The theoretical component included analysing and synthesising scientific literature, which enabled the systematisation of existing approaches to zero-day threat detection and identifying their strengths and limitations under resource-constrained conditions. A comparative analysis evaluated the effectiveness of conventional and advanced cyber defence methods based on key criteria such as resource dependence, operational autonomy, and detection accuracy. Among the applied methods, behavioural modelling was utilised to design the foundational structure of a system capable of identifying anomalous activity. To validate the proposed model, computer-based experimentation was conducted, assessing the performance of anomaly detection algorithms - specifically, the autoencoder and Isolation Forest - within a simulated, isolated environment. Furthermore, modelling and prototyping techniques formed the basis for developing an adaptive heuristic monitoring architecture and constructing a multi-level framework for real-time anomaly detection. This methodological framework facilitated a detailed examination of the operational characteristics of the proposed models in low-resource environments, assessed their zero-day threat detection accuracy compared to traditional methods, and supported the implementation of a simulation-based experiment.*

Literature review. *Recent research on detecting multi-profile zero-day threats demonstrates significant progress in applying behavioural analysis, effectively reducing reliance on signature-based detection methods. Scientific studies have examined various approaches for identifying zero-day threats, including signature-based detection, behavioural analysis, vulnerability analysis of program code, and machine learning techniques. This body of research highlights the ongoing need for further investigation into detection methods tailored explicitly to resource-constrained environments, conditions that are particularly characteristic of periods of military conflict.*

Research results*. As a result of the conducted research, a comprehensive analysis of contemporary approaches to detecting multi-profile zero-day threats was performed, with particular consideration given to the operational characteristics of information systems under martial law conditions and limited resources. The study identified key limitations of traditional cybersecurity tools – specifically Security Information and Event Management (SIEM) systems – which demonstrate reduced effectiveness in isolated environments lacking sufficient computational resources and qualified personnel. A novel authorial model for zero-day threat detection was developed in response to these challenges.*

The proposed model integrates preliminary traffic filtering, behavioural analysis, and localised self-learning through lightweight anomaly classification algorithms, including autoencoders and the Isolation Forest method. The model was formalised as a multi-level architecture capable of autonomous functioning, local adaptation, and real-time response to anomalous activities. A new concept of adaptive heuristic monitoring was introduced, enabling the identification of potential zero-day vulnerability exploitation without prior knowledge of specific threat signatures. Additionally, a risk-based process isolation mechanism was formulated, temporarily suspending potentially malicious actions while maintaining the continuity of critical system functions. Experimental model validation was conducted under conditions simulating a realistic cyber conflict environment. The results confirmed the model's high accuracy in detecting diverse anomaly profiles while maintaining a low false positive rate – even when deployed on outdated or low-performance hardware. These findings demonstrate the model's applicability for cybersecurity systems operating under military conflict, resource constraints, or autonomous deployment, thereby significantly enhancing the resilience of critical information infrastructure.

Research novelty. *The model proposed in the study is a multi-level system that integrates preliminary filtering, behavioural analysis, and local self-learning. This design enables the implementation of effective response strategies to unknown threats, even under constrained operational conditions. As a result, the model can function on outdated hardware, accumulate local experience, and adapt to changes in the behaviour of potential threats without compromising operational efficiency.*

Theoretical and practical significance. *The study presents a scientifically grounded approach to evaluating traditional cyber defence tools in resource-constrained environments, where operational autonomy, rapid response, and minimal dependence on external update sources are critically important. The practical significance of the research lies in the potential application of the proposed model to implement a hybrid strategy that integrates the strengths of multiple detection methods. This approach enables the development of a multi-level, flexible, and scalable cybersecurity framework, capable of providing reliable protection against zero-day threats even under limited resources.*

Conclusions and future work*. The proposed hybrid approach, which integrates the advantages of multiple detection methods, enables the implementation of a multi-level, flexible, and scalable cybersecurity framework. Future research will be directed toward developing neural network models with long-term memory capabilities to enhance adaptability to emerging behavioural anomalies. In addition, efforts will focus on designing a multi-agent architecture that supports the decentralised exchange of generalised detection patterns between nodes, thereby eliminating the need for a centralised control system.*

Keywords: *cybersecurity, zero-day threats, hybrid approach, behavioural analysis, self-learning.*

References

1. Muniz, J., McIntyre, G., AlFardan, N., (2016). *Security Operations Centre.* Indianapolis: Cisco Press. **2. Zimmerman, C.** (2014). *Ten Strategies of a World-Class Cybersecurity Operations Centre.* Bedford: The MITRE Corporation. **3. Sanders, M.** *How to Get the Most Value out of Your MSSP and Security Operations.* [online]. Available at: https://securityintelligence.com/how-to-get-the-most-value-out-of-yourmssp-and-security-operations [Accessed: 05 March 2025]. **4. Kobie, N.,** (2017). *Darktrace's AI is now automatically responding to hacks – and stopping them.* [online]. Available at: https://www.wired.com/story/darktrace-machine-learning-security/ [Accessed: 03 June 2025]. **5. Newman, L.,** (2018). *AI Can Help Cybersecurity – If It Can Fight Through the Hype.* [online]. Available at: https://www.wired.com/story/ai-machine-learning-cybersecurity/ [Accessed: 03 June 2025]. **6. Greenberg, A.,** (2016). *MIT's Teaching AI How to Help Stop Cyberattacks.* [online]. Available at: https://www.wired.com/2016/04/mits-teaching-ai-help-analysts-stop-cyberattacks/ [Accessed: 03 June 2025]. **7. Newman, L.,** (2020). *Gmail Is Catching More Malicious Attachments With Deep Learning.* [online]. Available at: https://www.wired.com/story/gmail-catching-more-malicious-attachments-deep-learning/ [Accessed: 03 June 2025]. **8.** *Intel Corporation. Intel® Threat Detection Technology.* [online]. Available at: https://www.intel.com/content/dam/www/public/us/en/documents/product-briefs/tdt-product-brief.pdf [Accessed: 05 March 2025]. **9. Mohamed, A. A., Al-Saleh, A., Sharma, S. K., Tejani, G. G.,** (2025). *Zero-day exploits detection with adaptive WavePCA-Autoencoder (AWPA) adaptive hybrid exploit detection network (AHEDNet).* DOI: https://doi.org/10.1038/s41598-025-87615-2. **10. Babaey, V., Faragardi, H. R.,** (2025). *Detecting Zero-Day Web Attacks with an Ensemble of LSTM, GRU, and Stacked Autoencoder* DOI: https://doi.org/10.3390/computers14060205. **11. Peng, S., Han, Yu , Li, R., Liu, L, Liu, J., Gu, Zh.,** (2025), *ROSE-BOX: A Lightweight and Efficient Intrusion Detection Framework for Resource-Constrained IIoT Environments.* DOI: https://doi.org/10.3390/app15126448. **12. Anderson, J.,** (2025). *Lightweight AI Models for Real-Time Threat Detection in Resource-Constrained IoT Environments.* [online]. Available at: https://www.researchgate.net/publication/390110099 [Accessed: 17 June 2025]. **13. Rahmati, M.** (2025), *Towards Explainable and Lightweight AI for Real-Time Cyber Threat Hunting in Edge Networks. Independent Researcher.* DOI: https://doi.org/10.48550/arXiv.2504. 16118. **14. Song, D., Brumley D., Yin, H., Caballero1, J., Jager, I., Gyung Kang, M., Zhenkai Liang, Z., Newsome, J., Poosankam P., Saxena P.** (2008), *BitBlaze: A New Approach to Computer Security via Binary Analysis.* [online]. Available at: https://bitblaze.cs.berkeley.edu/papers/bitblaze_iciss08.pdf [Accessed: 25 June 2025].

Рукопис надійшов до редакції *26.05.2025*
Рукопис прийнято до друку після рецензування *06.08.2025*
Дата публікації *29.08.2025*