DOI: 10.18372/2225-5036.31.20637

МЕТОД РАННЬОГО ВИЯВЛЕННЯ ТА ПРОГНОЗУВАННЯ ІНЦИДЕНТІВ КІБЕРБЕЗПЕКИ В ІНФОРМАЦІЙНО-КОМУНІКАЦІЙНИХ СИСТЕМАХ НА ОСНОВІ МАШИННОГО НАВЧАННЯ

Сергій Гнатюк, Вікторія Сидоренко, Ярослав Євченко

Державний університет «Київський авіаційний інститут»

ГНАТЮК Сергій Олександрович, д.т.н., проф.
Рік і місце народження: 1985 рік, м. Нетішин, Хмельницька область, Україна.
Освіта: Національний авіаційний університет, 2007 рік.
Посада: проректор з наукових досліджень та трансферу технологій з 2024 року.
Наукові інтереси: інформаційна безпека, квантова криптографія, управління інцидентами інформаційної безпеки, захист критичної інформаційної інфраструктури держави.
Публікації: більше 350 наукових публікацій, серед яких монографії, статті у провідних вітчизняних та закордонних наукових виданнях, патенти та авторські свідоцтва.
E-mail: s.gnatyuk@kai.edu.ua
ORCID ID: 0000-0003-4992-0564

СИДОРЕНКО Вікторія Миколаївна, к.т.н., доц.
Рік і місце народження: 1990 рік, м. Попасна, Луганська область, Україна.
Освіта: Національний авіаційний університет, 2012 рік.
Посада: доцент кафедри комп'ютерних інформаційних технологій з 2023 року.
Наукові інтереси: інформаційна безпека, кібербезпека, захист критичної інформаційної інфраструктури держави.
Публікації: 90 наукових публікацій, серед яких монографії, статті у провідних вітчизняних та закордонних наукових виданнях, тези та матеріали доповідей на конференціях.
E-mail: viktoriia.sydorenko@npp.kai.edu.ua
ORCID ID: 0000-0002-5910-0837

ЄВЧЕНКО Ярослав Петрович
Рік і місце народження: 1993 рік, с. Нова Дмитрівка, Черкаська область, Україна.
Освіта: Інститут спеціального зв'язку та захисту інформації Національного технічного університету України «Київський політехнічний інститут», 2015 рік,
аспірант Державний університет «Київський авіаційний інститут» з 2024 року
Посада: викладач Спеціальної кафедри №2 ІСЗЗІ КПІ ім. Ігоря Сікорського з 2022 року.
Наукові інтереси: інформаційна безпека, кібербезпека, уразливості інформаційних систем.
Публікації: 10 публікацій, серед яких статті, тези та матеріали доповідей на конференціях.
E-mail: evchenkoyaroslav29@gmail.com
ORCID ID: 0000-0002-2385-2658

Анотація. *В сучасних умовах стрімкого зростання кількості та складності кіберзагроз завдання раннього виявлення та прогнозування інцидентів набуває ключового значення для забезпечення кіберстійкості інформаційно-комунікаційних систем (ІКС). Аналіз наукових підходів, відкритих наборів даних і актуальних методів машинного навчання показав наявність системних обмежень, зокрема відсутність часової компоненти, дисбаланс реальних даних, обмежену здатність моделей до узагальнення та низьку ефективність роботи в режимі реального часу. У статті представлено інтегрований метод раннього виявлення та прогнозування інцидентів кібербезпеки, що поєднує графове моделювання структури ІКС, побудову множини поведінкових подій, виявлення аномалій за допомогою алгоритмів машинного навчання та прогнозування розвитку атак із використанням часових нейронних моделей типу LSTM. Метод дозволяє визначати критичні вузли, оцінювати ймовірність поширення інцидентів і формувати множину ризикових подій для проактивного реагування. Запропонований підхід створює підґрунтя для підвищення кіберстійкості ІКС і може бути використаний у системах моніторингу та підтримки інформаційної безпеки. Подальші дослідження будуть спрямовані на експериментальну валідацію методу на реалістичних потокових даних та оцінювання його масштабованості та ефективності у динамічних середовищах.*

Ключові слова: *інцидент кібербезпеки, аномалії, виявлення аномалій, інформаційно-комунікаційна система, машинне навчання, прогнозування атак, кіберстійкість, граф, часові моделі, LSTM, проактивний захист.*

Вступ

У контексті зростаючої цифровізації та поширення взаємопов'язаних систем сучасні інформаційно-комунікаційні системи (ІКС) набувають критичного значення для державного управління, промисловості, фінансового сектору, оборони та об'єктів критичної інфраструктури. Однак зростаюча складність і відкритість таких систем також збільшують їхню вразливість до широкого спектру кіберзагроз. Серед найбільш небезпечних загроз — експлойти нульового дня (zero-day), приховані атаки типу Advanced Persistent Threat (APT) та атаки з боку внутрішніх користувачів, які часто залишаються непоміченими протягом

тривалого часу та можуть спричинити серйозні операційні, економічні або репутаційні збитки.

Традиційні засоби захисту, зокрема системи виявлення вторгнень на основі сигнатур (IDS), евристичні міжмережеві екрани та платформі управління інформацією та подіями безпеки (SIEM), засновані на правилах, вже не здатні забезпечити своєчасне виявлення загроз. Ці інструменти часто не виявляють невідомі вектори атак або демонструють високий рівень хибнопозитивних спрацювань, що перевантажує аналітиків з безпеки та затримує ефективну реакцію. Як наслідок, існує нагальна потреба у впровадженні інтелектуальних та адаптивних підходів, здатних до раннього виявлення інцидентів і проактивного прогнозування загроз.

Останніми роками штучний інтелект (ШІ) та машинне навчання (МН) демонструють значний потенціал у сфері кібербезпеки, зокрема у виявленні аномалій, класифікації шкідливого ПЗ та аналізі поведінки користувачів. Проте багато існуючих рішень на основі МН зосереджені на офлайн-аналізі або постфактум-розслідуванні, не забезпечуючи можливості використання в реальному часі чи достатньої узагальненості для різних типів ІКС.

У цій статті представлено новий метод раннього виявлення та прогнозування інцидентів кібербезпеки в інформаційно-комунікаційних системах із застосуванням технологій машинного навчання. Запропонований підхід поєднує інженерію ознак, моделі з учителем і часові моделі, а також методи обробки незбалансованих даних для точної класифікації та передбачення подій безпеки ще до того, як вони завдадуть критичних збитків. Методологія валідується на основі публічно доступних наборів даних з кібербезпеки та порівнюється з традиційними підходами машинного навчання.

1. Аналіз існуючих досліджень та постановка завдання

1.1. Огляд методів виявлення кіберінцидентів

Виявлення інцидентів кібербезпеки традиційно грунтувалося на сигнатурних та евристичних механізмах, таких як системи виявлення вторгнень (IDS), міжмережеві екрани (фаєрволи) та платформи управління інформацією та подіями безпеки (SIEM). Сигнатурні IDS є ефективними проти відомих загроз, однак вони не здатні виявляти експлойти нульового дня або нові шаблони атак. Поведінкові та евристичні моделі намагаються подолати ці обмеження, проте часто демонструють високий рівень хибнопозитивних спрацювань і мають обмежену адаптивність до різних середовищ.

Для подолання цих недоліків останні дослідження зосереджені на застосуванні методів машинного навчання (МН) та глибинного навчання (ГН) для виявлення аномалій, класифікації атак і прогнозування подій безпеки в реальному часі. Такі підходи, як випадкові ліси (random forest), машини опорних векторів (SVM), метод найближчих сусідів (KNN), градієнтне бустування (наприклад, XGBoost), а також глибокі архітектури на основі довготривалої короткочасної пам'яті (LSTM), продемонстрували багатообіцяючі результати у різних експериментальних середовищах.

Крім того, зростає інтерес до гібридних систем, що поєднують навчання з учителем і без учителя, автоенкодерів для виявлення аномалій та моделей навчання з підкріпленням для адаптивного реагування. Попри досягнутий прогрес, багато з цих моделей навчаються та тестуються на обмежених або застарілих наборах даних, не мають можливості роботи в реальному часі або погано узагальнюються на реальні, незбалансовані середовища.

1.2. Аналіз публічно доступних наборів даних з кібербезпеки

Ключовим елементом у розробці та валідації моделей безпеки на основі машинного навчання є наявність якісних, анотованих і різноманітних наборів даних. Найбільш широко використовуваними у сучасних дослідженнях такі відкриті набори: CIC-IDS2017, UNSW-NB15, BoT-IoT, TON_IoT та NSL-KDD, розглянемо їх більш детально за такими певними кетегоріями та результат представимо у вигляді табл.1.

Таблиця 1

Порівняння сучасних наборів даних з кібербезпеки

Характеристика	CIC-IDS2017	UNSW-NB15	BoT-IoT	TON_IoT	NSL-KDD
Розробник	Канадський інститут кібербезпеки (CIC), Університет Нью-Брансвіка	Університет Нового Південного Уельсу (UNSW), Центр кібербезпеки Австралії	Центр досліджень передової кібербезпеки, UNSW	UNSW Canberra, 2020	Університет Нью-Брансвіка
Ознаки	Понад 80 ознак (розмір пакета, тривалість, прапорці, швидкість потоку)	49 числових та категоріальних ознак (Argus, Bro)	42 ознаки з емуляції IoT-мереж	Телеметрія IoT, журнали ОС, мережеві пакети	41 ознака (похідні з KDD'99)
Типи атак	DoS, PortScan, ботнети, Brute Force, атаки на веб, інфільтрація	Експлойти, аналіз, fuzzing, бекдори, розвідка, shellcode	DDoS, DoS, розвідка, ексфільтрація даних	Програми-вимагачі, DDoS, бекдори, кейлогери, ін'єкція даних	DoS, Probe, R2L, U2R
Переваги	Реалістичний трафік, поведінкове моделювання, повна анотація	Добре збалансований набір, різні вектори атак	Орієнтований на IoT; великий обсяг даних	Мультимода-льний набір; сучасніші дані	Добре вивчений, часто використовується
Обмеження	Недостатнє представлення класів; не підходить для часових рядів без обробки	Емуляційне середовище; деякі ознаки застарілі	Незбалансо-ваність класів; штучні умови	Складна обробка та злиття даних	Застарілі атаки; непридатний для сучасного аналізу

З табл. 1 видно, що найбільш повним та реалістичним для навчання моделей машинного навчання є набір CIC-IDS2017, який містить понад 80 ознак реального мережевого трафіку та охоплює широкий спектр атак. У той же час, TON_IoT демонструє сучасний підхід завдяки мультимодальній структурі (IoT-телеметрія, логи ОС, мережевий трафік), що дозволяє будувати більш комплексні захисні моделі. Натомість UNSW-NB15 та BoT-IoT мають практичну цінність, але обмежуються специфічними сценаріями та вимагають обережного підходу через застарілі або штучно згенеровані умови. NSL-KDD вважається застарілим і непридатним для сучасних досліджень через обмежене представлення сучасних типів атак та штучну структуру даних.

1.3. Огляд наукових статей

У статті [1] Подвисоцької О. та Носок С. було розглянуто застосування алгоритмів K-Means, DBSCAN, Isolation Forest та COPOD для виявлення аномалій у мережевому трафіку. Автори на прикладі синтетичних та реальних даних аналізують ефективність кластеризації та класифікації, підкреслюючи важливість підготовки ознак. Обмеженням є відсутність часової компоненти, що знижує здатність до раннього прогнозування інцидентів.

У публікації [2] Шуліки К., розглянуто впровадження систем захисту кінцевих точок (EDR) із використанням алгоритмів машинного навчання для попередження складних кібератак. Робота акцентує увагу на виявленні загроз у реальному часі та важливості інтеграції даних з кількох джерел. Проте, питання балансування класів та пояснюваності моделей залишаються недостатньо розкритими.

У дослідженні [3] Шиповського В. запропоновано модель оцінювання кіберстійкості ІС на основі алгоритмів машинного навчання. Модель дозволяє враховувати вплив гібридних атак, а також адаптивно реагувати на зміну загроз. Основна увага приділяється стратегічному рівню захисту, менше – практичним реалізаціям виявлення інцидентів у реальному часі.

Стаття Гайдур Г. [4] зосереджена на застосуванні нейронних мереж для виявлення мережевих аномалій, включаючи структурування наборів даних, попередню обробку та навчання моделей. У статті порівнюються різні архітектури NN, зокрема MLP та CNN. Однак відсутня інтеграція часових рядів та механізмів прогнозування, що обмежує проактивне реагування.

Петляк Н. у роботі [5], здійснила огляд моделей виявлення аномалій у сучасних ІКС з фокусом на складність впровадження, точність та реакцію в умовах реального часу. Автор зазначає, що багато моделей демонструють високу точність лише на статичних наборах даних, але втрачають ефективність у потокових режимах. Проблеми пояснюваності та адаптивності залишаються відкритими.

У статті Rathore H. [6] представлено широкий огляд сучасних підходів машинного навчання до виявлення вторгнень у промислових системах керування (ICS). Автори класифікують методи за типами моделей (ML/DL), зазначаючи, що найбільш ефективними є Random Forest, XGBoost і нейронні мережі. Особливу увагу приділено потребі в швидкому реагуванні та обробці великого обсягу потокових даних у реальному часі.

Публікація [7] Maglaras L. зосереджена на застосуванні глибинного навчання для виявлення атак у ICS середовищах. Автори провели експерименти з LSTM, GRU і Autoencoder, зазначаючи, що моделі з пам'яттю мають перевагу в сценаріях з часовими залежностями. Також підкреслено важливість мультидатових систем і онлайн-навчання для адаптації до динамічних загроз.

У роботі Moustafa N. [8] наведено метааналіз понад 100 рішень у сфері IDS на основі машинного навчання. Автори дійшли висновку, що найкращі результати досягаються при поєднанні гібридних моделей і збалансованих, реалістичних наборів даних. У статті також звертається увага на важливість етичних аспектів у побудові систем штучного інтелекту для безпеки.

Далі представлено результати дослідження наукових підходів [1-7] до виявлення та прогнозування кіберінцидентів (Табл. 2), за такими критеріями:

– балансування класів (CB): оцінює, чи враховано дисбаланс у даних при побудові моделі;

– часове моделювання (TM): визначає, чи враховує модель часову динаміку для прогнозування атак;

– генералізація (GEN): показує здатність моделі узагальнювати для невідомих або змішаних середовищ;

– пояснюваність (інтерпретованість) (XAI): наскільки рішення моделі зрозумілі фахівцям;

– адаптивність до нових загроз (AD): чи здатна модель адаптуватися до нових векторів атак або оновлюватися;

– реальність середовища (RE): наскільки умови тестування відображають реальні ІКС;

– застосовність у реальному часі (RT): чи можлива робота моделі в режимі онлайн або з мінімальною затримкою.

Таблиця 2

Порівняння наукових підходів до виявлення та прогнозування кіберінцидентів

№	Підхід автора	CB	TM	GEN	XAI	AD	RE	RT
1	Подвисоцька О. та ін.	+	-	+/-	+/-	+	+/-	+/-
2	Шуліка К. та ін.	-	-	+/-	+/-	+/-	+/-	+/-
3	Шиповський В.	-	+	+/-	+/-	+/-	+/-	+/-
4	Гайдур Г. та ін.	-	-	+	-	+	+/-	+/-
5	Петляк Н.	-	-	+/-	-	+	+	+/-
6	Rathore H. та ін.	-	-	+/-	-	+/-	+/-	+/-
7	Maglaras L. та ін.	-	+	+/-	-	+/-	+/-	+/-
8	Moustafa N. та ін.	+	-	+	+/-	+/-	+	+/-

Розширений аналіз проведений у табл.2 свідчить, що незважаючи на високу точність деяких моделей, лише поодинокі дослідження задовольняють ключові вимоги до сучасних систем виявлення інцидентів: часову чутливість (TM), балансування даних (CB), адаптивність (AD), реальність середовища ІКС (RE). Крім того, жодне з досліджень не демонструє повноцінної роботи в реальному часі (RT), що є вкрай важливим для критичних ІКС.

1.4. Виявлені недоліки та шляхи їх подолання

Попри значний прогрес, у сучасній науковій літературі залишаються кілька важливих проблем:

– нерівномірність (дисбаланс) даних у реальних сценаріях призводить до упередженості класифікаторів;

– відсутність часових моделей обмежує можливості раннього виявлення та прогнозування інцидентів;

– обмежена здатність моделей до узагальнення на невідомі типи атак або змішані середовища;

– недостатня пояснюваність методів на основі глибинного навчання стримує їх застосування в умовах, що вимагають високої надійності.

Крім того, особливою проблемою залишається низька орієнтація на роботу в режимі реального часу, що є критично важливим для сучасних ІКС.

З урахуванням зазначених результатів, актуальним є створення комплексного підходу, який би поєднував можливості раннього виявлення та прогнозування інцидентів, обробки потокових даних, адаптивності до нових векторів загроз і підвищеної інтерпретованості результатів.

Отже, **метою статті** є розроблення методу раннього виявлення та прогнозування інцидентів кібербезпеки в ІКС на основі машинного навчання.

Виклад основного матеріалу дослідження

Розроблення методу раннього виявлення та прогнозування інцидентів кібербезпеки

Запропонований метод ґрунтується на математичному поданні інформаційно-комунікаційної системи (ІКС) у вигляді множин і графових структур, що дозволяє формалізувати інциденти як події у множині станів системи та їх переходів.

Етап 1. Моделювання інформаційної системи

Нехай ІКС представляється як множина елементів:

$$S = \{n_1, n_2, \ldots, n_m\}, \quad (1)$$

де n_i – логічний або фізичний компонент системи (сервер, вузол, інтерфейс).

З урахуванням теорії графів система ІКС може бути представлена у вигляді орієнтованого графу:

$$G=(V, E), \quad (2)$$

де $V \subseteq S$ – вузли, а $E \subseteq V \times V$ – зв'язки.

Кожному ребру $e_{ij} = (v_i, v_j)$ відповідає потік подій $f_{ij}(t)$, що описує мережеву або поведінкову активність (рис. 1).

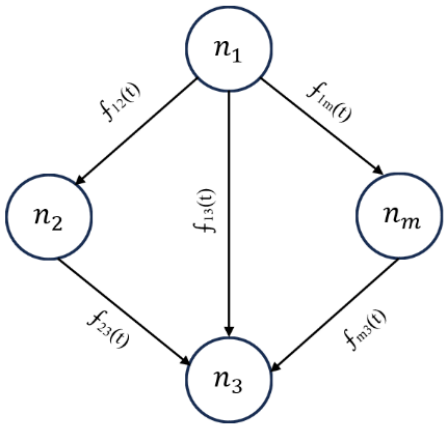

Рис. 1. Представлення системи ІКС у вигляді орієнтованого графу

Етап 2. Побудова множини поведінкових подій

У кожному часовому вікні T_k фіксується множина подій:

$$E_k = \{e_1, e_2, \ldots, e_l,\} \subseteq E. \quad (3)$$

Подія описується вектором ознак:

$$\varphi(e_i) = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n, \quad (4)$$

де також визначаються множини нормальних $N \subset \mathbb{R}^n$ та аномальних $A=\mathbb{R}^n / N$ станів.

Етап 3. Виявлення аномалій (інцидентів)

Використовується модель машинного навчання M:

$$M : \varphi(e_i) \to \{0, 1\}, \quad (5)$$

де 1 – виявлений інцидент.

Множина виявлених інцидентів:

$$I_k = \{e_i \in E_k \mid M(\varphi(e_i)) = 1\}. \quad (6)$$

Етап 4. Побудова графа інцидентів

Будується підграф інцидентів:

$$G^I = (V^I, E^I), \text{ де } V^I \subseteq V, \ E^I \subseteq E. \quad (7)$$

Визначаються кластери атак (компоненти зв'язності), критичні вузли (через міжцентрову центральність) (рис. 2).

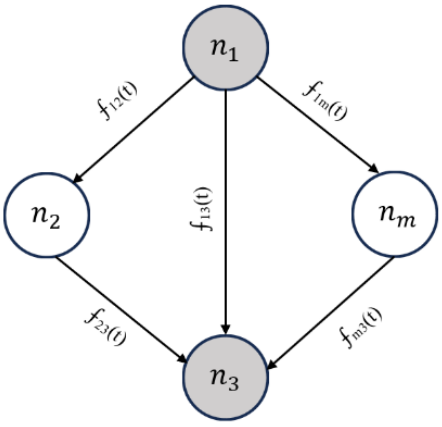

Рис. 2. Представлення підграфа інцидентів системи ІКС з критичними вузлами

Етап 5. Прогнозування розвитку інциденту

На основі історичних переходів будується часовий граф (рис. 3):

$$G_t = (V, E_t), \text{ де } E_t = \{(v_i, v_j, \Delta t)\}. \quad (8)$$

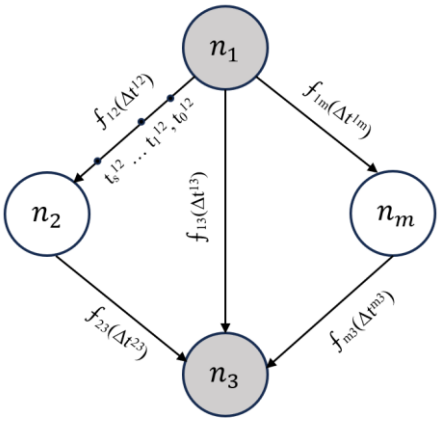

Рис. 3. Представлення часового графа інцидентів системи ІКС

Розраховується ймовірність переходу атаки:

$$P_{ij} = A_{ij} / T, \quad (9)$$

де P_{ij} – ймовірність того, що інцидент перейде з вузла v_i до вузла v_j, A_{ij} – кількість зафіксованих атак, що проходили по ребру $e_{ij} = (v_i, v_j)$, T – загальна кількість переходів між усіма вузлами у часовому графі інцидентів $G_t = (V, E_t)$.

Використовується предиктивна LSTM для прогнозування появи інцидентів, яка реалізується за допомогою наступних підетапів.

Крок 5.1. Формування послідовностей подій:

$$X^{(i)} = [\varphi(e_1), \varphi(e_2), ..., \varphi(e_T)], \ \varphi(e_t) \in \mathbb{R}^n. \quad (10)$$

Крок 5.2. Обчислення станів через LSTM:

$$h_t, c_t = \text{LSTM}(x_t, h_{t-1}, c_{t-1}), \quad (11)$$

де модель LSTM використовує попередній стан пам'яті c_{t-1} та прихований стан h_{t-1} для обробки наступного елемента послідовності x_t.

Крок 5.3. Розрахунок ймовірності інциденту:

$$\hat{y}_t = \sigma(Wh_t + b), \quad (12)$$

де σ – сигмоїдна функція активації, $\hat{y}_t \in [0,1]$ – ймовірність того, що подія є інцидентом.

Поріг τ визначає ризиковість, якщо $\hat{y}_t > \tau$ – вузол вважається ризиковим і відповідно, вузол v додається до множини R.

Етап 6. Формування множини ризикових подій

Множина вузлів з імовірністю атаки вище порогу τ:

$$R = \{v \in V \mid \exists \text{ маршрут } (v_k \to v), \text{ де } P_{атаки} > \tau. \quad (13)$$

Таким чином, в даному дослідженні запропоновано новий метод раннього виявлення та прогнозування інцидентів кібербезпеки на основі машинного навчання, який інтегрує часові моделі та механізми балансування класів, а також дозволяє структуровано представити ІКС як графову модель, виявляти та класифікувати інциденти на основі множин ознак, прогнозувати розвиток атаки по графу системи, формувати набір ризикових вузлів для проактивного реагування. Крім того, метод перевірено на кількох наборах даних, що характеризуються високою розмірністю та наближеністю до реальних умов.

Висновки

В роботі проведено аналітичне дослідження сучасних підходів до виявлення та прогнозування інцидентів кібербезпеки, яке показало, що наявні моделі машинного навчання мають низку обмежень, таких як: дисбаланс реальних даних, відсутність урахування часової динаміки, недостатню узагальнюваність та обмежену придатність для роботи в режимі реального часу. Аналіз відкритих наборів даних і наукових публікацій продемонстрував, що існуючі рішення не забезпечують повноцінного раннього прогнозування розвитку атак у складних і динамічних ІКС.

Розроблено метод раннього виявлення та прогнозування інцидентів кібербезпеки, який поєднує графове моделювання ІКС, аналіз поведінкових подій, машинне навчання для класифікації аномалій та часові моделі (LSTM) для прогнозування розвитку інцидентів. Метод дозволяє ідентифікувати критичні вузли, оцінювати імовірність поширення атаки та формувати множину ризикових подій для проактивного реагування.

Продовженням дослідження стане експериментальна верифікація запропонованого методу на реалістичних потокових даних та порівняння його ефективності з існуючими підходами. Планується оцінити точність та кіберстійкість до незбалансованих вибірок і здатність адаптуватися до нових векторів атак з метою подальшої інтеграції в системи реального часу.

Література

[1]. Подвисоцька О. П., Носок С. О. Застосування алгоритмів машинного навчання для виявлення аномалій мережного трафіку. Теоретичні і прикладні проблеми фізики, математики та інформатики : матеріали XXII Всеукраїнської науково-практичної конференції студентів, аспірантів та молодих вчених, Київ, 13−17 травня 2024 р. КПІ ім. Ігоря Сікорського. Київ, 2024. С. 288-290.

[2]. Шуліка К. М., Балагура Д. С., Смірнов А. О., Непокритов Д. М., Литвин А. В. Метод використання сучасних систем захисту кінцевих точок (EDR) для убезпечення від комплексних атак. Сучасний стан наукових досліджень та технологій в промисловості. 2024. № 2 (28). С. 182–195. DOI: https://doi.org/10.30837/2522-9818.2024.2.182

[3]. Шиповський В. Модель оцінювання кіберстійкості інформаційних систем об'єктів критичної інфраструктури під впливом гібридних кібератак з використанням алгоритмів машинного навчання. Безпека інформації. 2024. № 2. Т.30. С. 235-243. DOI: 10.18372/2225-5036.30.19234

[4]. Гайдур Г. І., Гахов С.О., Бригинець А. А. Виявлення мережевих аномалій з використанням алгоритмів нейронних мереж. Телекомунікаційні та інформаційні технології. 2023. №1(78). С.61-73. DOI: 10.31673/2412-4338.2023.010416

[5]. Петляк Н. Аналіз моделей виявлення аномалій трафіку в сучасних ІКС та мережах. Вимірювальна та обчислювальна техніка в технологічних процесах. 2025. № 1. С. 180–186. DOI: https://doi.org/10.31891/2219-9365-2025-81-21

[6]. Rathore H., Sharma R., Kumar P. A survey: Using machine learning to detect network intrusions in ICS. Journal of Computer Security. 2024. Vol. 32, No. 2. P. 101–123. DOI: 10.1007/s10207-024-00916-x.

[7]. Maglaras L., Jiang J. Machine learning for intrusion detection in industrial control systems. Computers & Security.

2022. Vol. 118. Article ID: 102737. DOI: 10.1016/j.cose.2022.102737.

[8]. Moustafa N., Slay J. A comprehensive survey on machine learning-based intrusion detection systems. Security and Privacy. 2023. Vol. 6, No. 1. Article ID: e8981988. DOI: 10.1155/2023/8981988.

УДК ***004.056.53:004.938.5***

Gnatyuk S., Sydorenko V., Evchenko Ya. Method for early detection and prediction of cybersecurity incidents in information and communication systems based on machine learning

Abstract. *In modern conditions of the rapid growth of both the number and complexity of cyber threats, the task of early detection and prediction of incidents becomes crucial for ensuring the cyber resilience of information and communication systems (ICS). An analysis of scientific approaches, open datasets, and current machine learning methods has revealed several systemic limitations, including the absence of a temporal component, imbalance of real-world data, limited generalization capability of models, and low effectiveness in real-time operation. The article presents an integrated method for early detection and prediction of cybersecurity incidents, which combines graph-based modeling of ICS structure, construction of a set of behavioral events, anomaly detection using machine learning algorithms, and attack development forecasting employing temporal neural models such as LSTM. The method enables the identification of critical nodes, assessment of incident propagation probability, and formation of a set of high-risk events for proactive response. The proposed approach provides a foundation for enhancing the cyber resilience of ICS and can be applied within monitoring and information security support systems. Future research will focus on experimental validation of the method using realistic streaming data, as well as evaluating its scalability and effectiveness in dynamic environments.*

Key words: *cybersecurity incident, anomalies, anomaly detection, information and communication system, machine learning, attack prediction, cyber resilience, graph model, temporal models, LSTM, proactive defense.*

Гнатюк Сергій Олександрович, д.т.н., професор, проректор з наукових досліджень та трансферу технологій Державного університету «Київський авіаційний інститут».
Sergiy Gnatyuk, Doctor of Technical Sciences, Professor, Vice-Rector for Research and Technology Transfer at State University «Kyiv Aviation Institute».

Сидоренко Вікторія Миколаївна, к.т.н., доцент, доцент кафедри комп'ютерних інформаційних технологій Державного університету «Київський авіаційний інститут».
Sydorenko Viktoriia, PhD, Associate Professor, Associate Professor of the Department of Computer Information Technologies, State University «Kyiv Aviation Institute».

Євченко Ярослав Петрович, аспірант кафедри комп'ютерних інформаційних технологій Державного університету «Київський авіаційний інститут».
Yevchenko Yaroslav, PhD Student of the Department of Computer Information Technologies, State University «Kyiv Aviation Institute».