Міністерство освіти і науки України
Тернопільський національний технічний університет імені Івана Пулюя

Факультет комп'ютерно–інформаційних систем і програмної інженерії
(повна назва факультету)
Кафедра кібербезпеки
(повна назва кафедри)

КВАЛІФІКАЦІЙНА РОБОТА

на здобуття освітнього ступеня

магістр
(назва освітнього ступеня)

на тему: Можливості застосування штучного інтелекту в Операційному центрі безпеки

Виконав: студент VI курсу, групи СБм–61
спеціальності 125 Кібербезпека
(шифр і назва спеціальності)

Задорожний С.Ю.
(підпис) (прізвище та ініціали)

Керівник *Скарга–Бандурова І.С.*
(підпис) (прізвище та ініціали)

Нормоконтроль Лечаченко Т.А
(підпис) (прізвище та ініціали)

Завідувач кафедри Загородна Н.В
(підпис) (прізвище та ініціали)

Рецензент
(підпис) (прізвище та ініціали)

Тернопіль
2023

Міністерство освіти і науки України
Тернопільський національний технічний університет імені Івана Пулюя

Факультет комп'ютерно-інформаційних систем і програмної інженерії
(повна назва факультету)

Кафедра кібербезпеки
(повна назва кафедри)

ЗАТВЕРДЖУЮ
Завідувач кафедри
_____ Загородна Н.В.
(підпис) (прізвище та ініціали)
«__» _____ 2023 р.

ЗАВДАННЯ
НА КВАЛІФІКАЦІЙНУ РОБОТУ

на здобуття освітнього ступеня Магістр
(назва освітнього ступеня)

за спеціальністю 125 Кібербезпека
(шифр і назва спеціальності)

Студенту Задорожному Святославу Юрійовичу
(прізвище, ім'я, по батькові)

1. Тема роботи Можливості застосування штучного інтелекту в Операційному центрі безпеки

Керівник роботи *Скарга–Бандурова І.С*
(прізвище, ім'я, по батькові, науковий ступінь, вчене звання)

Затверджені наказом ректора від «_» _ 2023 року № _

2. Термін подання студентом завершеної роботи 26.12.2023р.

3. Вихідні дані до роботи наукові літературні джерела

1 АНАЛІЗ ІСНУЮЧИХ ВИКЛИКІВ ТА ПРОБЛЕМ В ОБЛАСТІ ІНФОРМАЦІЙНОЇ БЕЗПЕКИ В SOC 1.1 Актуальність проблеми використання ШІ в системах кібербезпеки 1.1.1 Ландшафт сучасних кіберзагроз та їх вплив на системи безпеки 1.1.1.1 Шкідливе програмне забезпечення 1.1.1.2 Програми вимагачі 1.1.1.3 Атака на виявлення та запобігання вторгнень (IDS/IPS) 1.1.1.5 Атаки на системи SIEM 1.1.2 Пояснення важливості впровадження ШІ для підвищення ефективності заходів з кібербезпеки 1.2 Автоматизоване виявлення аномалій висновок до першого розділу 2 ПЕРЕДУМОВИ ІНТЕГРАЦІЇ IPS/IDS, SIEM ТА DLP ДЛЯ ЕФЕКТИВНОГО ЗАХИСТУ ВІД КІБЕР ЗАГРОЗ 2.1 аналіз функцій та завдань систем IPS/IDS в контексті кібербезпеки 2.2 Огляд проблем та викликів, що виникають при використанні традиційних систем IPS/IDS 2.2.1 Детальний огляд ролі та функцій систем SIEM в SOC 2.2.2 Аналіз проблем та труднощів, що можуть виникати при їх впровадженні та експлуатації 2.3 Розгляд функцій та важливості DLP в запобіганні витокам конфіденційної інформації висновок другого розділу 3.ПЕРСПЕКТИВИ ВИКОРИСТАННЯ ШІ В SOC ЦЕНТРАХ НА ОСНОВІ УНІФІКОВАНОЇ СИСТЕМИ 3.1 методи поліпшення кібербезпеки в SOC та вектори розвитку ШІ 3.1.1 Класичний підхід до кібербезпеки 3.1.2 Напів–уніфікована система зі ШІ 3.1.2.1 Представлення дерева подій напів–уніфікованої системи 3.1.2.2 Переваги та недоліки напів–уніфікованої системи 3.1.3 Уніфікована система 3.1.3.1 Представлення дерева подій 3.1.3.2 Переваги та недоліки 3.1.4 Порівняння напів–уніфікованої та уніфікованої систем кіберзахисту 3.1.5 Приклади атак та відповідні методи 3.2 Схема SOC центру на основі уніфікованої системи 3.3 Оцінки ефективності систем що входять до складу SOC 3.4 ПРАКТИЧНА РЕАЛІЗАЦІЯ АЛГОРИТМІВ МАШИННОГО НАВЧАННЯ ДЛЯ ВИЯВЛЕННЯ ІНТРУЗІЙ В МЕРЕЖІ 3.4.1 Основні етапи процесу

розробки і тестування моделей 3.4.2 Набір даних 3.4.3 Попередній аналіз та обробка даних 3.4.4 Вибір і тренування моделі 3.4.5 Перевірка якості моделей висновок 4.1 4.1 Охорона праці 4.2 Безпека в надзвичайних ситуаціях 4.3 Висновок до четвертого розділу ВИСНОВОК

5. Перелік графічного матеріалу (з точним зазначенням обов'язкових креслень, слайдів)

6. Консультанти розділів роботи

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		завдання видав	завдання прийняв
Охорона праці	Осухівська Г.М., к.т.н., доцент		
Безпека в надзвичайних ситуаціях	Клепчик В.М., старший викладач з адміністративно-господарської роботи та будівництва		

7. Дата видачі завдання 1 Листопада 2023 р.

КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назва етапів роботи	Термін виконання етапів роботи	Примітка
1.	Ознайомлення з завданням до кваліфікаційної роботи	01.10 – 02.10	*Виконано*
2.	Підбір джерел про принципи роботи штучного інтелекту		
	в SOC	02.10 – 05.10	*Виконано*
3.	Опрацювання джерел про принципи роботи штучного		
	інтелекту в SOC	05.10 – 10.10	*Виконано*
4.	Виконання дослідження принципипів роботи штучного		
	інтелекту в SOC	11.10 – 18.10	*Виконано*
5.	Підготовка матеріалу	19.10 – 24.10	*Виконано*
6.	Оформлення розділу «Аналіз предметної області»	25.10 – 29.10	*Виконано*
7.	Оформлення розділу «Теоретична частина»	30.10 – 02.10	*Виконано*
8.	Оформлення розділу «Практична частина»	03.10 – 8.12	*Виконано*
9.	Виконання завдання до підрозділу «Охорона праці»	9.12 – 10.12	*Виконано*
10.	Виконання завдання до підрозділу «Безпека в надзви-	10.12 – 11.12	*Виконано*
	чай них ситуаціях»		
11.	Оформлення кваліфікаційної роботи	14.12 – 15.12	*Виконано*
12.	Нормоконтроль	20.12 – 21.12	
13.	Перевірка на плагіат	12.12	
14.	Попередній захист кваліфікаційної роботи		
15.	Захист кваліфікаційної роботи	26.12	

Студент _____ Задорожний С.Ю.
(підпис) (прізвище та ініціали)

Керівник роботи _____ *Скарга–Бандурова І.С*
(підпис) (прізвище та ініціали)

АНОТАЦІЯ

Можливості застосування штучного інтелекту в операційному центрі безпеки // Кваліфікаційна робота на отримання освітнього рівня «Магістр» // Задорожний Святослав Юрійович // Тернопільський національний технічний університет імені Івана Пулюя, факультет комп'ютерно–інформаційних систем і програмної інженерії, кафедра кібербезпеки, група СБм–61 // Тернопіль, 2023 // С. 85, рис. – 7, табл. – 4, бібліогр. – 21, дод. – 2.

Кваліфікаційна робота складається з пояснювальної записки та графічної частини (ілюстративний матеріал – слайди). Об'єм графічної частини кваліфікаційної роботи становить 16 слайдів. Об'єм пояснювальної записки складає 61 друкованих сторінок формату А4 (210×297).

Ключові слова: ЦЕНТР ОПЕРАЦІЙ З БЕЗПЕКИ, ІНФОРМАЦІЙНА БЕЗПЕКА, КОМПЛЕКС ЗАСОБІВ ЗАХИСТУ, ШТУЧНИЙ ІНТЕЛЕКТ, МАШИННЕ НАВЧАННЯ

Основна мета цього дослідження полягає у вдосконаленні систем запобігання інтрузій (IPS), систем управління інформаційною безпекою (SIEM) та систем запобігання витоку даних (DLP) у контексті центрів операцій з безпеки (SOC) за допомогою методів штучного інтелекту (ШІ) та машинного навчання (МН). Дослідження спрямоване на виявлення та аналіз проблем та викликів, що виникають у зазначених системах, та визначення можливостей їх подолання за допомогою передових технологій.

В рамках кваліфікаційної роботи проаналізовано концепції і підходи організації центрів операційної безпеки на базі ШІ. Також виконано порівняння різних підходів до побудови SOC центру, визначені переваги і недоліки цих концепцій. Запропонована метрика і підходи до оцінювання системи побудованної на принципах уніфікації та поєднання ШІ.

ABSTRACT

Possibilities of Artificial Intelligence Application in the Security Operations Center // Qualification Thesis for the Educational Level 'Master' // Zadorozhnyi Sviatoslav Yuriiovych // Ivan Puluj Ternopil National Technical University, Faculty of Computer and Information Systems, Department of Cybersecurity, Group SBm–61 // Ternopil, 2023 // P. 85, Fig. – 7, Table. – 4, Bibl. – 21, appen. – 2.

The qualification thesis consists of an explanatory note and a graphic part (illustrative material – slides). The volume of the graphic part of the qualification work is 16 slides. The volume of the explanatory note is 61 printed pages of A4 format (210×297).

Keywords: SECURITY OPERATIONS CENTER, INFORMATION SECURITY, PROTECTION SUITE, ARTIFICIAL INTELLIGENCE, MACHINE LEARNING

The main goal of this research is to improve Intrusion Prevention Systems (IPS), Security Information and Event Management (SIEM) systems, and Data Loss Prevention (DLP) systems in the context of Security Operations Centers (SOC) using Artificial Intelligence (AI) and Machine Learning (ML) methods. The research aims to identify and analyze problems and challenges in the mentioned systems and determine the possibilities of overcoming them using advanced technologies.

Within the scope of the qualification thesis, concepts and approaches to the organization of security operations centers based on AI are analyzed. A comparison of different approaches to building SOC centers is also performed, identifying the advantages and disadvantages of these concepts. A metric and approaches to evaluating a system built on the principles of unification and the combination of AI are proposed.

ЗМІСТ

HYPERLINK \l "_Toc154346009"

ВСТУП

Актуальність роботи. У сучасному світі, де обробка та передача конфіденційної інформації стають ключовими аспектами бізнес–процесів, питання забезпечення безпеки цих даних набуває вирішального значення. Інформація стає цінним активом, а недостатній рівень її захисту може привести до серйозних наслідків, таких як витоки даних, порушення конфіденційності та втрати репутації компаній. Саме тому, актуальність цієї кваліфікаційної роботи, спрямованої на вивчення можливостей штучного інтелекту у контексті SOC, є беззаперечною. Розглядаючи широкий спектр існуючих викликів та проблем у зазначених системах, робота ставить за мету знайти ефективні шляхи їх вирішення за допомогою передових технологій. У даній роботі розглянуто глобальний підхід до безпеки, що враховує всі аспекти витоку інформації та встановлює високий стандарт для її захисту в інформаційно–технологічному середовищі. Важливість цього підходу полягає в тому, що традиційні методи безпеки, такі як брандмауери, виявляються обмежено ефективними у світлі сучасних викликів інформаційної безпеки. У контексті цього дослідження, ШІ та МН стають ключовими технологіями для оптимізації ефективності IPS/IDS, SIEM, і DLP. Вони надають змогу глибоко аналізувати поведінку користувачів, ідентифікувати нормальні та аномальні патерни поведінки, а також реагувати на потенційні загрози в режимі реального часу. SIEM забезпечує постійний моніторинг та аналіз подій, а IDS/IPS виявляють та блокують вторгнення. Об'єднані, ці системи стають надійним засобом виявлення, реагування та запобігання загрозам конфіденційної інформації. Інтеграція ШІ та МН в ці системи дозволяє оптимізувати їх загальну ефективність, забезпечуючи максимальний захист даних.

Окрім того, важливо розуміти, як впровадження ШІ може вплинути на робочі процеси в операційних центрах безпеки, які традиційно покладалися на

людський фактор та експертну оцінку. Використання ШІ для аналізу великих обсягів даних, які не можуть бути ефективно оброблені людиною, відкриває нові горизонти для захисту інформаційних систем. Така автоматизація не лише покращує швидкість та точність виявлення загроз, але й дозволяє спеціалістам концентруватися на складніших завданнях, що потребують глибокого аналізу та критичного мислення. Застосування ШІ у SOC передбачає інтеграцію алгоритмів машинного навчання, які можуть навчатися на базі існуючих даних про інциденти, постійно вдосконалюючись та адаптуючись до нових типів загроз. Це не лише зменшує час на виявлення та реагування на інциденти, але й дозволяє прогнозувати потенційні загрози на основі аналізу тенденцій та патернів.

Отже, **метою** даної роботи є дослідження можливостей використання штучного інтелекту (ШІ) та методів машинного навчання (МН) для вдосконалення ефективності систем запобігання інтрузій (IPS/IDS), систем управління інформаційною безпекою та систем запобігання витоку даних (DLP) у контексті Security Operations Center (SOC).

Для отримання поставленої мети, в роботі необхідно виконати наступні задачі:

— Провести аналіз існуючих викликів та проблем в області інформаційної безпеки в SOC та визначення ефективних шляхів їх вирішення з використанням передових технологій.

— Виконати огляд передумов інтеграції IPS/IDS, SIEM та DLP для ефективного захисту від кіберзагроз

— Розробити концептуальну схему використання ШІ та МН в SOC центрах на основі уніфікованої системи

— Оцінити ефективність методів МН для задачі виявлення мережевих інтрузій.

Об'єкт дослідження: Системи та процеси захисту інформації в Центрах операцій з безпеки (Security Operations Center, SOC).

Предмет дослідження: Використання штучного інтелекту та методів машинного навчання для оптимізації роботи SOC центрів.

Наукова новизна: Розробка інтегрованого підходу до застосування ШІ та МН для оптимізації систем інформаційної безпеки в SOC. Аналіз впливу ШІ та МН на ефективність та швидкість реакції на потенційні загрози.

Практичне значення роботи: Набули подальшого розвитку методи запобігання інтрузій, управління інформаційною безпекою та запобігання витоку даних за допомогою технологій ШІ та МН. У роботі підкреслено, що успіх впровадження ШІ залежить не лише від технологій, але й від стратегічного підходу, що включає розуміння потреб користувачів, управління змінами та неперервне навчання.

Апробація результатів магістерської роботи. Окремі результати роботи доповідались на XI науково–технічній конференції «Інформаційні моделі, системи та технології», Тернопіль, ТНТУ, 7 – 8 грудня 2023 р.

Публікації. За темою роботи з викладенням її основних результатів опубліковані тези в збірнику матеріалів науково–практичної конференції (Додаток А).

1 АНАЛІЗ ІСНУЮЧИХ ВИКЛИКІВ ТА ПРОБЛЕМ В ОБЛАСТІ ІНФОРМАЦІЙНОЇ БЕЗПЕКИ В SOC

1.1 Актуальність проблеми використання ШІ в системах кібербезпеки

Використання ШІ в системах кібербезпеки є надзвичайно важливою проблемою в сучасному цифровому світі. З кожним днем збільшується обсяг цифрових загроз, зокрема атаки на інформаційні системи, кібер шпигунство та інші форми кіберзлочинності. У зв'язку з цим, виникає потреба в удосконаленні засобів та методів захисту.

ШІ відіграє ключову роль у підвищенні ефективності кібербезпеки через використання алгоритмів машинного навчання, глибокого навчання та інших технологій. Автоматизовані системи на основі ШІ можуть швидко аналізувати величезні обсяги даних, виявляти аномалії та реагувати на потенційні загрози набагато ефективніше, ніж традиційні методи. Застосування ШІ в кібербезпеці також є важливим для підтримки стійкості та адаптивності систем. Автоматизовані алгоритми можуть навчатися на основі нових загроз та апгрейдів без значних затримок у виправленні вразливостей.

1.1.1 Ландшафт сучасних кіберзагроз та їх вплив на системи безпеки.

Сучасні кіберзагрози представляють собою серйозну загрозу для інформаційної безпеки, оскільки вони можуть призвести до втрати конфіденційної інформації, перерв у роботі систем та значних фінансових ризиків та збитків. Далі розглянуто найбільш відомі з них.

1.1.1.1 Шкідливе програмне забезпечення

Види шкідливого програмного забезпечення. Включає в себе віруси, черв'яки, троянські коні, рекламне програмне забезпечення, шпигунське програмне забезпечення та інші форми шкідливих програм.

Методи поширення. Може вводитися через заражені файли, веб–сайти, електронну пошту, та інші засоби.

Вплив:

— Видалення чи пошкодження файлів. Шкідливе програмне забезпечення може видаляти чи пошкоджувати файли, що може призвести до втрати важливої інформації.

— Сповільнення роботи системи. Збільшення обчислювального навантаження може призвести до сповільнення роботи комп'ютера чи інших пристроїв.

— Викрадення інформації. Шпигунське програмне забезпечення може збирати конфіденційну інформацію, таку як логіни, паролі чи фінансові дані.

1.1.1.2 Програми вимагачі

Програма вимагач шифрує файли на комп'ютері чи в мережі з метою обміну за викуп.

Часто розповсюджується через електронну пошту, приховані веб–сайти та вразливості програмного забезпечення.

Вплив:

— Втрата доступу. Важливі файли стають недоступними для власника, оскільки вони зашифровані.

— Фінансові втрати. Зловмисники вимагають викуп за ключ розшифрування, що може призвести до фінансових збитків.

— Негативний вплив на бізнес. Якщо системи або дані компанії заражені, це може суттєво вплинути на роботу організації та її репутацію, а також завдати шкоди клієнтам.

1.1.1.3 Атака на виявлення та запобігання вторгнень (IDS/IPS)

Зловмисники можуть використовувати техніки, які дозволяють приховати шкідливий трафік, зробивши його непомітним для систем виявлення вторгнень.

Використання шифрування для захисту від виявлення, приховуючи зловмисну активність в зашифрованому трафіку.

Впровадження аномалій в трафік. Зловмисники можуть модифікувати трафік для введення аномалій, що може спричинити помилкове виявлення або не виявлення шкідливої активності.

Використання інструментів для обходу сигнатур. Зловмисники використовують спеціалізовані інструменти, які обходять сигнатурні методи виявлення, дозволяючи їм пройти невиявленими [5, 14].

Вплив атак на IDS/IPS:

— Неочікувані вторгнення. Успішне маскування та обхід систем IDS/IPS може призвести до неочікуваних вторгнень, що залишають систему без захисту та вразливою перед зловмисною діяльністю.

— Експлуатація вразливостей. Зловмисники можуть використовувати обхідні техніки для експлуатації виявлених вразливостей, надаючи їм несанкціонований доступ та контроль над системою.

— Порушення доступності. У разі успішного обходу та маскування можливе порушення нормальної роботи системи, зниження її доступності та виклик серйозних проблем для користувачів чи організації.

— Знищення конфіденційності. У випадку виявлення обходу або маскування може виникнути ризик знищення конфіденційності, оскільки зловмисники можуть отримати доступ до чутливої інформації.

1.1.1.4 DDoS–атаки

DDoS–атаки є важливою складовою кіберзагроз та використовуються для перевантаження мережевих або серверних ресурсів, мета якої – заблокувати нормальну роботу систем. Ці атаки можуть приймати різні форми, включаючи атаки на рівень мережі, такі як SYN атаки, атаки на рівень застосунків, наприклад HTTP або DNS–атаки, або одночасні комбінації різних технік.

Основний вплив DDoS–атак полягає в негайній загрозі доступності служб чи ресурсів. Перевищення обсягу запитань чи трафіку може призвести до тимчасового чи повного припинення нормальної роботи системи.

Великі фінансові втрати. Організації, особливо ті, які надають онлайн–сервіси, можуть зазнати значних фінансових втрат через втрату обсягів продажу або обслуговування клієнтів під час періоду атаки.

Вплив на репутацію. Переривання доступу може призвести до втрати довіри клієнтів та партнерів, що може суттєво нашкодити організації та завдати непоправних репутаційних втрат.

Можливість використання як прикриття. DDoS–атаки також можуть бути використані для відволікання уваги від інших видів кіберзагроз або вторгнень, що спрямовані на викрадення конфіденційної інформації чи знищення даних [9].

1.1.1.5 Атаки на системи SIEM

Атаки на системи управління подіями та інцидентами безпеки (SIEM) мають на меті обійти чи нейтралізувати ефективність SIEM у виявленні та реагуванні на кіберзагрози. Зловмисники можуть використовувати різні методи для цього, такі як намагання обману систем виявлення, знищення журналів подій або вимкнення ключових компонентів SIEM.

Зниження ефективності SIEM. Зловмисники можуть переповнити журнали подій великим обсягом непотрібної інформації або виводити події,

які приводять у збудження або хаотичний стан системи. Це може призвести до зниження ефективності виявлення загроз та реагування на інциденти, оскільки важко виділити справжні загрози серед великого потоку подій.

Обман системи виявлення. Зловмисники можуть використовувати методи маскування або імітації, щоб обійти системи виявлення SIEM та залишати свою діяльність непоміченою. Зниження точності виявлення загроз та відсутність реагування на дійсні кіберзагрози через імітацію нормальної активності.

Знищення або вимкнення компонентів SIEM. Зловмисники можуть направляти атаки на інфраструктуру SIEM, зокрема на бази даних, сервери або мережеві компоненти. Це може призвести до повного вимкнення SIEM або його ключових функцій, зменшуючи або усуваючи можливість ефективного виявлення та реагування на інциденти безпеки [5].

1.1.2 Пояснення важливості впровадження ШІ для підвищення ефективності заходів з кібербезпеки.

Використання систем на основі ШІ стає вирішальним аспектом забезпечення безпеки інформаційних систем. ШІ дозволяє ефективно виявляти та аналізувати потенційні загрози за допомогою інтелектуальних алгоритмів та аналітичних засобів.

Системи ШІ можуть визначати нові, раніше невідомі атаки, аналізувати величезні обсяги даних в реальному часі та виявляти невідповідності у поведінці систем і користувачів. Це допомагає забезпечити вчасне реагування на потенційні загрози та нейтралізовувати їх, навіть якщо вони є унікальними або розвиваються.

Інтеграція ШІ у кібербезпеку також полегшує роботу аналітиків та операторів безпеки, забезпечуючи їм інтуїтивно зрозумілі інструменти для виявлення та відстеження кіберзагроз. З використанням ШІ можливо не лише

вдосконалити захист від відомих атак, а й підготуватися до нових, що можуть з'явитися в майбутньому.

Таким чином, впровадження системи ШІ стає кроком вперед у напрямку сучасної та ефективної кібербезпеки, забезпечуючи захист від зростаючих загроз у цифровому середовищі.

ШІ відіграє ключову роль у підвищенні ефективності заходів з кібербезпеки, завдяки ряду важливих функцій та можливостей:

— Системи ШІ здатні аналізувати великі обсяги даних в реальному часі, включаючи журнали подій, мережевий трафік та інші дані безпеки. Це дозволяє вчасно виявляти та аналізувати навіть найбільш витончені кіберзагрози, які важко виявити за допомогою традиційних методів.

— ШІ використовує алгоритми машинного навчання та глибокого навчання для виявлення невідомих атак, не базуючись на заздалегідь відомих сигнатур. Це робить можливим виявлення кіберзагроз, які не мають типових сигнатур або є унікальними, забезпечуючи додатковий рівень захисту.

— Системи ШІ визначають звичайні шаблони поведінки та автоматично виявляють аномальні зміни, які можуть свідчити про кіберзагрози. Забезпечує ефективне виявлення аномалій та потенційно шкідливих дій, навіть якщо вони не відповідають стандартним підписам атак.

— ШІ використовує інтелектуальні аналітичні методи для розуміння контексту подій та взаємозв'язків між ними. Забезпечує глибше розуміння кіберзагроз, їх потенційних наслідків та шаблонів атак, що полегшує ефективну реакцію на інциденти.

— Сучасне цифрове середовище стало ареною постійних кіберзагроз, і для успішного протидії їм необхідні ефективні та швидкі заходи безпеки. У цьому контексті впровадження системи ШІ стає важливою стратегією, особливо коли мова йде про інтелектуальну відповідь на загрози.

— Системи ШІ використовують алгоритми машинного навчання та штучного інтелекту для автоматичної реакції на виявлені кіберзагрози. Це

дозволяє системам негайно впроваджувати стратегії захисту, мінімізуючи час, необхідний для реагування на інцидент.

— ШІ може аналізувати контекст і інформацію про загрозу, визначаючи оптимальні стратегії для запобігання або нейтралізації загрози. Це гарантує використання ефективних та оптимальних методів оборони безпеки.

Системи ШІ, виявляючи загрози, надають організаціям можливість автоматично реагувати на потенційні інциденти безпеки, частково або повністю усуваючи необхідність вручну втручатися. Це особливо важливо в умовах, коли швидкість реакції на загрози є ключовим аспектом успішного захисту [3,6].

Інтелектуальна відповідь Системи ШІ також означає, що можна використовувати оптимальні стратегії захисту, щоб мінімізувати можливі наслідки кіберзагроз. Враховуючи контекст і характер загроз, системи можуть приймати рішення щодо відсічення атаки, блокування аномальних дій або автоматичного відновлення до стабільного стану.

Таким чином, інтелектуальна відповідь на загрози завдяки впровадженню Системи Штучного Інтелекту не лише полегшує роботу кіберзахисників, але й робить реакцію на інциденти більш ефективною та пристосованою до постійних змін кіберпростору.

1.2 Автоматизоване виявлення аномалій

ШІ використовують алгоритми машинного навчання та глибокого навчання для аналізу поведінки систем та користувачів. Ці системи здатні визначати звичайні шаблони та автоматично виявляти аномальні зміни в них.

Сучасні загрози вже не обмежуються стандартними сценаріями, і традиційні методи виявлення можуть бути неефективними. Використання ШІ дозволяє системі "навчатися" звичайному функціонуванню та виявляти аномальність, навіть якщо це новий, раніше невідомий тип загрози.

Вчасне виявлення загроз. Системи ШІ дозволяють виявляти аномалії в реальному часі, що дозволяє реагувати на загрози негайно. Це дозволяє уникнути затримок у виявленні загроз та вживати заходи для їх мінімізації.

Навчання на основі змін поведінки. ШІ аналізує зміни у поведінці системи та користувачів, а не лише зафіксовані сигнатури відомих загроз. Системи можуть адаптуватися до нових загроз та еволюції атак, навчаючись на основі найновіших даних.

Зниження кількості хибно позитивних. Алгоритми ШІ допомагають розрізняти нормальну активність від потенційно шкідливої, що зменшує кількість хибних сповіщень. Менше хибно позитивних сигналів дозволяє фахівцям з кібербезпеки більш ефективно використовувати час та людські ресурси [3].

Автоматизоване виявлення аномалій за допомогою системи ШІ стає невід'ємною складовою сучасної кібербезпеки. Здатність систем "розуміти" навчені зразки та виявляти невизначені аномалії допомагає ефективно протидіяти навіть найвитонченішим кіберзагрозам, забезпечуючи високий рівень захисту.

Системи ШІ використовують алгоритми та моделі машинного навчання для автоматизації різноманітних завдань у сфері кібербезпеки. Вони не лише виявляють загрози, але й реагують на них, застосовуючи заздалегідь визначені та оптимальні стратегії безпеки [10].

Автоматизація процесів безпеки за допомогою ШІ означає, що система може самостійно виявляти, аналізувати та реагувати на кіберзагрози. Це включає в себе автоматичне виявлення аномалій в шаблонах, блокування атак та навіть відновлення до безпечного стану.

Важливість автоматизації процесів безпеки:

— Швидка реакція на загрози. Автоматизовані системи ШІ здатні реагувати на загрози негайно, без втручання людини. Забезпечує миттєву відповідь на кіберзагрози, що є критичним у світі швидко змінюваних атак.

— Ефективність та оптимальність. Системи ШІ можуть оптимізувати стратегії безпеки, використовуючи дані та аналізуючи звичайні шаблони та тенденції. Забезпечує ефективне використання ресурсів та максимізацію захисту при мінімальній інтервенції людини.

— Масштабованість та Постійний Моніторинг. Автоматизація дозволяє системам ШІ проводити постійний моніторинг безпеки в режимі реального часу. Забезпечує захист на різних рівнях та масштабах, враховуючи широкий спектр потенційних загроз.

Автоматизація процесів безпеки за допомогою Систем Штучного Інтелекту не тільки розширює можливості виявлення та протидії загрозам, але й забезпечує гнучкість та ефективність [6].

Висновок до першого розділу

Основна мета дослідження: вдосконалення систем запобігання інтрузій, систем управління інформаційною безпекою, та систем запобігання витоку даних у контексті центрів операцій з безпеки за допомогою методів штучного інтелекту та машинного навчання .

Дослідницьке питання: Як методи ШІ та МН можуть оптимізувати та покращити ефективність систем IPS, SIEM, та DLP у SOC?

Аналіз існуючих проблем та викликів у системах IPS, SIEM, та DLP. Вивчення та оцінка поточного стану технологій у контексті безпеки інформаційних систем. Дослідження можливостей застосування ШІ та МН у контексті цих систем. Розгляд різних методів та підходів ШІ і МН, які можуть використовуватися для покращення виявлення загроз та реакції на них.

Розробка та оцінка нових підходів на основі ШІ та МН для поліпшення цих систем. Розробка інноваційних рішень, їх тестування та оцінка ефективності у покращенні безпеки інформаційних систем.

Систематизовано існуючі проблеми та виклики, що виникають у зазначених системах безпеки.

Запропоновано нові підходи на основі ШІ та МН для покращення ефективності цих систем.

Виконано порівняння та оцінку ефективності розроблених рішень у порівнянні з традиційними методами.

Цей підхід дозволить структуровано та послідовно представити матеріал, вказуючи на важливість та новизну дослідження в контексті використання ШІ та МН у системах кібербезпеки.

2 ПЕРЕДУМОВИ ІНТЕГРАЦІЇ IPS/IDS, SIEM ТА DLP ДЛЯ ЕФЕКТИВНОГО ЗАХИСТУ ВІД КІБЕРЗАГРОЗ

У цьому розділі поглиблено розглянуто кілька ключових технологій в сфері кібербезпеки, які мають важливе значення для ефективного функціонування SOC. Сучасний кіберпростір постійно еволюціонує, і з цією постійною зміною зростає потреба в розумінні та використанні передових технологій захисту. В цьому контексті, системи IPS, IDS та SIEM відіграють ключову роль.

На початку зроблено з аналізу функцій та завдань IPS/IDS. Ці системи є фундаментальними в ідентифікації та запобіганні кібератак, і ми розглядаємо як традиційні, так і сучасні підходи до їхнього використання. Також розглядаються типові виклики та проблеми, з якими стикаються організації при впровадженні та експлуатації цих систем.

Далі переходимо до ролі систем SIEM у SOC. SIEM забезпечує комплексний погляд на безпеку, збираючи та аналізуючи дані з різних джерел, що допомагає в ранньому виявленні та реагуванні на інциденти. В цьому підрозділі ми детально розглянемо, як SIEM може підвищити ефективність виявлення та реагування на інциденти, а також які проблеми можуть виникати при їх використанні.

Оцінюємо значення DLP у контексті захисту конфіденційної інформації. Важливість DLP в сучасних організаціях не може бути переоцінена, оскільки вони допомагають ідентифікувати та запобігати несанкціонованому витоку чутливих даних. Аналізуємо як функціональні аспекти DLP, так і виклики, пов'язані з їх впровадженням та управлінням.

2.1 Аналіз функцій та завдань систем IPS/IDS в контексті кібербезпеки

Глобальний ринок продуктів інформаційної безпеки розвивається під впливом швидко зростаючого різноманіття складних і комплексних загроз, що

призводить до безпосереднього впливу на бізнес і стають затребуваними не тільки для великих і середніх, але і для малих організацій. В даний час традиційні засоби захисту, такі як міжмережевий екран і антивірус, не здатні забезпечити належний рівень захисту внутрішньої мережі організації, адже шкідливе програмне забезпечення може «замаскуватися» і відправляти пакети, які з погляду міжмережевого екрану виглядають повністю легітимними. Хоча існує безліч комерційних рішень, здатних забезпечити належний рівень захисту внутрішньої мережі організації, проте варто зупинитися на такому класі рішень, як системи виявлення вторгнень та системи запобігання вторгненням.

Рішення цього класу бувають як комерційними і з відкритим вихідним кодом. Даний клас засобів захисту відноситься до методу відстеження несанкціонованих спроб отримання доступу до ресурсів організації, що захищається, званий моніторингом управління доступом. Він націлений на виявлення та реєстрацію недоліків у безпеці внутрішньої інфраструктури – мережеві атаки, спроби несанкціонованого доступу чи підвищення привілеїв, робота шкідливого програмного забезпечення тощо. Таким чином, порівняно з мережевим екраном, що контролює тільки параметри сесії, IDS і IPS аналізують внутрішні потоки даних, що передаються, знаходячи в них послідовності бітів, які можуть представляти собою шкідливі дії або події. Крім того, вони можуть здійснювати моніторинг системних журналів та інших файлів реєстрації діяльності користувачів.

Отже, IDS – система виявлення вторгнень, призначена для реєстрації підозрілих дій у мережі, та повідомляє про них відповідального за інформаційну безпеку співробітника за допомогою передачі повідомлення на консоль управління, надсилання електронного листа, SMS–повідомлення на мобільний телефон тощо [14].

Традиційна IDS складається з сенсорів, які переглядають мережевий трафік чи журнали та передають аналізаторам, аналізатори шукають в отриманих даних шкідливий характер і у разі успішного виявлення – надсилає

результати до адміністративного інтерфейсу. Залежно від розташування IDS поділяються на мережеві і хостові. За назвою зрозуміло, що одна відстежує весь мережевий трафік того сегмента, де вона встановлена, а інша в межах єдиного комп'ютера. Для більш зрозумілої класифікації IDS необхідно виділити ще дві підмножини, які діляться за типом аналізованого трафіку. IDS, що базується на мережевих протоколах, яка аналізує комунікаційні протоколи зі зв'язаними системами або користувачами, а також IDS, заснована на прикладних протоколах, що призначені для аналізу даних, що передаються з використанням специфічних для певних додатків протоколів [1].

Шкідливу активність у аналізованому трафіку можна виявити різними способами. Тому в IDS існують такі характеристики, що відрізняють один від одного різні типи технологій IDS.

Сигнатурні IDS спрямовані на виявлення шаблонів у мережевому трафіку та функціонують аналогічно антивірусному програмному забезпеченню. Однак цей підхід має свої обмеження, оскільки сигнатури повинні бути постійно актуалізовані, і такі системи не можуть ефективно виявляти невідомі атаки. Сигнатурні IDS поділяються на два типи: ті, що відстежують шаблони у мережевих пакетах, та ті, що відстежують стан системи, порівнюючи її дії з шаблонами [14].

Системи, що відстежують стан, оперують поняттям стану системи, де зміни у роботі системи призводять до зміни її стану. IDS визначає початковий стан перед атакою та скомпрометований стан після успішної атаки.

IDS, засновані на аномаліях, працюють без використання сигнатур і навчаються «нормальній» діяльності системи. Це дозволяє їм виявляти невідомі атаки. Аномалії поділяються на статистичні, аномалії протоколів та аномалії трафіку.

IDS, що базуються на правилах, використовують логіку "ЯКЩО ситуація ТОДІ дія". Вони схожі на експертні системи, де база знань та логічні висновки застосовуються до аналізу даних. Знання в таких системах

представлене правилами, а дані вважаються фактами, до яких застосовуються правила.

IDS сповіщає про потенційно шкідливі активності, але його основна функція – це виявлення порушень безпеки після їхнього виникнення. На відміну від цього, IPS націлена на активне запобігання шкідливим діям на ранній стадії. Як підклас IDS, IPS використовує ті ж методи виявлення атак, але його основний акцент – на превентивних та проактивних заходах. IPS може працювати на різних рівнях, включаючи рівень хоста та рівень мережі. Наприклад, мережевий IPS вбудовується в інфраструктуру мережі, аналізує весь трафік і може ефективно блокувати атаки на ранніх етапах. Використання зовнішнього та внутрішнього інтерфейсів дозволяє системі взаємодіяти з трафіком, визначаючи його безпеку та, відповідно, пропускати чи блокувати його. Незважаючи на можливість роботи в режимі моніторингу, де IPS аналізує та спостерігає за трафіком без активного блокування, це може вести до втрати основної функціональності, яка полягає в запобіганні потенційно шкідливих атак. Таким чином, ефективна робота IPS вимагає балансу між моніторингом та активним запобіганням для забезпечення найвищого рівня безпеки мережі [1].

IPS може бути класифіковано на дві основні категорії в залежності від методів аналізу трафіку. Перша категорія використовує сигнатури для аналізу та порівняння трафіку з відомими атаками, тоді як друга базується на аналізі протоколів для виявлення нелегітимного трафіку, використовуючи базу знань про виявлені раніше вразливості. Завдяки другій категорії забезпечується захист від атак невідомого типу, допомагаючи виявляти нові загрози на основі аномалій у поведінці трафіку.

Щодо методів реагування на атаки, існує розмаїття стратегій. Однією з основних є блокування з'єднань за допомогою TCP–пакета з RST–прапором або за допомогою міжмережевого екрану. Ці заходи дозволяють ефективно припиняти атаку та забезпечувати безпеку мережі. Додатково, реагування може включати зміни налаштувань комунікаційного обладнання для

миттєвого реагування на загрози. Також важливим є блокування записів користувачів чи конкретного хоста в інфраструктурі, забезпечуючи ізоляцію вразливих або потенційно небезпечних елементів. Ці різноманітні методи реагування допомагають ефективно контролювати та мінімізувати вплив потенційних атак на безпеку мереж.

Оптимальним рішенням для забезпечення високого рівня захисту інфраструктури є інтегроване використання засобів IDS та IPS в рамках єдиного продукту – міжмережевого екрану. Цей інтегрований підхід дозволяє проводити глибокий аналіз мережевих пакетів для виявлення атак і активно блокувати їх. Важливо відзначити, що міжмережевий екран є лише одним шаром захисту, зазвичай розташованим за ним. Для досягнення повного комплексного захисту мережі рекомендується використовувати повний арсенал інструментів, таких як UTM, який об'єднує в собі різноманітні засоби захисту.

UTM об'єднує різні функції безпеки, такі як VPN, IPS, антивірус, засоби фільтрації та антиспаму, спільно працюючи для забезпечення повного захисту від різноманітних загроз. Це комплексний підхід до безпеки дозволяє ефективно виявляти, блокувати та відвертати потенційні атаки на різних рівнях мережевого стеку. Інтеграція різних заходів безпеки допомагає створити міцну оборонну лінію та забезпечити стійкість мережі в умовах постійно зростаючого рівня кіберзагроз.

2.2 Огляд проблем та викликів, що виникають при використанні традиційних систем IPS/IDS

Недосконаліть IDS/IPS та помилки в їх ПЗ дозволяють знаходити умови, за яких вони не здатні виявити атаку в мережевому трафіку. Серед досить давно відомих технік обходу стадії розбору стриму можна перерахувати такі:

— Нестандартна фрагментація пакетів. IDS може стикатися з ситуаціями, коли пакети розділені або мають нестандартну структуру на рівнях IP, TCP або інших протоколів. Це може зробити важкою коректну реконструкцію та аналіз трафіку, що потрібно для виявлення атак.

— Пакети з некоректними значеннями TTL або MTU: Якщо пакети мають неправильні значення TTL або MTU це може призвести до некоректного розгляду IDS. Такі аномалії можуть виникнути через помилкову конфігурацію мережі або в результаті атак на саму мережу.

— Неоднозначність сприйняття TCP–фрагментів: IDS може інтерпретувати TCP–фрагменти інакше в порівнянні з сервером або клієнтом, особливо, якщо йдеться про номери TCP SYN. Це може виникнути через нестандартні або неочікувані ситуації в мережевому трафіку.

— Підставний пакет TCP FIN із невірною контрольною сумою: IDS може помилятися при інтерпретації пакета TCP FIN, особливо якщо його контрольна сума невірна. Це може викликати невірні висновки про завершення сесії, що може бути використано для здійснення атак.

— Різний час перерв TCP–сесії між IDS і клієнтом також може стати інструментом для приховування атак.

— Що стосується етапу аналізу протоколів і нормалізації полів, багато техніки обходу WAF можуть бути запозичені і для IDS. Їх число значно більше, ось лише лише деякі з них:

— Архівування HTTP–пакету без відповідного заголовка Content–Encoding може так і залишитися нерозтисненим на стадії нормалізації, такий прийом можна часом зустріти в трафіку шкідливих програм.

— Використання рідкісних кодувань, наприклад Quoted–Printable для протоколів POP3/IMAP, також може зробити деякі сигнатури безсилими.

Варто приділяти увагу індивідуальним помилкам, характерним для кожного виробника систем виявлення вторгнень та сторонніх бібліотек, які можна знайти у відкритих баз даних помилок.

2.2.1 Детальний огляд ролі та функцій систем SIEM в SOC.

На самому базовому рівні всі рішення SIEM виконують певний рівень функцій агрегації, консолідації та сортування даних, щоб ідентифікувати загрози та дотримуватися вимог щодо відповідності даних. Хоча деякі рішення відрізняються за можливостями, більшість пропонує однаковий базовий набір функцій:

SIEM отримує дані про події з широкого спектру джерел у всій IT–інфраструктурі організації, включаючи локальні та хмарні середовища. Дані журналу подій від користувачів, кінцевих точок, додатків, джерел даних, хмарних робочих навантажень і мереж, а також дані апаратного та програмного забезпечення безпеки, наприклад брандмауерів або антивірусного програмного забезпечення, збираються, співвідносяться та аналізуються в режимі реального часу.

Деякі рішення SIEM також інтегруються зі сторонніми каналами аналізу загроз , щоб співвідносити свої внутрішні дані безпеки з раніше розпізнаними сигнатурами та профілями загроз. Інтеграція з новинами про загрози в реальному часі дозволяє командам блокувати або виявляти нові типи сигнатур атак.

Кореляція подій є важливою частиною будь–якого рішення SIEM. Використовуючи вдосконалену аналітику для виявлення та розуміння складних шаблонів даних, кореляція подій надає інформацію для швидкого визначення та пом’якшення потенційних загроз безпеці бізнесу. Рішення SIEM значно покращують середній час виявлення і середній час відповіді для команд IT–безпеки, розвантажуючи ручні робочі процеси, пов’язані з поглибленим аналізом подій безпеки.

SIEM консолідує свій аналіз в єдину центральну інформаційну панель, де групи безпеки відстежують діяльність, сортують сповіщення, ідентифікують загрози та ініціюють реагування або виправлення. Більшість інформаційних панелей SIEM також включають візуалізацію даних у

реальному часі, яка допомагає аналітикам безпеки виявляти сплески або тенденції підозрілої активності. Використовуючи налаштовані попередньо визначені правила кореляції, адміністратори можуть негайно отримувати сповіщення та вживати відповідних заходів для пом’якшення загроз, перш ніж вони матеріалізуються у більш значні проблеми безпеки [6].

Рішення SIEM є інструментом для організацій, які прагнуть відповідати різним нормативним вимогам. Завдяки автоматичному збору та аналізу даних, SIEM виконує ключову роль у зборі та перевірці відповідності в усій бізнес–інфраструктурі. Це рішення здатне надавати звіти про відповідність в реальному часі, орієнтовані на різноманітні стандарти відповідності, такі як PCI–DSS, GDPR, HIPAA, SOX та інші.

SIEM дозволяє автоматизовано генерувати звіти, спрощуючи процес управління безпекою та надаючи можливість раннього виявлення можливих порушень. Це особливо важливо для ефективного вирішення вимог відповідності та вчасної реакції на можливі загрози безпеки. Рішення SIEM часто постачаються з готовими додатками, які можуть автоматично створювати звіти, враховуючи специфіку вимог відповідності.

Узагальнено, SIEM є важливим інструментом для організацій, що прагнуть забезпечити безпеку своїх інформаційних ресурсів та одночасно відповідати стандартам і вимогам відповідності в галузі кібербезпеки.

2.2.2 Аналіз проблем та труднощів, що можуть виникати при їх впровадженні та експлуатації

SIEM системи можуть стикатися з виникненням хибно позитивних інцидентів у зв'язку з різноманітними факторами, такими як неправильна конфігурація, несумісність логічних правил або відсутність врахування специфічних контекстів. Вирішення цих питань на етапі пост–інциденту передбачає не лише технічні вдосконалення, але й глибокий аналіз внутрішньої інфраструктури та процесів моніторингу.

Важливо акцентувати увагу на оптимізації правил та налаштувань SIEM системи, щоб вони були більш адаптовані до конкретних умов і потреб організації. Врахування контекстуальних величин, таких як структура мережі, характер трафіку та нормальний образ діяльності користувачів, дозволяє зменшити ймовірність хибно позитивних сигналів.

Однак, щоб забезпечити впевненість у точності та ефективності системи, також слід інвестувати у навчання персоналу, щоб спеціалісти могли ефективно взаємодіяти з SIEM, розуміти його виводи та своєчасно реагувати на будь–які події. Такий комплексний підхід дозволяє покращити загальну реакцію на інциденти та знизити ймовірність хибно позитивних реакцій системи.

Проблеми налаштувань. Тут може йтися як про джерело даних, так і про SIEM або будь–яку іншу систему, що генерує інциденти.

Джерело. Найчастіше неправильне налаштування джерела породжує потік некоректних даних або даних правильних, але переданих з помилкою – наприклад, дублі журналів, які кілька разів приходять на корелятор.

Системний збій. Буває й так, що сповіщення антивірусом, DLP чи WAF, насправді є результатом збою чи неправильної інтерпретації вхідних даних.

SIEM. Найчастіші випадки хибно позитивних сповіщень є правила кореляції подій.

Відсутність загрози як такої. Наприклад:

— Розслідування призвело до того, що суб'єктом дії виступав легітимний користувач – інцидент, насправді, не підтвердився. Дії користувача були правомірними, але шаблони цих дій ненавмисно співпадають з поведінкою зловмисника.

— Дія була легітимною, але одночасно з цим не було жодного повідомлення з боку суб'єкта ,наприклад, коли йдеться про сканування, ініційоване співробітниками з IT відділу, і своєчасно не було наповнено перелік винятків для правил.

— Інцидент, що стався, був викликаний навчаннями, тестуванням нового правила.

У всіх цих випадках алгоритм перевірки приблизно однаковий, а ось способи усунення проблеми можуть сильно відрізнятися [4].

2.3 Розгляд функцій та важливості DLP в запобіганні витокам конфіденційної інформації

DLP можна визначити як технології, які виконують як перевірку вмісту, так і контекстний аналіз даних, надісланих через програми обміну повідомленнями, такі як електронна пошта та обмін миттєвими повідомленнями, у русі по мережі, у використанні на керованому кінцевому пристрої, і в стані спокою на локальних файлових серверах або в хмарних програмах і хмарних сховищах. Ці рішення виконують реагування на основі політики та правил, визначених для запобігання ризику ненавмисних або випадкових витоків або розголошення конфіденційних даних за межами авторизованих каналів.

Технології DLP загалом поділяються на дві категорії – Enterprise DLP та Integrated DLP. У той час як корпоративні рішення DLP є всеосяжними та упаковані в агентське програмне забезпечення для настільних комп’ютерів і серверів, фізичних і віртуальних пристроїв для моніторингу мереж і трафіку електронної пошти або програмних пристроїв для виявлення даних, Integrated DLP обмежено безпечними веб–шлюзами, безпечними шлюзами електронної пошти, продукти шифрування електронної пошти, платформи керування корпоративним контентом, інструменти класифікації даних, інструменти виявлення даних і брокери безпеки доступу до хмари [12].

Після обробки вмісту існує кілька методів аналізу, які можна використати для ініціювання порушення політики, зокрема:

— На основі правил або регулярних виразів, одна з найпоширеніших технік аналізу, яка використовується в DLP, полягає в тому, що механізм

аналізує вміст за певними правилами, такими як 16–значні номери кредитних карток. Ця техніка є чудовою для першого проходу. фільтр, оскільки правила можна швидко налаштувати та обробити, хоча вони можуть бути схильні до високого рівня помилкових спрацьовувань без перевірки контрольної суми для визначення дійсних шаблонів.

— Відбитки бази даних також відомий як точна відповідність даних, цей механізм шукає точні збіги з злитих баз даних. Хоча активні підключення до бази даних впливають на продуктивність, це варіант для структурованих даних із баз даних.

— Точна відповідність файлу вміст файлу не аналізується; однак хеші файлів відповідають точним відбиткам. Забезпечує низьку кількість помилкових спрацьовувань, хоча цей підхід не працює для файлів із кількома схожими, але не ідентичними версіями.

— Часткова відповідність документів шукає повну або часткову відповідність у певних файлах, наприклад у кількох версіях форми, заповнених різними користувачами.

— Статистичний аналіз використовує машинне навчання або інші статистичні методи, як–от байєсівський аналіз, щоб ініціювати порушення правил у безпечному вмісті. Для сканування потрібен великий об’єм даних, чим більше, тим краще, інакше можуть виникати помилкові спрацьовування та негативні результати.

— Попередньо створені категорії з правилами та словниками для поширених типів конфіденційних даних, таких як номери кредитних карток/захист PCI, HIPAA тощо.

Сьогодні на ринку існує безліч методів, які забезпечують різні види перевірки вмісту. Варто взяти до уваги одне: хоча багато постачальників DLP розробили власні механізми вмісту, деякі використовують технологію сторонніх розробників, яка не призначена для DLP. Наприклад, замість створення відповідності за шаблоном для номерів кредитних карток, постачальник DLP може отримати ліцензію на технологію від постачальника

пошукової системи для зіставлення за шаблоном номерів кредитних карток. Оцінюючи рішення DLP, зверніть пильну увагу на типи шаблонів, виявлених кожним рішенням у порівнянні з реальним масивом конфіденційних даних, щоб підтвердити точність механізму вмісту [12].

Разом із великим числом переваг, DLP системи також мають свої власні виклики та обмеження, які можуть впливати на їхню ефективність [13].

— Недостатня точність виявлення витоків. На шляху впровадження DLP може виникнути проблема хибно позитивних сигналів, коли система помилково ідентифікує нормальні дії як потенційні загрози. Надмірна строгість або неадекватність налаштувань правил може призвести до того, що рутинні операції, такі як обмін файлами або електронна пошта, визначатимуться як небезпечні, що може викликати непокій серед користувачів та спричинити втрату часу на перевірку хибних сигналів.

— Неспроможність виявлення зашифрованих витоків. Специфікації конфіденційних даних часто включають захист за допомогою шифрування. DLP системи можуть виявляти витоки лише в тому випадку, якщо дані передаються незашифрованими. Застосування шифрування може ускладнити процес виявлення, зменшуючи ефективність DLP.

— Відсутність оновлень аналізаторів. Деякі DLP системи можуть бути вразливі через недостатню частоту або швидкість оновлення аналітичних модулів. Якщо система не отримує своєчасні поновлення для виявлення нових шаблонів атак чи еволюції загроз, вона може залишитися неефективною перед сучасними кіберзагрозами.

— Важкість виявлення невідомих шаблонів. Інновації у кіберзлочинності можуть призводити до використання нових, непередбачених шаблонів атак. Недостатня гнучкість аналітичних інструментів DLP може ускладнити виявлення таких новаторських загроз.

— Неадекватна реакція на нові загрози: Якщо DLP система не має механізмів швидкої реакції на нові загрози, це може призвести до великої затримки у впровадженні заходів безпеки. Наприклад, система може

неадекватно відреагувати на нові варіації шкідливих програм або методи обходу захисту.

Висновок другого розділу

У розділі ми ретельно розглянули критичні компоненти кібербезпеки: системи IPS/IDS, SIEM та DLP, які є невід'ємною частиною ефективного функціонування SOC. Дослідження цих систем дозволило нам зрозуміти їх важливість, складнощі та виклики, з якими стикаються організації при їх впровадженні та експлуатації.

Аналіз систем IPS та IDS підкреслив їх значення у виявленні та запобіганні кібератак, а також виявив проблеми, які виникають у контексті постійно змінюваного кіберпростору. Проведена оцінка показує як вони інтегруються з іншими інструментами безпеки для створення більш міцної оборони.

Далі, розгляд систем SIEM пролив світло на їхню критичну роль у зборі, аналізі та відповіді на безпекові інциденти. Ми встановили, що вони надають цінні дані для більш швидкого виявлення та реагування на потенційні загрози, але також зіткнулись з проблемами, пов'язаними з їх складністю та управлінням.

На завершення, дослідження DLP показало їх важливість у захисті конфіденційності даних та у запобіганні витоку чутливої інформації. Ми виявили, що, хоча вони надзвичайно важливі для забезпечення безпеки даних, існують виклики, пов'язані з їхньою інтеграцією та ефективністю.

3 ПЕРСПЕКТИВИ ВИКОРИСТАННЯ ШІ В SOC ЦЕНТРАХ НА ОСНОВІ УНІФІКОВАНОЇ СИСТЕМИ

В розділі надано комплексний погляд на технічні і стратегічні перспективи впровадження ШІ в SOC. Розглядається класичний підхід до кібербезпеки, встановлюючи фундамент для подальшого порівняння з більш сучасними методами. Проводиться детальний аналіз напів–уніфікованих та уніфікованих систем, оцінюючи їхні переваги та недоліки у контексті застосування представлення дерев подій, які є критичними для розуміння та відстеження кіберзагроз [6]. На дано приклади атак та методи їх виявлення та нейтралізації з використанням ШІ. Важливою частиною розділу є обговорення збору та аналізу даних для ефективного реагування на інциденти.

3.1 Методи поліпшення кібербезпеки в SOC та вектори розвитку ШІ

У динамічному середовищі кібербезпеки, де інформаційні системи піддаваються постійним загрозам та різнорідним атакам, виявлення та ефективне реагування на кіберзагрози стає надзвичайно важливим завданням. На сьогоднішній день існує різноплановий підхід до вирішення цих завдань, який охоплює класичні методи аналізу подій, напів–уніфіковані системи з штучним інтелектом та повністю уніфіковані автономні системи.

Класичний метод аналізу подій ґрунтується на традиційних підходах до обробки інформації та виявлення загроз без безпосереднього використання штучного інтелекту. Аналітики використовують експертний досвід та збирають і аналізують дані з різних систем для ідентифікації аномальностей та потенційних загроз. Традиційні методи аналізу синхронізовані з визначенням шаблонів та використанням правил виявлення аномалій.

У напів–уніфікованій системі з штучним інтелектом відбувається еволюція в здатності виявлення кіберзагроз через використання алгоритмів машинного навчання та штучного інтелекту. Ця система інтегрує дані з різних

джерел, проте реагування на інциденти залишається під управлінням людей. Алгоритми ШІ використовуються для аналізу динаміки змін у системі, розпізнавання ненормальних шаблонів та підвищення точності виявлення загроз.

Повністю уніфікована автономна система є етапом вперед порівняно з попередніми підходами. Вона використовує передові технології штучного інтелекту для аналізу, виявлення та автоматичного реагування на кіберзагрози. Система самостійно інтегрує та обробляє дані з трьох джерел, використовуючи адаптивні алгоритми машинного навчання для прийняття швидких та ефективних рішень в реальному часі. Цей підхід дозволяє автоматизувати реагування на загрози, зменшуючи час реакції та підвищуючи загальний рівень безпеки.

3.1.1 Класичний підхід до кібербезпеки

Класичний підхід ґрунтується на досвіді та експертизі фахівців з кібербезпеки, які відповідають за виявлення, аналіз та реагування на потенційні кіберзагрози. Основною метою цього підходу є створення систем, які здатні вчасно реагувати на виявлені аномалії чи загрози безпеки.

Основні принципи класичного підходу:

— Ручний аналіз подій. Кібербезпечні аналітики вручну переглядають та аналізують дані, зокрема системні журнали, мережевий трафік та інші інформаційні ресурси.

— Експертний досвід. Використання експертного досвіду у визначенні аномальних ситуацій та потенційних загроз.

— Правила та шаблони. Визначення правил та шаблонів для виявлення типових загроз на основі попередніх інцидентів.

— Локальний контроль. Аналіз та реагування на загрози проводиться на локальному рівні, що дозволяє забезпечити контроль над безпекою системи в конкретному місці.

Класичний підхід виявляється своєчасним та ефективним у виявленні традиційних загроз, основаних на відомих атаках та вразливостях. Однак у світі постійно зростаючих технологічних викликів та нових форм кіберзагроз, цей підхід може виявитися менш ефективним у виявленні складних та еволюційних видів атак. Далі, розглянемо більш детально переваги і недоліки класичного підходу до кібербезпеки.

Переваги класичного підходу до кібербезпеки:

— Експертний досвід. Високий рівень експертного досвіду аналітиків з безпеки дозволяє виявляти та реагувати на унікальні загрози, які можуть викликати великий ризик.

— Гнучкість. Можливість вручну адаптувати правила та стратегії реагування, що робить підхід гнучким та придатним для різноманітних умов.

— Локальний контроль. Здатність забезпечити локальний контроль над реагуванням на інциденти, що важливо для децентралізованих систем.

— Відсутність залежності від алгоритмів. Відсутність великої залежності від складних алгоритмів машинного навчання, що спрощує розгортання та підтримку.

Недоліки класичного підходу до кібербезпеки:

— Низька швидкість реакції. Ручний аналіз та реакція можуть бути повільними, особливо при великому обсязі даних або розподіленому характері системи.

— Вразливість до нових загроз. Підхід може бути менш ефективним у виявленні нових та еволюційних видів кіберзагроз, оскільки він базується на відомих шаблонах.

— Високий рівень залежності від експертів. Необхідність у постійному вдосконаленні експертного досвіду для забезпечення ефективності системи.

— Неспроможність автоматизації. Важкість автоматизації процесів через велику кількість ручних етапів та залежність від людського фактору.

— Обмежена складність аналізу. Може бути непросто виявляти та аналізувати складні атаки, що використовують неочікувані методи або комбінації загроз.

Хоча класичний підхід має свої переваги, особливо у виявленні унікальних загроз та в локальному контролі, його обмеженості у швидкості реакції та адаптації до нових загроз можуть стати критичними [9].

3.1.2 Напів–уніфікована система зі ШІ

Напів–уніфікована система зі штучним інтелектом (ШІ) представляє сучасний підхід, об'єднуючи традиційний аналіз подій із передовими технологіями машинного навчання. Використовуючи алгоритми ШІ, система автоматично виявляє аномалії та загрози, зменшуючи залежність від ручного аналізу та надаючи можливість більш швидко реагувати на потенційні інциденти. Така система створює міцний союз між експертними знаннями та автоматизованими засобами машинного навчання. Це дозволяє більш ефективно виявляти та реагувати на загрози, зменшуючи вплив людського фактору та прискорюючи час реакції на інциденти.

Основні компоненти напів–уніфікованої системи:

— МН та алгоритми ШІ. Використання розширених алгоритмів машинного навчання для аналізу великих обсягів даних та виявлення складних шаблонів.

— Автоматизоване виявлення аномалій. Застосування ШІ для автоматичного виявлення аномалій та надто складних для виявлення шаблонів.

— Експертна оцінка та корегування. Залучення кібербезпекових експертів для остаточного аналізу та корегування результатів системи.

— Адаптивність до нових загроз. Здатність системи швидко адаптуватися до нових та еволюційних видів кіберзагроз за рахунок навчання на льоту.

3.1.2.1 Представлення дерева подій напів–уніфікованої системи

Концепція напів–уніфікованої системи зі штучним інтелектом виявляється дуже перспективною. Ця система об'єднує в собі найкращі аспекти різних систем кіберзахисту, таких як DLP, SIEM, IPS/IDS, і використовує ШІ для обробки та аналізу великих обсягів даних.

Це вирішує проблеми недоліків, що можуть виникнути при використанні окремих систем. Але як саме ця напів–уніфікована система функціонує у виявленні та реагуванні на кіберзагрози? Розглянемо цей процес на основі концепції дерева подій.

Початкова подія: Надходження невідомого мережевого трафіку.

Рівень 1. Виявлення аномалії:

— SIEM. Виявлення надмірного трафіку за допомогою аналізу подій та порівняння зі звичайним шаблоном.

— IPS. Визначення, чи є це вторгнення або лише незвичайний обсяг трафіку.

Рівень 2. Аналіз аномалії:

— ШІ. Використання алгоритмів машинного навчання для аналізу динаміки змін та виявлення нормальної/ненормальної активності.

Рівень 3. Визначення типу загрози:

— DLP. Виявлення і зупинка витоку конфіденційної інформації.

— SIEM та ШІ. Кореляція аномалій з базою відомих загроз для визначення їхнього типу.

Рівень 4. Повідомлення та прийняття рішення:

— Аналітик отримує повідомлення та контекстну інформацію щодо виявленої загрози. Вибирає стратегію реагування на основі отриманої інформації.

Рівень 5. Реагування на загрозу:

— Аналітик. Введення обраної стратегії в дію, наприклад, блокування трафіку чи ізоляція вразливих ресурсів.

Рівень 6. Моніторинг та аналіз результатів:

— SIEM та ШІ. Ведення журналів та аналіз після реагувальних результатів для вдосконалення стратегій управління інцидентами.

Рівень 7. Машинне навчання та адаптація:

— ШІ. Використання результатів для навчання системи на основі нових даних та адаптації до змін в кіберзагрозах та зміни в структурі системи.

Це дерево подій представляє напів–уніфіковану систему, яка об'єднує різні системи кіберзахисту для виявлення, аналізу та реагування на кіберзагрози. Аналітик, взаємодіючи з усіма компонентами системи та використовуючи дані з ШІ, визначає стратегію реагування на загрозу, роблячи при цьому відокремлені рішення засновані на контексті та аналітиці.

3.1.2.2 Переваги та недоліки напів–уніфікованої системи

У даній роботі проаналізовано переваги та недоліки напів–уніфікованої системи з ШІ в контексті кіберзахисту. Детальний аналіз цієї концепції дозволить краще розуміти, як такі системи можуть забезпечити ефективний захист у складних та непередбачуваних умовах кіберпростору.

Плюси:

— Інтеграція різних технологій. Можливість використовувати кращі аспекти кіберзахисту різних систем, поєднуючи їх для уніфікованого підходу.

— Автоматизація та штучний інтелект. Використання алгоритмів машинного навчання та ШІ дозволяє автоматизувати виявлення та реагування на загрози в режимі реального часу.

— Гнучкість налаштування. Можливість легко додавати нові компоненти чи системи для вдосконалення функціональності або реагувати на зміни загроз.

— Контекстуальний аналіз дозволяє вдосконалювати точність виявлення загроз та зменшує кількість фальсифікацій.

— МН та Адаптація це – навчання на основі нових даних та адаптуватися до змін в кіберзагрозах, що поліпшує її ефективність з часом.

Мінуси:

— Впровадження такої системи може бути складним завданням, оскільки вона вимагає інтеграції різних технологій та великої кількості даних.

— Витрати на обслуговування системи може вимагати значних коштів через потребу в постійному покращеннях, оновленнях та навчанні персоналу.

— Велика кількість даних та алгоритмів може створити ризик фальсифікації, де система може помилятися або реагувати на хибні сигнали.

— Рішення, які приймаються аналітиком, можуть бути залежними від людського фактору, інтуїції та досвіду, що може призвести до непередбачуваних результатів.

3.1.3 Уніфікована система

Сучасні підходи до забезпечення безпеки, такі як використання систем DLP, та IPS/IDS, надають значні переваги, проте часто працюють на власних рівнях і відтак не завжди забезпечують повний огляд кіберзахисту.

У цьому контексті концепція створення уніфікованої системи збору та аналізу інформації на базі ШІ виявляється вельми актуальною. Ця система може функціонувати як інтегратор, здатний збирати дані з різних систем кібербезпеки, відтак створюючи єдиний погляд на кіберзагрози та інциденти.

Використання ШІ дозволяє автоматизувати процеси збору та аналізу даних, створюючи систему, що може ефективно виявляти аномалії на різних рівнях інфраструктури. Наприклад, DLP системи можуть ідентифікувати витік чутливої інформації, SIEM аналізує величезний потік подій, а IPS/IDS виявляє та блокує атаки у мережевому трафіку. Уніфікована система, завдяки ШІ,

здатна об'єднати ці дані, створюючи повний обсяг інформації та надаючи можливість зробити комплексні висновки та приймати оперативні рішення.

Такий підхід до кіберзахисту не лише спрощує управління безпекою, але й забезпечує миттєву реакцію на загрози, підвищуючи рівень впевненості в кіберпросторі. ШІ–доповнена уніфікована система стає необхідним інструментом у боротьбі зі сучасними кіберзагрозами, забезпечуючи комплексний та інтегрований підхід до кібербезпеки.

3.1.3.1 Представлення дерева подій

Дерево подій на основі уніфікованої системи на базі ШІ.

Початкова подія: Масове надходження трафіку на мережевий вузол.

Рівень 1. Виявлення аномалії:

— SIEM. Аналіз подій для виявлення надмірного трафіку, аномалій та підозрілої активності.

— IDS/IPS. Визначення підозрілої активності на різних рівнях мережі.

Рівень 2. Аналіз аномалії та виявлення загрози:

— ШІ. Аналіз текстових даних для розуміння контексту та виявлення надмірної активності.

— SIEM. Кореляція текстових та технічних даних для визначення загроз та їхнього контексту.

Рівень 3. Виявлення витоку інформації та взаємодія з ШІ:

— DLP. Виявлення витоку чутливої інформації та створення звіту в текстовому форматі.

— ШІ. Генерація повідомлень та звітів у зрозумілій формі для аналітиків.

Рівень 4. Повідомлення та корекція загроз з урахуванням хибно–позитивних інцидентів:

— SIEM. Повідомлення адміністратора та подання відомостей щодо виявленої загрози.

— IDS/IPS. Автоматичне введення заходів, з урахуванням аналізу хибно позитивних сповіщень для уникнення помилкових реакцій.

Рівень 5. Аналіз впливу та взаємодія з ШІ:

— ШІ .Участь в діалозі з аналітиками для додаткової інформації та розробки стратегій відновлення.

— SIEM .Оцінка впливу та визначення стратегій відновлення.

Рівень 6. МН та адаптація:

— ШІ. Використання результатів аналізу для навчання системи та адаптації алгоритмів для виявлення нових загроз.

Рівень 7. Моніторинг та оптимізація:

— SIEM та ШІ. Постійний моніторинг та оптимізація системи з урахуванням результатів аналізу текстових та технічних даних.

Уніфікована система кіберзахисту, яка поєднує в собі елементи ШІ, системи управління інформацією та SIEM, IDS/IPS, DLP забезпечує комплексний та ефективний захист від кіберзагроз. Дерево подій відображає системний підхід до виявлення, аналізу та відповіді на потенційні загрози, зокрема, враховуючи інтелектуальний аналіз через мовленнєву модель ШІ [9].

Основною перевагою такої системи є інтеграція різноманітних технологій, що дозволяє використовувати їх переваги для максимально ефективного виявлення та реагування на загрози. Використання мовленнєвої моделі у співробітництві з ШІ спрощує процес збору інформації та взаємодії з аналітиками, що робить систему більш доступною та зручною для користувачів.

Важливим аспектом є також управління хибно позитивних сповіщень на кожному етапі, що дозволяє уникнути помилкових реакцій та оптимізувати виявлення загроз. Однак, важливо враховувати і підтримувати систему, забезпечуючи її актуальність та адаптацію до нових видів загроз.

У підсумку, уніфікована система на основі мовленнєвої моделі ШІ є потужним інструментом для забезпечення безпеки в кіберпросторі, забезпечуючи не лише технічний захист, але і інтелектуальну взаємодію з користувачем та високий рівень адаптації до змін у загрозах.

3.1.3.2 Переваги та недоліки

У цьому розділі проведено аналіз переваг та недоліків уніфікованої системи кіберзахисту, звертаючи увагу на те, як ці системи можуть поліпшити ефективність виявлення та реагування на загрози, а також на важливість управління можливими негативними аспектами в їхньому застосуванні.

Переваги:

— Інтеграція технологій. Здатність об'єднувати різні технології для комплексного та ефективного кіберзахисту.

— Інтелектуальний аналіз за допомогою мовленнєвої моделі ШІ. Зручний та інтуїтивно зрозумілий інтерфейс для спілкування з системою, що полегшує взаємодію з аналітиками та забезпечує більш ефективний аналіз подій.

— Автоматизація та машинне навчання. Використання ШІ для автоматизації виявлення та адаптації до нових загроз через машинне навчання.

— Управління хибно позитивними сповіщеннями. Активне управління помилковими спрацюваннями, що забезпечує точність та відсіює помилкові сигнали.

— Зручність та ефективність. Забезпечення доступності та зручності взаємодії з системою для користувачів, що прискорює процес реагування на інциденти.

Недоліки:

— Складність впровадження. Великий обсяг інтеграції технологій та потреба у висококваліфікованому персоналі може зробити впровадження складним та витратним.

— Витрати на обслуговування та навчання. Постійне обслуговування, апгрейди навчання персоналу можуть призвести до додаткових витрат.

— Конфіденційність та захист даних. Обробка великого обсягу чутливих даних може породити питання щодо захисту особистої інформації.

— Залежність від штучного інтелекту. Рішення, прийняті системою на основі ШІ, можуть бути не завжди передбачуваними та залежати від алгоритмів машинного навчання.

— Необхідність постійного оновлення. Зміна кіберзагроз та технологічного ландшафту вимагає постійного оновлення системи для ефективного функціонування.

Уніфікована система кіберзахисту, оснащена ШІ та мовленнєвої моделі, представляє собою сучасний та потужний інструмент у боротьбі зі зростаючими кіберзагрозами. Переваги такої системи, включаючи інтеграцію технологій, інтелектуальний аналіз та ефективне управління загрозами, виходять за межі традиційного підходу до кібербезпеки.

Однак важливо розпізнати, що впровадження та ефективне використання таких систем потребує значних ресурсів, а також уважного управління та постійного оновлення. Збалансоване розглядання переваг та недоліків допомагає розуміти, які виклики можуть виникнути та як їхнє вирішення може сприяти створенню надійного кіберзахисту в епоху високих технологій

3.1.4 Порівняння напів–уніфікованої та уніфікованої систем кіберзахисту

Виходячи з представлення дерев подій для двох типів систем, відрізняються за ступенем автоматизації, рівнем деталізації та участю людського фактору (табл. 3.1).

Таблиця 3.1 – Порівняння характеристик систем відповідно до дерев подій

Характеристика	Уніфікована система	Напів–уніфікована система
Ступінь автоматизації	Всі рівні управління та взаємодії здійснюються автоматично за допомогою системи штучного інтелекту (ШІ) та інших технічних засобів. Людський фактор обмежений.	Включає в себе рівень взаємодії з аналітиками та аналіз та прийняття рішень людьми на певних етапах.
Участь людини	Залучення аналітиків або адміністраторів можливе, але вони переважно взаємодіють з системою лише для стратегічного управління та навчання ШІ.	Включає активну участь аналітиків та адміністраторів на різних етапах процесу, зокрема в процесах аналізу, визначення загроз, прийняття рішень та реагування на інциденти
Адаптація та навчання	Зазвичай включає механізми машинного навчання для адаптації до нових загроз та змін в середовищі без активної участі людини.	Вимагає більш активного втручання аналітиків та адміністраторів для навчання системи та розробки стратегій відновлення.
Моніторинг та оптимізація	Проводить постійний моніторинг та оптимізацію системи без значного втручання адміністраторів	Включає більш активний контроль та аналіз результатів для вдосконалення стратегій управління інцидентами.

Як видно з табл. 3.1, уніфікована система має більш високий рівень автоматизації та меншу залежність від людського фактору порівняно з напів–уніфікованою системою.

3.1.5 Приклади атак та відповідні методи

Розглянувши потребу в уніфікації систем кіберзахисту, особливо на фоні великої різноманітності систем, які вже використовуються, важливо звернутися до конкретних типів атак та технічних засобів їх виявлення. Однією з важливих складових цього підходу є використання систем DLP, SIEM та IPS/IDS, які працюють на різних рівнях захисту та забезпечують комплексний огляд кіберпростору.

Конкретні приклади атак та відповідні методи їх виявлення за допомогою зазначених систем наведено в табл. 3.2

Таблиця 3.2 – Типи атак і способи їх виявлення у різних системах

№	Тип Атаки	DLP	SIEM	IPS/IDS
1	SQL Injection	Виявлення спроб SQL–ін'єкцій в тексті	Аналіз надзвичайних запитів до бази даних	Виявлення SQL–ін'єкцій у мережевому трафіку
2	Cross–Site Scripting (XSS)	Виявлення вбудованих скриптів у веб–сторінках	Аналіз аномальних HTTP запитів та відповідей	Виявлення XSS–атак в HTTP трафіку
3	Чоловік–по–середині (Man–in–the–Middle)	Виявлення незвичайного маршрутизації або змін у трафіку	Аналіз змін у мережевих з'єднаннях та трафіку	Виявлення аномальної активності у з'єднаннях
4	Фішинг	Виявлення підозрілих електронних листів та веб–сайтів	Аналіз підозрілих електронних листів та URL	Виявлення фішингових спроб в мережевому трафіку
5	Атака грубою силою на паролі	Виявлення великої кількості невдалих спроб входу	Аналіз аномальних спроб входу та блокування	Виявлення великої кількості невдалих спроб входу
6	Витік конфіденційної інформації	Виявлення надто великої кількості передачі конфіденційної інформації	Моніторинг подій, пов'язаних з передачею конфіденційних даних	Виявлення надто інтенсивного передачі конфіденційних даних

Продовження таблиці 3.2

7	Атака на вразливість	Виявлення використання відомих або невідомих вразливостей	Аналіз подій, пов'язаних із спробами використання вразливостей	Виявлення аномалій у мережевому трафіку, що може вказувати на експлуатацію вразливостей
8	Зловживання привілеями	Виявлення несправедливого або надто активного доступу	Аналіз аномально високих привілеїв.	Виявлення несправедливого використання привілеїв
9	Внутрішні загрози та недобросовісна діяльність	Моніторинг невичерпного обігу чи доступу до конфіденційної інформації	Аналіз незвичайної активності легітимних користувачів	Виявлення надто агресивної чи недобросовісної діяльності

DLP системи спрямовані на захист конфіденційної інформації, виявляючи та блокуючи спроби втрати даних. SIEM системи відстежують та аналізують події в мережі для виявлення аномалій та потенційних загроз. IPS/IDS системи виявляють та запобігають вторгненням в мережу.

Уніфікована система, побудована на ШІ, інтегрує ці функціональності для створення цілісного підходу до кіберзахисту. Вона забезпечує автоматизоване збір та аналіз даних з різних джерел, надаючи можливість оперативного реагування на загрози.

Зазначені системи взаємодіють та доповнюють одна одну, створюючи повний обсяг інформації для ефективного керування кіберзахистом. Ця система дозволяє виявляти атаки на різних рівнях, а також взаємодіяти з Інтернетом Речей, хмаровими сервісами та іншими аспектами сучасного інформаційного середовища.

ШІ в уніфікованій системі робить можливим використання передових алгоритмів машинного навчання для виявлення несподіваних аномалій та нових типів загроз. Це допомагає удосконалити адаптивність системи до змінюючись обставин та стратегій кіберзлочинців.

3.2 Схема SOC центру на основі уніфікованої системи

Системи кіберзахисту грають важливу роль у зборі інформації про можливі загрози та підготовці до майбутніх атак. Розглянемо роль кожної системи, зокрема уніфікованої системи (рис. 3.1) зі ШІ:

— DLP спрямований на виявлення та контроль збуту конфіденційної інформації. На етапі підготовки та розвідки, DLP може моніторити обмін даними, щоб виявити будь–які надмірні або підозрілі пересилання інформації, які можуть служити джерелом розвідки або підготовки до атаки.

— SIEM відіграє ключову роль у зборі та аналізі подій з різних джерел. На етапі розвідки SIEM може агрегувати дані про події з мережевих пристроїв, систем журналів та інших джерел для виявлення ненормальних змін, що можуть свідчити про розвідку.

— IPS/IDS може виявляти аномальний мережевий трафік та спроби вторгнення. На етапі розвідки ці системи можуть виявити підозрілі дії, такі як надмірні запити на ресурси, які можуть служити ознаками підготовки до атаки.

— ШІ .Уніфікована система з ШІ може використовувати алгоритми машинного навчання для аналізу шаблонів розвідки та передаються ШІ дані. Наприклад, вона може аналізувати великі обсяги даних, щоб виявити зміни у звичайній активності, які можуть вказувати на розвідку або підготовку до атак.

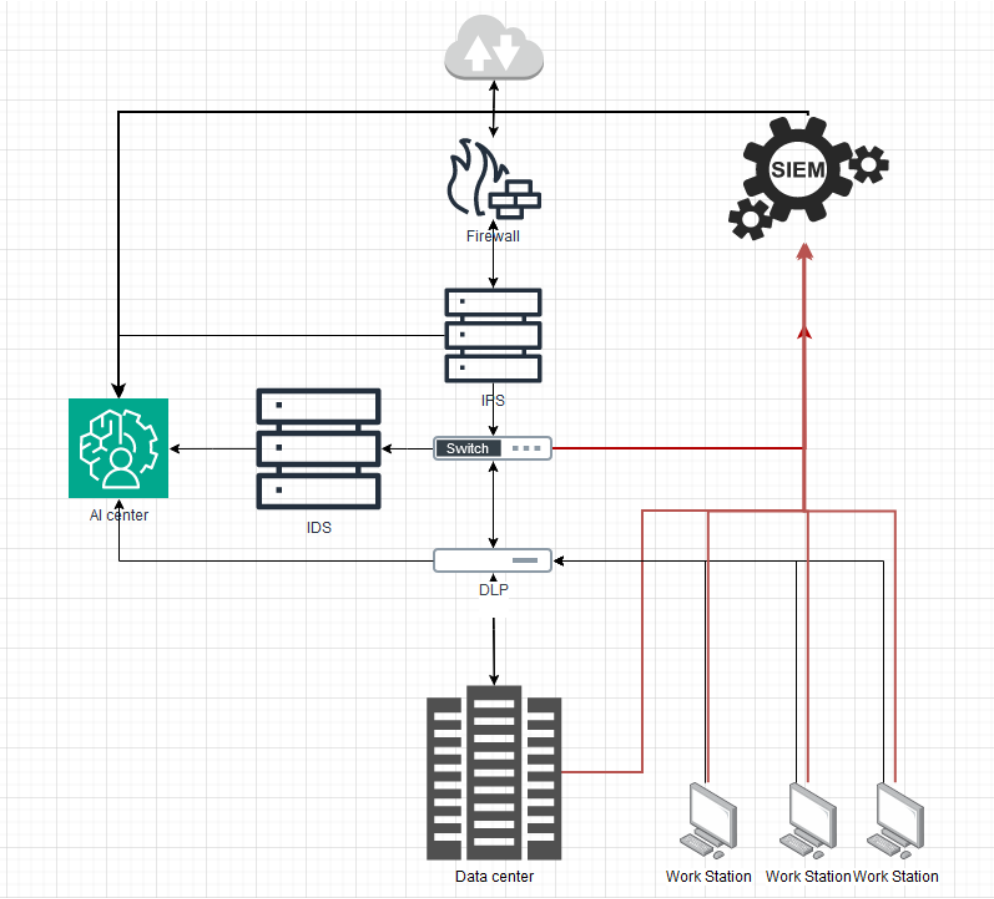

Рисунок 3.1 – Схема SOC центру на основі уніфікованої системи

Уніфікована система з ШІ використовує аналітику для виявлення аномалій, враховуючи контекстуальні дані та широкий спектр інформації. Наприклад, ШІ може реагувати на надзвичайне підвищення трафіку, яке може свідчити про підготовку атаки або розвідку.

Розглянемо, як вона може збирати дані для визначення, чи це атака, чи можливе хибне спрацювання:

— Збір даних. Система на основі ШІ може інтегруватися з різноманітними джерелами даних, такими як системи журналювання, IPS/IDS, DLP, SIEM, а також з зовнішніми джерелами інформації, наприклад, з базами даних вразливостей або обліковими записами загроз.

— Моніторинг поведінки. Система аналізує нормальну поведінку системи, користувачів та мережі. Алгоритми машинного навчання визначають типові шаблони активності та розпізнають невідомі чи підозрілі зміни.

— Виявлення аномалій. ШІ використовує алгоритми для виявлення аномалій, які можуть вказувати на атаку або несправності в системі. Це може бути збільшення трафіку, незвичайні запити, або інші незвичайні події.

— Контекстуальний аналіз щоб визначити, чи це атака, система враховує контекст даних. Наприклад, може враховувати звичайну активність певного користувача, типові часи доступу до ресурсів, та інші параметри.

— Кореляція подій. ШІ може співпрацювати з системою управління інформацією та подій для кореляції подій і виявлення взаємодії різних аномалій, що дозволяє виявити складні атаки.

— МН та адаптація. ШІ навчається на основі нових даних та подій, адаптуючи свої алгоритми до змін в кіберзагрозах та зміни в структурі системи.

— Реагування та впровадження заходів. ШІ та IPS/IDS у разі виявлення атаки може автоматично впроваджувати заходи безпеки, такі як блокування атакуючого трафіку або ізоляція вразливих систем.

— Підтримка та Управління. ШІ та SIEM може вести журнали подій та результатів, які використовуються для аналізу та покращення стратегій кіберзахисту в майбутньому.

Ця система інтегрується з різноманітними джерелами даних, такими як системи журналювання, IPS/IDS, DLP, SIEM, а також отримує інформацію з зовнішніх джерел, наприклад, баз даних вразливостей або облікових записів загроз [7,12].

На першому етапі, система моніторить нормальну поведінку системи, користувачів та мережі, аналізуючи звичайні шаблони активності. Застосовуючи алгоритми машинного навчання, вона розпізнає типові шаблони та виявляє невідомі чи підозрілі зміни.

ШІ виявляє аномалії, враховуючи контекст даних, таких як звичайну активність користувачів, типові часи доступу до ресурсів та інші параметри. У співпраці з SIEM, системою управління інформацією та подій, вона корелює події та взаємодіє з різними аномаліями для виявлення складних атак.

Уніфікована система з ШІ може виявити атаки та аномалії на основі зібраних даних, враховуючи їхній контекст та аналізуючи їхні шаблони. Вона також автоматично реагує на виявлені загрози, впроваджуючи заходи безпеки, такі як блокування атакуючого трафіку чи ізоляція вразливих систем. Цей комплексний підхід дозволяє системі ефективно виявляти та протидіяти кіберзагрозам на етапі їхньої підготовки та розвідки.

3.3 Оцінки ефективності систем що входять до складу SOC

Оцінка ефективності систем є критичною для забезпечення захисту від сучасних кіберзагроз та ефективного управління кібербезпекою.

У цьому контексті розробка та впровадження метрик ефективності стають необхідністю для оцінки працездатності та готовності системи відповідати на виклики кіберзахисту. Проаналізовані метрики, такі як точність виявлення загроз, швидкість реакції на загрози та інші, створюють комплексний підхід до оцінки ефективності уніфікованих систем у сфері кібербезпеки.

Запропоновані визначені метрики та їхні формули, надаючи високорівневий огляд їхнього значення для оцінки роботи системи. Крім того, розглядаються додаткові аспекти, які можуть допомогти в повноті оцінки та виявленні областей для подальшого вдосконалення. Необхідність вдосконалення кіберзахисту стає актуальною в умовах постійного розвитку загроз та технологій, і цей аналітичний документ спрямований на підтримку фахівців у сфері кібербезпеки при оцінці та вдосконаленні їхніх систем.

Метрика ефективності уніфікованої системи на основі ШІ в Кіберзахисті:

— Точність виявлення загроз. Визначає, наскільки ефективно система виявляє реальні кіберзагрози і визначається як кількість вірно виявлених загроз розділити на загальну кількість загроз

— Швидкість реакції на загрози. Вимірює ефективність системи у швидкому реагуванні на виявлені загрози. Основною характеристикою є час від виявлення загрози до прийняття відповідних заходів.

— Відсоток хибних спрацювань. Визначає, наскільки часто система генерує помилкові сигнали про загрози і визначається як відношення кількості хибно позитивних подій на загальну кількість виявлених подій

— Точність системи ідентифікації витоків даних. Вимірює, наскільки точно система визначає витоки чутливої інформації – відношення кількості вірно ідентифікованих витоків до загальної кількості виявлених витоків

— Час відновлення після загрози. Вказує на ефективність стратегій відновлення та реакції після інциденту.

— Кількість автоматично вирішених загроз. Вказує на рівень автоматизації та самостійності системи в управлінні кіберзагрозами.

— Інтеграція з мовним інтерфейсом. Враховується зручність та результативність використання мовного інтерфейсу для звітування та взаємодії з аналітиками.

— Стійкість до потенційних атак. Оцінює, наскільки добре система витримує спроби обходу чи атаки на її компоненти.

3.4 Практична реалізація алгоритмів машинного навчання для виявлення інтрузій в мережі

Для вирішення задачі виявлення інтрузій в даному розділі розглянуто декілька методів машинного навчання. Моделювання виконувалося в середовищі Python. В якості базового коду використовувався

1 Основні етапи процесу розробки і тестування моделей

Процес розробки та тестування моделей машинного навчання для виявлення інтрузій в мережі розділений на декілька основних етапів (рис. 3.2).

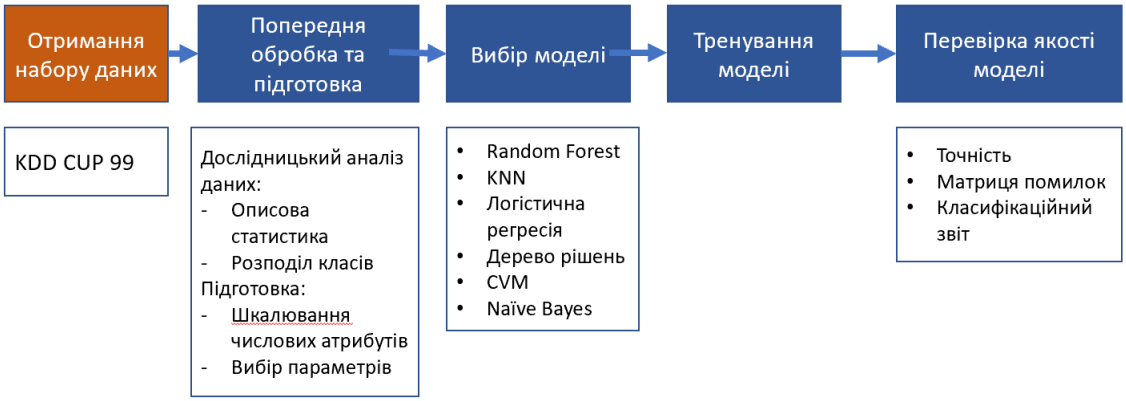

Рисунок 3.2 – Блок–схема основних етапів процесу розробки та тестування моделей.

Етап 1. Отримання набору даних

Зібрати набір даних, який містить дані аномалій в мережевому трафіку.

Етап 2. Попередня обробка даних

Ознайомитися з основними елементами набору та розподілом класів. Провести шкалювання числових атрибутів та виконати вибір параметрів для подальшого моделювання.

Етап 3. Вибір моделі

Визначити набір моделей МН для запуску та порівняння.

Етап 3. Навчання моделі

Набір даних було розділено на тренувальний та тестовий. Перший було використано для навчання моделей, другий для валідації.

Етап 5. Перевірка моделі

На даному етапі було використано валідаційний набір для визначення точності моделі. Тестовий набір використовувався для оцінки загальної ефективності та якості моделей.

Набір даних

Дослідження було проведено з використанням відкритого набору даних даних KDD CUP 99 [16]. Навчальний набір даних KDD складається приблизно з 4 900 000 одиночних векторів з’єднання, кожен з яких містить 41 атрибут, і позначається як звичайний або атакуючий, з одним конкретним типом атаки. Змодельовані атаки належать до однієї з наступних чотирьох категорій:

— Відмова в обслуговуванні (DoS) – це атака, при якій нападаючий направляє потік запитів на систему з метою зробити обчислювальний чи пам'ятовий ресурс занадто зайнятим або заповненим, щоб обробити законні запити, та в процесі завдає завади законним користувачам у доступі до машини.

— Атака–проникнення (Probe) – вивчення мережі комп'ютерів для збору інформації, яка буде використана для компрометації її засобів безпеки.

— Користувач до кореня атаки (U2R) – клас експлойтів, при якому зловмисник має доступ до звичайного облікового запису користувача на системі (здобутий шляхом перехоплення паролів, словникового атаку чи соціальної інженерії) і може використовувати деяку вразливість для отримання кореневого доступу до системи [17].

— Віддалена до локальної атаки (R2L) – виникає, коли атакуючий, який має можливість відправляти пакети на машину через мережу, але не має облікового запису на цій машині, використовує яку–небудь вразливість для отримання локального доступу як користувач цієї машини.

Файли представлено у форматі txt (рис. 3.3).

```
0,tcp,ftp_data,SF,491,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,2,2,0.00,0.00,0.00,0.00,1.00,0.00,0.00,150,25,0.17,0.03,0.17,0.00
,0.00,0.00,0.05,0.00,normal,20
0,udp,other,SF,146,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,13,1,0.00,0.00,0.00,0.00,0.08,0.15,0.00,255,1,0.00,0.60,0.88,0.00,0.
00,0.00,0.00,0.00,normal,15
0,tcp,private,S0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,123,6,1.00,1.00,0.00,0.00,0.05,0.07,0.00,255,26,0.10,0.05,0.00,0.00,
1.00,1.00,0.00,0.00,neptune,19
0,tcp,http,SF,232,8153,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,5,5,0.20,0.20,0.00,0.00,1.00,0.00,0.00,30,255,1.00,0.00,0.03,0.04,
0.03,0.01,0.00,0.01,normal,21
0,tcp,http,SF,199,420,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,30,32,0.00,0.00,0.00,0.00,1.00,0.00,0.09,255,255,1.00,0.00,0.00,0.0
0,0.00,0.00,0.00,0.00,normal,21
0,tcp,private,REJ,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,121,19,0.00,0.00,1.00,1.00,0.16,0.06,0.00,255,19,0.07,0.07,0.00,0.0
0,0.00,0.00,1.00,1.00,neptune,21
```

Рисунок 3.3 – Перші п’ять рядків тексту з тренувального набору даних

Попередній аналіз та обробка даних

На першому кроці було виконано дослідницький аналіз даних, в результаті якого було з’ясовано, що 45927 записів належать класу DoS, 67343 записів це нормальній трафік, 11656 записів, що належать до Атаки–проникнення, 995 належать до віддаленої атаки локальної мережі та 52, що свідчать про спробу доступу до кореневих облікових записів (рис. 3.4).

	attack_class	frequency_percent_train	attack_class	frequency_percent_test
DoS	45927	36.46	7458	33.08
Normal	67343	53.46	9711	43.08
Probe	11656	9.25	2421	10.74
R2L	995	0.79	2754	12.22
U2R	52	0.04	200	0.89

Рисунок 3.4 – Розподілення даних за класом атаки

Для візуалізації було використано код, що генерує стовпчасту діаграму на основі відсотка частоти в навчальних і тестових наборах даних (рис. 3.5)

Рисунок 3.5 – Код, що генерує стовпчасту діаграму для тестового та тренувального наборів

Результат надано на рис. 3.6

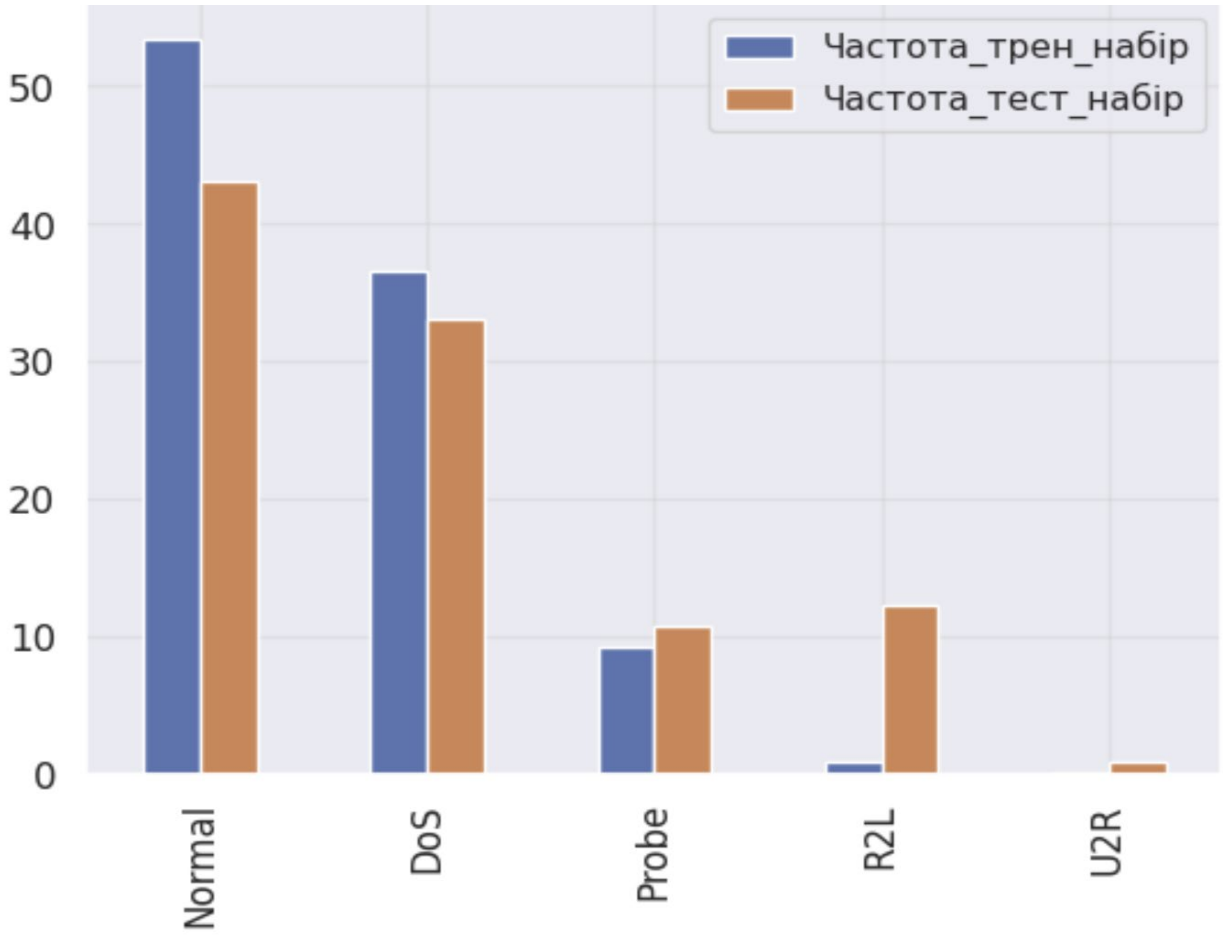

Рисунок 3.6 – Розподіл класів для тренувального та тестового наборів

Далі було виконано шкалювання цифрових атрибутів. Шкалювання числових атрибутів використовується для приведення значень числових змінних до певного діапазону або стандартного формату, що дозволяє полегшити їх порівняння або використання в алгоритмах машинного навчання та інших аналітичних задачах. Основні цілі шкалювання числових атрибутів включають:

— привести значення атрибутів до одного масштабу. Це важливо, оскільки деякі алгоритми машинного навчання чутливі до різниці в масштабах між атрибутами. Наприклад, алгоритми, які використовують відстані, такі як мають різний масштаб.

— покращення виявлення аномальних значень атрибутів, оскільки вони можуть відзначатися значеннями, які суттєво відрізняються від інших значень після шкалювання.

— підготовка даних для використання в алгоритмах. Шкалювання робить дані більш придатними для роботи з різними алгоритмами, особливо тими, які базуються на відстанях або градієнтних методах.

У даному випадку, для стандартизації числових характеристик використовувався StandardScaler, гарантуючи, що вони мають середнє значення нуль і стандартне відхилення одиниці (рис. 3.7). Код працює з двома наборами даних (dfkdd_train і dfkdd_test) окремо, масштабуючи їхні числові атрибути незалежно.

```
from sklearn.preprocessing import StandardScaler
scaler = StandardScaler()

# extract numerical attributes and scale it to have zero mean and unit variance
cols = dfkdd_train.select_dtypes(include=['float64','int64']).columns
sc_train = scaler.fit_transform(dfkdd_train.select_dtypes(include=['float64','int64']))
sc_test = scaler.fit_transform(dfkdd_test.select_dtypes(include=['float64','int64']))

# turn the result back to a dataframe
sc_traindf = pd.DataFrame(sc_train, columns = cols)
sc_testdf = pd.DataFrame(sc_test, columns = cols)
```

Рисунок 3.7 – Фрагмент коду для шкалювання атрибутів

Вибір і тренування моделі

Для навчання різних моделей класифікації та створення моделі ансамблю було використано сценарій scikit–learn. Моделі включають K–найближчих сусідів, логістичну регресію, наївний метод Байєса Гауса, дерево рішень і машину опорних векторів.

Класифікатор K–найближчих сусідів (KNN) навчався з використанням наступних налаштувань:

— Модель: `KNeighborsClassifier`

— Параметри: `n_jobs=–1` для паралельної обробки

— Підгонка моделі за допомогою навчальних даних (`X_train`,

Класифікатор логістичної регресії навчався використовуючи модель

— Параметри: `n_jobs=–1` для паралельної обробки, `random_state=0` для відтворюваності.

— Підгонка моделі за допомогою навчальних даних (`X_train`,

Гаусівський наївний класифікатор Байєса:

— Модель: `BernoulliNB`

— Підгонка моделі за допомогою навчальних даних (`X_train`,

Класифікатор дерева рішень навчався за допомогою

– Параметри: `criterion='entropy'` для розщеплення на основі ентропії, `random_state=0` для відтворюваності.

— Підгонка моделі за допомогою навчальних даних (`X_train`,

Класифікатор SVM:

— Модель: `SVC`

— Параметр: `random_state=0` для відтворюваності

— Підгонка моделі за допомогою навчальних даних (`X_train`,

Перевірка якості моделей

Перевірка якості моделей проводилася за допомогою декількох стандартних метрик:

Точність (Accuracy) визначається як відношення правильно класифікованих екземплярів до загальної кількості екземплярів. В нашому випадку, точність становить приблизно 0.93, що є високим показником.

Precision показує, яка частина позначок, виділених моделлю як певний клас, є правильними. У нашому випадку, для класу "0" точність становить 0.00, що може бути викликано відсутністю передбачуваних екземплярів цього класу.

Recall вказує на те, яка частина всіх екземплярів даного класу була виявлена моделлю. Для класу "1" recall дорівнює 1.00, що означає, що модель виявила всі екземпляри цього класу.

–Score – середнє гармонічне між precision і recall.

– кількість фактичних екземплярів кожного класу у тестовому наборі.

Результати для тренувального та тестового наборів надано у табл. 3.3,

Таблиця 3.3 – Порівняння тренувальних моделей МН для виявлення інтрузій

	Точність	Precision 0	Precision 1	Recall 0	Recall 1	f1 0	f1 1
CVM		0.99	0.99	0.99	0.99	0.99	0.99
Наївний класифікатор Байєса	0 . 9	0.98	0.97	0.97	0.98	0.97	0.97
Дерево рішень							
– найближчих сусідів							
Логістична регресія							

Таблиця 3.4 — Порівняння тестових моделей МН для виявлення інтрузій

	Точність	Precision 0	Precision 1	Recall 0	Recall 1	f1 0	f1 1
CVM	.89	0.89	0.81	0.71	0.79	0.79	0.81
Наївний класифікатор Байєса	0.83	0.86	0.82	0.74	0.91	0.79	0.86
Дерево рішень	0.82	0.81	0.82	0.75	0.78	0.78	0.84
– найближчих сусідів	0.87	0.90	0.84	0.78	0.94	0.83	0.89
Логістична регресія	0.84	0.83	0.85	0.80	0.87	0.81	0.86

Вивчаючи показники ефективності з таблиці 3.4 для тестового набору, ми можемо заглибитися в значення кожного показника для моделей машинного навчання, що використовуються в контексті виявлення вторгнень.

CVM (машина опорних векторів)

Показує збалансовану точність між обома класами з дещо кращою ідентифікацією ненавмисних дій (клас 0).

Відгук для класу 0 є нижчим, що вказує на те, що деякі ненавмисні дії можуть бути помилково класифіковані як вторгнення.

Оцінки F1 узгоджуються для обох класів, що відображає збалансований компроміс між точністю та відгуком.

Наївний класифікатор Байєса

Має хорошу точність, особливо для неінтрузивних дій.

Високий відгук для вторгнень (клас 1) означає, що він здатний ідентифікувати більшість реальних вторгнень, хоча і ціною більшої кількості хибних спрацьовувань, на що вказує нижча точність.

Показник F1 для класу 1 помітно вищий, що свідчить про кращу ефективність у виявленні реальних вторгнень.

Дерево рішень

Показує дуже близьку точність для обох класів, але з помітним падінням порівняно з навчальною вибіркою, що вказує на можливі проблеми з надмірним налаштуванням.

Recall трохи кращий для класу 1, що свідчить про те, що він може добре виявляти вторгнення, але потребує вдосконалення у визначенні ненавмисної поведінки.

Оцінка F1 є вищою для вторгнень, що означає ефективне виявлення, але також вказує на потенційне перенастроювання, оскільки вона не відповідає ідеальній оцінці з навчального набору.

–Найближчі сусіди

Ця модель має найвищу точність для ненавмисних дій, що свідчить про низький рівень хибних спрацьовувань для класу 0.

Має найвищий відгук для класу 1, що свідчить про те, що вона рідко пропускає реальні вторгнення.

Оцінка F1 для класу 1 є найвищою серед усіх моделей, що свідчить про надійність виявлення вторгнень.

Логістична регресія

Показує хорошу точність для обох класів, з дещо кращими показниками для класу 1.

Відгук також вищий для вторгнень, що свідчить про ефективність у виявленні істинних спрацьовувань.

Оцінки F1 є високими і збалансованими, що свідчить про хорошу продуктивність в цілому, особливо для виявлення вторгнень.

Показники в Таблиці 3.3 дають повне уявлення про здатність моделей розрізняти нормальну поведінку і вторгнення. Висока точність у класі 1 має вирішальне значення для систем виявлення вторгнень, щоб уникнути помилкового позначення нормальної діяльності як вторгнення, що може призвести до непотрібних дій або сповіщень. І навпаки, високий відгук для класу 1 важливий для того, щоб гарантувати, що справжні вторгнення будуть виявлені і не пропущені, що може призвести до порушень безпеки.

Показник F1 є гармонійним середнім значенням точності та відтворення, що дає єдиний показник, який збалансовує обидва аспекти. Високий показник F1 для класу 1 особливо важливий в контексті безпеки, оскільки він свідчить про те, що модель є одночасно точною і всеосяжною у виявленні вторгнень.

Варто зазначити, що хоча модель K–найближчих сусідів показує найвищий показник F1 для вторгнень у тестовому наборі, дуже важливо враховувати інші фактори, такі як обчислювальна ефективність моделі, вартість помилкових спрацьовувань у порівнянні з помилковими неспрацьовуваннями, а також специфіку мережевого середовища, в якому буде розгорнута модель. Ці результати повинні слугувати орієнтиром, але не диктувати вибір моделі для практичного використання в системах виявлення вторгнень.

ОХОРОНА ПРАЦІ ТА БЕЗПЕКА В НАДЗВИЧАЙНИХ СИТУАЦІЯХ

Охорона праці

Сучасний розвиток технічного та технологічного стану виробництва передбачає постійну автоматизацію та оптимізацію виробничих процесів. Сьогодні, напевно, важко уявити компанію, господарська діяльність в якій здійснювалась би без використання комп'ютерної техніки. Через масовий характер робіт, що виконуються працівниками за допомогою комп'ютера, законодавством України чітко врегульовано норми та вимоги до використання комп'ютерної техніки на підприємстві, безпосередньо й охорона праці при роботі з комп'ютером [17].

Згідно з нормативними актами про охорону праці (НПАОП 0.00-7.15- 18) є такі вимоги безпеки до робочих місць працівників з електронними пристроями:

— Площа, відведена на одне робоче місце має становити не менше 6 кв.м., а об'єм – не менше 20 куб.м [18].

— Конструкція робочого місця повинна забезпечувати підтримання оптимальної робочої пози, тобто такої, яка дозволяє працівникові виконувати роботу з мінімальним напруженням тіла, і яка дозволяє уникнути перевтоми в ході і після закінчення робочого процесу.

— Для забезпечення безпеки та захисту здоров'я працівників усе випромінювання від екранних пристроїв має бути зведене до гранично допустимого рівня (вплив на людину факторів довкілля - шуму, вібрації, забруднювачів, температури тощо, який не спричиняє соматичних або психічних розладів, а також змін стану здоров'я, працездатності, поведінки,що виходять за межі пристосувальних реакцій) з погляду безпеки та охорони здоров'я працівників [28].

— Організація робочого місця працівника з екранними пристроями має забезпечувати відповідність усіх елементів робочого місця та їх

розташування ергономічним, антропологічним, психофізіологічним вимогам, а також характеру виконуваних робіт [19].

— Освітлення робочого місця працівника з екранними пристроями має створювати відповідний контраст між екраном і навколишнім середовищем (з урахуванням виду роботи) та відповідати вимогам ДСанПІН

— Мікроклімат приміщень з робочими місцями працівників з екранними пристроями має підтримуватись на постійному рівні та відповідати вимогам Санітарних норм мікроклімату виробничих приміщень ДСН 3.3.6.042-99, затверджених постановою Головного державного санітарного лікаря України від 01 грудня 1999 року No 42 [20].

Вимоги щодо розміщення ІТС

Приміщення, в яких планується установка та подальша робота з комп'ютером, повинні відповідати проектній документації будинку, погодженій з уповноваженими державними органами. Крім того, роботодавець повинен враховувати санітарні нормативи освітлення, вимоги до параметрів мікроклімату (температура, відносна вологість), ступеня і сили вібрації, звукового шуму і вогнестійкості приміщення, а також характеристики електромагнітного, ультрафіолетового та інфрачервоного полів. Робочі місця, обладнані персональними комп'ютерами, заборонено облаштовувати у підвальних або цокольних приміщеннях будівель [18]. При обладнанні приміщень забороняється використання полімерних матеріалів, що виділяють шкідливі хімічні речовини.

Природне і штучне освітлення

Згідно документу ДБН В.2.5-28:2018 "Природне і штучне освітлення" приміщення з постійним перебуванням людей повинні мати природне освітлення. Природне освітлення поділяється на бокове, верхнє і комбіноване. Що до штучного освітлення воно поділяється на робоче, аварійне, охоронне і чергове [20].

Для загального штучного освітлення доцільно використовувати розрядні та світлодіодні джерела світла, які за однакової потужності з тепловими джерелами мають більшу світлову віддачу та більший термін експлуатації.

Види інструктажів з охорони праці

Працівники, під час прийняття на роботу та періодично, повинні проходити на підприємстві інструктажі з питань охорони праці, надання першої медичної допомоги потерпілим від нещасних випадків, а також з правил поведінки та дій при виникненні аварійних ситуацій, пожеж і стихійних лих.

За характером і часом проведення інструктажі з питань охорони праці (далі - інструктажі) поділяються на вступний, первинний, повторний, позаплановий та цільовий.)

Вступний інструктаж Проводиться:

— з усіма працівниками, які приймаються на постійну або тимчасову роботу, незалежно від їх освіти, стажу роботи та посади;

— з працівниками інших організацій, які прибули до організації і беруть безпосередню участь у робочому процесі або виконують інші роботи для підприємства;

Вступний інструктаж проводиться спеціалістом служби охорони праці або іншим фахівцем відповідно до наказу (розпорядження) по організації, який в установленому типовим положенням порядку пройшов навчання і перевірку знань з питань охорони праці [20]. Первинний інструктаж.

Первинний інструктаж проводиться до початку роботи безпосередньо на робочому місці з працівником:

— новоприйнятим (постійно чи тимчасово) до організації або до фізичної особи, яка використовує найману працю;

— який переводиться з одного структурного підрозділу організації до іншого.

Повторний інструктаж

Повторний інструктаж на робочому місці індивідуально з окремим працівником або групою працівників, які виконують однотипні роботи, за обсягом і змістом переліку питань первинного інструктажу.

Повторний інструктаж проводиться в терміни, визначені нормативно-правовими актами з охорони праці, які діють у галузі, або роботодавцем (фізичною особою, яка використовує найману працю) з урахуванням конкретних умов праці, але не рідше:

— на роботах з підвищеною небезпекою - 1 раз на 3 місяці;

— для решти робіт - 1 раз на 6 місяців. Позаплановий інструктаж.

— Позаплановий інструктаж проводиться з працівниками на робочому місці або в кабінеті охорони праці:

— при введенні в дію нових або переглянутих нормативно-правових актів з охорони праці, а також при внесенні змін та доповнень до них;

— при порушеннях працівниками вимог нормативно-правових актів з охорони;

— праці, що призвели до травм, аварій, пожеж тощо;

— при перерві в роботі виконавця робіт більш ніж на 30 календарних днів для робіт з підвищеною небезпекою, а для решти робіт - понад 60 днів.

Цільовий інструктаж.

Цільовий інструктаж проводиться з працівниками:

— при ліквідації аварії або стихійного лиха;

— при проведенні робіт, на які відповідно до законодавства оформлюються наряд-допуск, наказ або розпорядження.

Цільовий інструктаж проводиться індивідуально з окремим працівником або з групою працівників. Обсяг і зміст цільового інструктажу визначаються залежно від виду робіт, що виконуватимуться.

Безпека в надзвичайних ситуаціях

Здоров'я людини ґрунтується на основі генетичних факторів, способу життя та екологічних умов. Однак певною мірою воно залежить також від свідомого ставлення людини до себе та оточуючого середовища. Здоров'я людини — стан повного соціально-біологічного комфорту коли функція всіх органів і систем організму виважені з природним і соціальним середовищем, відсутні будь-які хвилювання, хворобливі стани та фізичні дефекти. Критерій здоров'я визначається комплексом показників. Однак за найзагальнішими рисами здоров'я індивідуума можна визначити як природний стан організму, що характеризується повною зрівноваженістю будь-яких виражених хворобливих змін. Слід пам'ятати, що здоров'я залежить від багатьох факторів які об'єднуються в одне інтегральне поняття —здоровий спосіб життя. Його метою є навчити людину розумно ставитися до свого здоров'я, фізичної та психічної культури, загартовувати свій організм, вміло організовувати працю і відпочинок.

До основних складових здорового способу життя належать декілька основних чинників.

Спосіб життя має велике значення для здоров'я людини і складається з чотирьох категорій:

— Економічної (рівень життя).
— Соціологічної (якість життя).
— Соціально-психологічної.
— Соціально-економічної.

Отже, до способу життя людини належать: активна участь людини в процесі формування умов життя, її адекватна реакція на зміну умов навколишнього середовища, а також праця, побут, задоволення матеріальних і духовних потреб у суспільному житті, норми і правила поведінки.

Слід пам'ятати, що людина — суб'єкт і одночасно — головний результат своєї діяльності. Культура з цієї точки зору — це самосвідоме ставлення до самого себе. Однак люди дуже часто нехтують своїм здоров'ям, ведуть неправильний спосіб життя, не дотримуються режиму переїдають, курять. Тому для здоров'я потрібні знання, які увійшли б у повсякденну звичку людини.

Не завжди в житті людини здоров'я займає перше місце порівняно з речами та іншими матеріальними благами. У результаті це призводить до шкоди не лише своєму здоров'ю, а й здоров'ю майбутніх поколінь. Отже, здоров'я повинно займати перше місце в ієрархії потреб людини.

На превеликий жаль, ціну здоров'я більшість людей усвідомлює лише тоді, коли воно значно похитнулось. Лише тоді виникає прагнення вилікувати захворювання, стати здоровим.

Нерозумне і довге випробовування стійкості свого організму нездоровим способом життя (алкоголь, нікотин). Тільки через певний час спрацьовують зворотні зв'язки у людини, коли вона полишає шкідливі звички, проте, часто запізно.

Джерелом навичок з цього питання є, передусім, приклад батьків, допомагає також і санітарна освіта. Важливим фактором, що визначає реакцію людини на екстремальну ситуацію, є її психофізичні якості та загальний стан. Вони проявляються через чутливість людини до виявлення сигналів небезпеки перед реакцією на них. Показники, які зумовлюють можливості людини виявити небезпечну ситуацію та адекватно відреагувати на неї, залежать від її індивідуальних особливостей, зокрема від її нервової системи. На поведінку людини у небезпечній ситуації впливає й її психічний та фізичний стан.

Відомо, що 80 % більшості хвороб мають психосоматичний характер, тобто значною мірою залежить від стану душі людини, який визначає її безпечну поведінку.

Сучасна людина зустрічається з багатьма факторами ризику, що негативно впливають на стан й нервової та серцево-судинної систем, знижує

опірність організму. При цьому виникає стресова реакція організму. Так, наприклад, психічна травма, отримана внаслідок конфлікту, виводить людину з нормального психічного стану, що може призвести до суттєвих змін у виконанні професійних функцій і загального функціонального стану. У перекладі «стрес» означає «напруження», тобто відповідь організму на поставлену перед ним проблему.

Велике значення для розвитку стресового стану має поведінка в екстремальних умовах (аварія, кримінальна ситуація, стихійне лихо). Неправильна поведінка у таких ситуаціях найчастіше є причиною шкідливих наслідків стресу. Вона зумовлює результат стресу більше, ніж фактори зовнішнього середовища. У цих випадках стрес може виявитись у вигляді паніки, суєти, істерики.

Це захисна реакція організму на зовнішні надзвичайні подразники і ситуації, тривалі негативні емоції. Він супроводжується підвищеним серцебиттям, виснаженням і зривом адаптаційних і імунних систем організму та іншими змінами. До певної межі стрес сприяє вирішенню людиною певних завищених завдань і навантажень. Однак, у разі перевищення цієї межі в організмі людини виникають порушення механізмів саморегуляції, відбувається погіршення трудової діяльності і стануться зриви, які призводять до виникнення небезпечних ситуацій. При стресових ситуаціях різко підвищується вміст адреналіну у крові, посилюється робота серця, звужуються кровоносні судини, підвищується температура тіла і рівень глюкози у крові. У результаті в організмі виникають фізіологічні порушення, розлади нервової, серцево-судинної систем та ін. До цих розладів належать нервовість, роздратованість, тривога, агресивність, втома, загострення хворобливих станів.

Тривала стресова ситуація призводить до багатьох психосоматичних захворювань: психозів, неврозів, захворювань мозку, серцево-судинних захворювань, інфаркту, гіпертонічної хвороби, шлунково-кишкових захворювань, зниження імунітету, онкологічних захворювань.

4.3 Висновок до четвертого розділу

Таким чином, у результаті аналізу вимог щодо охорони праці користувачів комп’ютерів, визначено особливості організації робочих місць, вимог з електробезпеки, природного та штучного освітлення для ефективної і безпечної роботи.

Також розглянуто питання здорового способу життя та його вплив на професійну діяльність, структури системи БЖД, елементів теорії, що відповідають моделі безпеки життєдіяльності.

ВИСНОВОК

Робота ефективно відображає прогрес у сфері кібербезпеки через інтеграцію штучного інтелекту, зосереджуючись на аналізі різних методів та систем, їх ефективності, викликів та можливостей. Вона демонструє, як впровадження ШІ може значно покращити здатність систем кібербезпеки до виявлення, аналізу та реагування на кіберзагрози, надаючи глибокий аналіз та порівняльну оцінку різних підходів.

У роботі розглядаються класичні методи кібербезпеки, які зазвичай зосереджені на локальних контрольних мірах та реактивному реагуванні на виявлені загрози. Хоча ці методи ефективні для виявлення та реагування на унікальні атаки, вони часто обмежені у своїй здатності адаптуватися до нових та розвиваючихся загроз. Також вони мають недоліки, такі як низька швидкість реакції та обмежена можливість адаптації до нових видів загроз.

Напівуніфіковані системи з ШІ представлені як поєднання традиційних методів з алгоритмами машинного навчання. Вони дозволяють автоматизувати виявлення аномалій та зменшують залежність від ручного аналізу. Проте ці системи також стикаються з викликами, такими як висока складність впровадження та потреба у великих обсягах даних для ефективного навчання.

Уніфіковані системи з ШІ, які повністю інтегрують машинне навчання для аналізу даних, показують великі обіцянки у забезпеченні швидкого реагування на загрози та комплексного підходу до кібербезпеки. Ці системи демонструють здатність швидко адаптуватися до нових кіберзагроз, пропонуючи комплексний та автоматизований підхід до виявлення та реагування.

Ефективність систем кібербезпеки оцінюється за допомогою різних метрик, таких як точність виявлення загроз, швидкість реагування, відсоток помилкових спрацювань, та інші. Результати показують, що системи, що використовують ШІ, можуть значно покращити ці показники, однак вони

також потребують складних алгоритмів та великої кількості даних для навчання.

Загалом, робота підкреслює значення адаптації до змінюваних кіберзагроз та важливість балансу між інноваційними технологічними рішеннями та глибоким розумінням викликів у сфері кібербезпеки. Вона відкриває нові перспективи для розвитку більш ефективних, адаптивних та надійних систем кібербезпеки, що здатні відповідати на сучасні та майбутні виклики у цій галузі.

ПЕРЕЛІК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Abbas, S., Naser, W., & Kadhim, A. (2023). Subject review: Intrusion Detection System (IDS) and Intrusion Prevention System (IPS). Global Journal of Engineering and Technology Advances. DOI:https://doi.org/10.30574/gjeta.2023.14.2.0031
2. Abdiyeva-Aliyeva, G., Hematyar, M., & Bakan, S. (2021). Development of System for Detection and Prevention of Cyber Attacks Using Artificial Intelligence Methods. 2021 2nd Global Conference for Advancement in Technology (GCAT), 1-5. DOI:https://doi.org/10.1109/GCAT52182.2021.9587584.
3. Acosta, J., Basak, A., Kiekintveld, C., Leslie, N., & Kamhoua, C. (2020). Cybersecurity Deception Experimentation System. 2020 IEEE Secure Development (SecDev), 34-40. DOI:https://doi.org/10.1109/SecDev45635.2020.00022.
4. Arun E Thomas: Security Operations Center – SIEM Use Cases and Cyber Threat Intelligence
5. Azodi, A., Jaeger, D., Cheng, F., & Meinel, C. (2013). Pushing the Limits in Event Normalisation to Improve Attack Detection in IDS/SIEM Systems. 2013 International Conference on Advanced Cloud and Big Data, 69-76.DOI:https://doi.org/10.1109/CBD.2013.27.
6. Geluvaraj, B., Satwik, P., & Kumar, T. (2018). The Future of Cybersecurity: Major Role of Artificial Intelligence, Machine Learning, and Deep Learning in Cyberspace. International *Conference on Computer Networks and Communication Technologies*.

   DOI:https://doi.org/10.1007/978-981-10-8681-6_67.

7. Hajji, S., Moukafih, N., & Orhanou, G. (2019). Analysis of Neural Network Training and Cost Functions Impact on the Accuracy of IDS and SIEM Systems. , 433-451.DOI https://doi.org/10.1007/978-3-030-16458-4_25.

8. Idris, H. (2022). Machine Learning Approach for Cybersecurity Implementation. *2022 International Conference on Business Analytics for Technology and Security (ICBATS)*, 1-4. DOI:https://doi.org/10.1109/ICBATS54253.2022.9759091.
9. Jeff Bollinger, Brandon Enright, and Matthew Valites :Crafting the InfoSec Playbook: Security Monitoring and Incident Response Master Plan
10. Joseph Muniz:Modern Security Operations Center
11. Kenaza, T., Machou, A., & Dekkiche, A. (2018). Implementing a Semantic Approach for Events Correlation in SIEM Systems. DOI:https://doi.org/10.1007/978-3-319-89743-1_55.
12. Data Loss Prevention (DLP): High-impact Strategies - What You Need to Know Kevin Roebuck Definitions, Adoptions, Impact, Benefits, Maturity, Vendors Emereo Pty Limited, 2011
13. Paracha, M., Sheeraz, M., Chai, Y., Ahmad, S., Khan, Z., Hussain, S., , A., & Durad, M. (2022). Implementation of Two Layered DLP Strategies. *2022 International Conference on Cyber Warfare and Security (ICCWS)*, 8-13. DOI:https://doi.org/10.1109/ICCWS56285.2022.9998436.
14. Scarfone, K., & Mell, P. (2010). Intrusion Detection and Prevention Systems. , 177-192. DOI:https://doi.org/10.1007/978-3-642-04117-4_9.
15. Soliman, K., Sobh, M., & Bahaa-Eldin, A. (2021). Survey of Machine Learning HIDS Techniques. *2021 16th International Conference on Computer Engineering and Systems (ICCES)*. DOI:https://doi.org/10.1109/ICCES54031.2021.9686138.
16. Tavallaee, Mahbod; Bagheri, Ebrahim; Lu, Wei; Ghorbani, Ali-A. A Detailed Analysis of the KDD CUP 99 Data Set .
17. Наказ Міністерства соціальної політики України «Про затвердження Вимог щодо безпеки та захисту здоров'я працівників під час роботи з екранними пристроями [Електронний ресурс] – Режим доступу до ресурсу: https://zakon.rada.gov.ua/laws/show/z0508-18.

18. Закон України «Про охорону праці» [Електронний ресурс] – Режим доступу до ресурсу: https://zakon.rada.gov.ua/laws/show/2694-12.
19. Наказ Міністерства внутрішніх справ України «Про затвердження Правил пожежної безпеки в Україні» [Електронний ресурс] – Режим доступу до ресурсу: https://zakon.rada.gov.ua/laws/show/z0252-15.
20. Державний комітет ядерного регулювання України. Проект від 01.03.2008 р. Консультації щодо підвищення безпеки джерел іонізуючого випромінювання в Україні. Київ, 2008.
21. Білявський Г.О, Бутченко Л.І., Навроцький В.М. Основи екології: Теорія і практикум: Навч. Посібник. Київ, 2002. 352

ДОДАТКИ

Додаток А

УДК 004.056
Задорожний С.Ю., Скарга-Бандурова І.С., д.т.н., проф.
Тернопільський національний технічний університет імені Івана Пулюя, Україна

МОЖЛИВОСТІ ЗАСТОСУВАННЯ ШТУЧНОГО ІНТЕЛЕКТУ В ОПЕРАЦІЙНОМУ ЦЕНТРІ БЕЗПЕКИ

S. Yu. Zadorozhnyi, I.S. Skarga-Bandurova, DSc, Prof

HARNESSING ARTIFICIAL INTELLIGENCE FOR SECURITY OPERATIONS CENTRES

Центр Операційної Безпеки (Security Operations Centre або SOC) є центральним пунктом будь-якої організації, який виконує моніторинг, аналіз та реагування на інциденти безпеки. Основна місія SOC полягає в захисті підприємства від порушень та атак, забезпеченні безпеки активів (даних, додатків, інфраструктури тощо) та забезпеченні нормального функціонування. За визначенням [1], SOC є комбінацією людей, процесів та технологій, що забезпечують захист інформаційних систем організації через про активний дизайн та конфігурацію, постійний моніторинг, виявлення непередбачених дій або небажаного стану та мінімізацію шкоди від небажаних ефектів. Для виявлення та протидії кіберзагрозам, SOC повинен збирати дані по всій організації і проводити аналіз в реальному або наближеному до реального часу, використовуючи великі обсяги даних з журналів (наприклад, мережевих та веб-брандмауерів, датчиків мережі, кінцевих точок), систем виявлення та запобігання вторгнень, систем управління ідентифікацією та доступом, та безлічі інших джерел. Однак, незважаючи на всі захисні заходи, організації постійно піддаються кібератакам і порушенням безпеки, що може негативно впливати на їх діяльність. Вузьким місцем в традиційних підходах є те, що процес реагування на інциденти обробляється вручну аналітиками, які отримують сповіщення та аналізують їх, щоб вирішити, яку реакцію вжити. Крім того, традиційні методи породжують велику кількість сповіщень, які перевантажують аналітиків SOC, змушуючи їх визначати пріоритети та приймати відповідні заходи щодо ліквідації. Використання методів штучного інтелекту (AI) та машинного навчання (ML) в SOC може вирішити ці завдання, забезпечуючи більшу точність, швидкість та адаптивність у боротьбі з кіберзагрозами.

Основна мета цього дослідження - оцінити, як методи AI та ML можуть вдосконалити системи запобігання інтрузій (IPS), системи управління інформаційною безпекою (SIEM), та системи запобігання витоку даних (DLP) у контексті SOC. Це дослідження фокусується на аналізі проблем та викликів, з якими стикаються ці системи, та розглядає потенціал AI/ML для їх подолання. У роботі розглядаються наступні дослідницькі запитання:

1. Аналіз проблем та викликів у системах IPS/IDS, SIEM, DLP. Визначення конкретних труднощів, з якими стикаються ці системи, та з'ясування їх взаємозв'язків та впливу на загальну безпеку.

2. Дослідження шляхів впровадження AI/ML для розв'язання визначених проблем, оцінка їх практичних переваг та аналіз викликів для їх впровадження.

3. Впровадження екосистеми вибраних систем. Дослідження можливостей інтеграції систем IPS/IDS, SIEM, DLP у єдину екосистему на базі ШІ та аналіз ефективності такого підходу для підвищення функціональності SOC.

Література

1. Security, R.S.I. (2021). NIST Security Operations Center Best Practices. [online] RSI Security. Available at: https://blog.rsisecurity.com/nist-security-operations-center-best-practices/.

Додаток Б

Лістинг 3.4 — Створення і навчання штучного інтелекту

```
# import relevant modules
%matplotlib inline
import matplotlib
import matplotlib.pyplot as plt
import pandas as pd
import numpy as np
import seaborn as sns
import sklearn
import imblearn

# Ignore warnings
import warnings
warnings.filterwarnings('ignore')

# Settings
pd.set_option('display.max_columns', None)
np.set_printoptions(threshold=np.nan)
np.set_printoptions(precision=3)
sns.set(style="darkgrid")
plt.rcParams['axes.labelsize'] = 14
plt.rcParams['xtick.labelsize'] = 12
plt.rcParams['ytick.labelsize'] = 12

print("pandas : {0}".format(pd.__version__))
print("numpy : {0}".format(np.__version__))
print("matplotlib : {0}".format(matplotlib.__version__))
print("seaborn : {0}".format(sns.__version__))
print("sklearn : {0}".format(sklearn.__version__))
print("imblearn : {0}".format(imblearn.__version__))

# Dataset field names
datacols = ["duration","protocol_type","service","flag","src_bytes",
"dst_bytes","land","wrong_fragment","urgent","hot","num_failed_logins","logged
_in","num_compromised","root_shell","su_attempted","num_root",
"num_file_creations","num_shells","num_access_files","num_outbound_cmds","is
_host_login","is_guest_login","count","srv_count","serror_rate",
"srv_serror_rate","rerror_rate","srv_rerror_rate","same_srv_rate","diff_srv_rate","s
rv_diff_host_rate","dst_host_count","dst_host_srv_count",
"dst_host_same_srv_rate","dst_host_diff_srv_rate","dst_host_same_src_port_rate"
,"dst_host_srv_diff_host_rate","dst_host_serror_rate","dst_host_srv_serror_rate","
dst_host_rerror_rate","dst_host_srv_rerror_rate","attack", "last_flag"]
```

```
# Load NSL_KDD train dataset
dfkdd_train = pd.read_table("~/NID/NSL_KDD_dataset/KDDTrain.txt", sep=",",
names=datacols) # change path to where the dataset is located.
dfkdd_train = dfkdd_train.iloc[:,:-1] # removes an unwanted extra field

# Load NSL_KDD test dataset
dfkdd_test = pd.read_table("~/NID/NSL_KDD_dataset/KDDTest.txt", sep=",",
names=datacols)
dfkdd_test = dfkdd_test.iloc[:,:-1]

# View train data
dfkdd_train.head(3)

# train set dimension
print('Train set dimension: {} rows, {} columns'.format(dfkdd_train.shape[0],
dfkdd_train.shape[1]))
Train set dimension: 125973 rows, 42 columns
```

Test dataset

```
# View test data
dfkdd_test.head(3)

# test set dimension
print('Test set dimension: {} rows, {} columns'.format(dfkdd_test.shape[0],
dfkdd_test.shape[1]))
Test set dimension: 22544 rows, 42 columns

mapping = {'ipsweep': 'Probe','satan': 'Probe','nmap': 'Probe','portsweep':
'Probe','saint': 'Probe','mscan': 'Probe',
      'teardrop': 'DoS','pod': 'DoS','land': 'DoS','back': 'DoS','neptune': 'DoS','smurf':
'DoS','mailbomb': 'DoS',
      'udpstorm': 'DoS','apache2': 'DoS','processtable': 'DoS',
      'perl': 'U2R','loadmodule': 'U2R','rootkit': 'U2R','buffer_overflow':
'U2R','xterm': 'U2R','ps': 'U2R',
      'sqlattack': 'U2R','httptunnel': 'U2R',
      'ftp_write': 'R2L','phf': 'R2L','guess_passwd': 'R2L','warezmaster':
'R2L','warezclient': 'R2L','imap': 'R2L',
      'spy': 'R2L','multihop': 'R2L','named': 'R2L','snmpguess': 'R2L','worm':
'R2L','snmpgetattack': 'R2L',
      'xsnoop': 'R2L','xlock': 'R2L','sendmail': 'R2L',
      'normal': 'Normal'
      }
# Apply attack class mappings to the dataset
dfkdd_train['attack_class'] = dfkdd_train['attack'].apply(lambda v: mapping[v])
dfkdd_test['attack_class'] = dfkdd_test['attack'].apply(lambda v: mapping[v])
# Drop attack field from both train and test data
```

```
dfkdd_train.drop(['attack'], axis=1, inplace=True)
dfkdd_test.drop(['attack'], axis=1, inplace=True)
# View top 3 train data
dfkdd_train.head(3)
dfkdd_train['num_outbound_cmds'].value_counts()
dfkdd_test['num_outbound_cmds'].value_counts()
0    125973
Name: num_outbound_cmds, dtype: int64
0    22544
Name: num_outbound_cmds, dtype: int64
# 'num_outbound_cmds' field has all 0 values. Hence, it will be removed from both
train and test dataset since it is a redundant field.
dfkdd_train.drop(['num_outbound_cmds'], axis=1, inplace=True)
dfkdd_test.drop(['num_outbound_cmds'], axis=1, inplace=True)
# Attack Class Distribution
attack_class_freq_train = dfkdd_train[['attack_class']].apply(lambda x:
x.value_counts())
attack_class_freq_test = dfkdd_test[['attack_class']].apply(lambda x:
x.value_counts())
attack_class_freq_train['frequency_percent_train'] = round((100 *
attack_class_freq_train / attack_class_freq_train.sum()),2)
attack_class_freq_test['frequency_percent_test'] = round((100 *
attack_class_freq_test / attack_class_freq_test.sum()),2)

attack_class_dist = pd.concat([attack_class_freq_train,attack_class_freq_test],
axis=1)
attack_class_dist
# Attack class bar plot
plot = attack_class_dist[['frequency_percent_train',
'frequency_percent_test']].plot(kind="bar");
plot.set_title("Attack Class Distribution", fontsize=20);
plot.grid(color='lightgray', alpha=0.5);


dfkdd_train.head()
from sklearn.preprocessing import StandardScaler
scaler = StandardScaler()

# extract numerical attributes and scale it to have zero mean and unit variance
cols = dfkdd_train.select_dtypes(include=['float64','int64']).columns
sc_train =
scaler.fit_transform(dfkdd_train.select_dtypes(include=['float64','int64']))
sc_test = scaler.fit_transform(dfkdd_test.select_dtypes(include=['float64','int64']))

# turn the result back to a dataframe
```

```
sc_traindf = pd.DataFrame(sc_train, columns = cols)
sc_testdf = pd.DataFrame(sc_test, columns = cols)

from sklearn.preprocessing import LabelEncoder
encoder = LabelEncoder()

# extract categorical attributes from both training and test sets
cattrain = dfkdd_train.select_dtypes(include=['object']).copy()
cattest = dfkdd_test.select_dtypes(include=['object']).copy()

# encode the categorical attributes
traincat = cattrain.apply(encoder.fit_transform)
testcat = cattest.apply(encoder.fit_transform)

# separate target column from encoded data
enctrain = traincat.drop(['attack_class'], axis=1)
enctest = testcat.drop(['attack_class'], axis=1)

cat_Ytrain = traincat[['attack_class']].copy()
cat_Ytest = testcat[['attack_class']].copy()
```

Data Sampling

```
from imblearn.over_sampling import RandomOverSampler
from collections import Counter

# define columns and extract encoded train set for sampling
sc_traindf = dfkdd_train.select_dtypes(include=['float64','int64'])
refclasscol = pd.concat([sc_traindf, enctrain], axis=1).columns
refclass = np.concatenate((sc_train, enctrain.values), axis=1)
X = refclass

# reshape target column to 1D array shape
c, r = cat_Ytest.values.shape
y_test = cat_Ytest.values.reshape(c,)

c, r = cat_Ytrain.values.shape
y = cat_Ytrain.values.reshape(c,)

# apply the random over-sampling
ros = RandomOverSampler(random_state=42)
X_res, y_res = ros.fit_sample(X, y)
print('Original dataset shape {}'.format(Counter(y)))
print('Resampled dataset shape {}'.format(Counter(y_res)))
Original dataset shape Counter({1: 67343, 0: 45927, 2: 11656, 3: 995, 4: 52})
Resampled dataset shape Counter({0: 67343, 1: 67343, 2: 67343, 3: 67343, 4:
67343})
```

```
from sklearn.ensemble import RandomForestClassifier
rfc = RandomForestClassifier();

# fit random forest classifier on the training set
rfc.fit(X_res, y_res);
# extract important features
score = np.round(rfc.feature_importances_,3)
importances = pd.DataFrame({'feature':refclasscol,'importance':score})
importances =
importances.sort_values('importance',ascending=False).set_index('feature')
# plot importances
plt.rcParams['figure.figsize'] = (11, 4)
importances.plot.bar();


from sklearn.feature_selection import RFE
import itertools
rfc = RandomForestClassifier()

# create the RFE model and select 10 attributes
rfe = RFE(rfc, n_features_to_select=10)
rfe = rfe.fit(X_res, y_res)

# summarize the selection of the attributes
feature_map = [(i, v) for i, v in itertools.zip_longest(rfe.get_support(), refclasscol)]
selected_features = [v for i, v in feature_map if i==True]
selected_features
['src_bytes',
 'dst_bytes',
 'logged_in',
 'count',
 'srv_count',
 'dst_host_srv_count',
 'dst_host_diff_srv_rate',
 'dst_host_same_src_port_rate',
 'dst_host_serror_rate',
 'service']

# define columns to new dataframe
newcol = list(refclasscol)
newcol.append('attack_class')

# add a dimension to target
new_y_res = y_res[:, np.newaxis]
```

```
# create a dataframe from sampled data
res_arr = np.concatenate((X_res, new_y_res), axis=1)
res_df = pd.DataFrame(res_arr, columns = newcol)

# create test dataframe
reftest = pd.concat([sc_testdf, testcat], axis=1)
reftest['attack_class'] = reftest['attack_class'].astype(np.float64)
reftest['protocol_type'] = reftest['protocol_type'].astype(np.float64)
reftest['flag'] = reftest['flag'].astype(np.float64)
reftest['service'] = reftest['service'].astype(np.float64)

res_df.shape
reftest.shape
(336715, 41)
(22544, 41)
from collections import defaultdict
classdict = defaultdict(list)

# create two-target classes (normal class and an attack class)
attacklist = [('DoS', 0.0), ('Probe', 2.0), ('R2L', 3.0), ('U2R', 4.0)]
normalclass = [('Normal', 1.0)]

def create_classdict():
   '''This function subdivides train and test dataset into two-class attack labels'''
   for j, k in normalclass:
      for i, v in attacklist:
         restrain_set = res_df.loc[(res_df['attack_class'] == k) |
(res_df['attack_class'] == v)]
         classdict[j +'_' + i].append(restrain_set)
         # test labels
         reftest_set = reftest.loc[(reftest['attack_class'] == k) | (reftest['attack_class']
== v)]
         classdict[j +'_' + i].append(reftest_set)

create_classdict()
for k, v in classdict.items():
   k
'Normal_DoS'
'Normal_R2L'
'Normal_Probe'
'Normal_U2R'
pretrain = classdict['Normal_DoS'][0]
pretest = classdict['Normal_DoS'][1]
grpclass = 'Normal_DoS'
```

Finalize data preprocessing for training

```
from sklearn.preprocessing import OneHotEncoder
enc = OneHotEncoder()

Xresdf = pretrain
newtest = pretest

Xresdfnew = Xresdf[selected_features]
Xresdfnum = Xresdfnew.drop(['service'], axis=1)
Xresdfcat = Xresdfnew[['service']].copy()

Xtest_features = newtest[selected_features]
Xtestdfnum = Xtest_features.drop(['service'], axis=1)
Xtestcat = Xtest_features[['service']].copy()


# Fit train data
enc.fit(Xresdfcat)

# Transform train data
X_train_1hotenc = enc.transform(Xresdfcat).toarray()

# Transform test data
X_test_1hotenc = enc.transform(Xtestcat).toarray()

X_train = np.concatenate((Xresdfnum.values, X_train_1hotenc), axis=1)
X_test = np.concatenate((Xtestdfnum.values, X_test_1hotenc), axis=1)

y_train = Xresdf[['attack_class']].copy()
c, r = y_train.values.shape
Y_train = y_train.values.reshape(c,)

y_test = newtest[['attack_class']].copy()
c, r = y_test.values.shape
Y_test = y_test.values.reshape(c,)
OneHotEncoder(categorical_features='all', dtype=<class 'numpy.float64'>,
       handle_unknown='error', n_values='auto', sparse=True)

from sklearn.svm import SVC
from sklearn.naive_bayes import BernoulliNB
from sklearn import tree
from sklearn.model_selection import cross_val_score
from sklearn.neighbors import KNeighborsClassifier
from sklearn.linear_model import LogisticRegression
from sklearn.ensemble import VotingClassifier
```

```
# Train KNeighborsClassifier Model
KNN_Classifier = KNeighborsClassifier(n_jobs=-1)
KNN_Classifier.fit(X_train, Y_train);

# Train LogisticRegression Model
LGR_Classifier = LogisticRegression(n_jobs=-1, random_state=0)
LGR_Classifier.fit(X_train, Y_train);

# Train Gaussian Naive Baye Model
BNB_Classifier = BernoulliNB()
BNB_Classifier.fit(X_train, Y_train)

# Train Decision Tree Model
DTC_Classifier = tree.DecisionTreeClassifier(criterion='entropy', random_state=0)
DTC_Classifier.fit(X_train, Y_train);

# Train RandomForestClassifier Model
#RF_Classifier = RandomForestClassifier(criterion='entropy', n_jobs=-1,
random_state=0)
#RF_Classifier.fit(X_train, Y_train);

# Train SVM Model
#SVC_Classifier = SVC(random_state=0)
#SVC_Classifier.fit(X_train, Y_train)

## Train Ensemble Model (This method combines all the individual models above
except RandomForest)
#combined_model = [('Naive Baye Classifier', BNB_Classifier),
#                  ('Decision Tree Classifier', DTC_Classifier),
#                  ('KNeighborsClassifier', KNN_Classifier),
#                  ('LogisticRegression', LGR_Classifier)
#                 ]
#VotingClassifier =  VotingClassifier(estimators = combined_model,voting =
'soft', n_jobs=-1)
#VotingClassifier.fit(X_train, Y_train);

from sklearn import metrics

models = []
#models.append(('SVM Classifier', SVC_Classifier))
models.append(('Naive Baye Classifier', BNB_Classifier))
models.append(('Decision Tree Classifier', DTC_Classifier))
#models.append(('RandomForest Classifier', RF_Classifier))
models.append(('KNeighborsClassifier', KNN_Classifier))
```

```
models.append(('LogisticRegression', LGR_Classifier))
#models.append(('VotingClassifier', VotingClassifier))

for i, v in models:
    scores = cross_val_score(v, X_train, Y_train, cv=10)
    accuracy = metrics.accuracy_score(Y_train, v.predict(X_train))
    confusion_matrix = metrics.confusion_matrix(Y_train, v.predict(X_train))
    classification = metrics.classification_report(Y_train, v.predict(X_train))
    print()
    print('============================== {} {} Model Evaluation
==============================='.format(grpclass, i))
    print()
    print ("Cross Validation Mean Score:" "\n", scores.mean())
    print()
    print ("Model Accuracy:" "\n", accuracy)
    print()
    print("Confusion matrix:" "\n", confusion_matrix)
    print()
    print("Classification report:" "\n", classification)
    print()
```