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

бакалавр
(назва освітнього ступеня)

на тему: Створення модулю виявлення аномалій для системи виявлення вторгнень "Snort"

Виконав: студент IV курсу, групи СБс-43
спеціальності 125 Кібербезпека
(шифр і назва спеціальності)

_____ Фаберський А.М.
(підпис) (прізвище та ініціали)

Керівник _____ Загородна Н.В.
(підпис) (прізвище та ініціали)

Нормоконтроль _____ Тимощук Д.І.
(підпис) (прізвище та ініціали)

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

Рецензент _____ _____
(підпис) (прізвище та ініціали)

Тернопіль - 2024

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 Бакалавр
(назва освітнього ступеня)

за спеціальністю 125 Кібербезпека
(шифр і назва спеціальності)

Студенту Фаберському Андрію Михайловичу
(прізвище, ім'я, по батькові)

1. Тема роботи Створення модулю виявлення аномалій для системи виявлення вторгнень "Snort"

Керівник роботи Загородна Н,В., к.т.н., доц.
(прізвище, ім'я, по батькові, науковий ступінь, вчене звання)

Затверджені наказом ректора від « 14 » 04 2024 року № 4/7-350

2. Термін подання студентом завершеної роботи 27.06.2024 р.

3. Вихідні дані до роботи наукові літературні джерела

4. Зміст роботи (перелік питань, які потрібно розробити)
1. Аналіз предметної області.
2. Теоретична частина.
3. Практична частина.
4. Безпека життєдіяльності, основи охорони праці

5. Перелік графічного матеріалу (з точним зазначенням обов'язкових креслень, слайдів)
1. Титулка. 2. Актуальність. 3. Мета, задачі дослідження. 4. Завдання, які виконує IDS Snort та її функціональна схема. 5. Графічне відображення процесу проходження даних через через IDS Snort. 6. Функціонал роботи пропонованого адаптивного модуля у системі Snort. 7. Схема алгоритму роботи нейромережевого модуля. 8. Засоби програмної розробки. 9. Результат роботи нейромережевого модуля на реальному наборі даних. 10. Звіт про роботу системи Snort. 11. Звіт роботи нейромережевого модуля 12. Основні результати проведеного дослідження

6. Консультанти розділів роботи

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		завдання видав	завдання прийняв
Безпека життєдіяльності, основи охорони праці			

7. Дата видачі завдання 2024 р.

КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назва етапів роботи	Термін виконання етапів роботи	Примітка
1.	Ознайомлення з завданням до кваліфікаційної роботи	14.04 – 18.04	*Виконано*
2.	Підбір джерел про модулі системи Snort	19.04 – 26.04	*Виконано*
3.	Опрацювання джерел про спеціалізовані модулі системи		
	Snort	27.04 – 30.04	*Виконано*
4.	Виконання дослідження щодо створення модулю		
	виявлення аномалій для системи Snort	01.05 – 06.05	*Виконано*
5	Розроблення програмного коду	07.05 – 11.05	
6.	Оформлення розділу «Аналіз предметної області»	12.05 – 15.05	*Виконано*
7.	Оформлення розділу «Теоретична частина»	16.05 – 18.05	*Виконано*
8.	Оформлення розділу «Практична частина»	19.05 – 21.05	*Виконано*
9.	Виконання завдання до підрозділу «Безпека		
	життєдіяльності, основи охорони праці»	14.05 – 21.05	*Виконано*
10.	Оформлення кваліфікаційної роботи	22.05 – 05.06	*Виконано*
11.	Нормоконтроль	06.06 – 12.06	*Виконано*
12.	Перевірка на плагіат	10.06 – 14.06	*Виконано*
13.	Попередній захист кваліфікаційної роботи	15.06 – 18.06	*Виконано*
14.	Захист кваліфікаційної роботи	28.06	

Студент _____ Фаберський А.М.
(підпис) (прізвище та ініціали)

Керівник роботи _____ Загоролна Н.В.
(підпис) (прізвище та ініціали)

АНОТАЦІЯ

Створення модулю виявлення аномалій для системи виявлення вторгнень "Snort" // Фаберський Андрій Михайлович // Тернопільський національний технічний університет імені Івана Пулюя, факультет комп'ютерно-інформаційних систем та програмної інженерії, кафедра кібербезпеки, група СБс–42 // Тернопіль, 2024 // С. – 48, рисунків – 15, таблиць – 5 , слайдів – 12, бібліографи – 20.

Ключові слова: аномалії, IDS, snort, мережа, безпека, аналіз, трафік, сигнатури, атака, фільтрація.

Кваліфікаційна робота присвячена розробці спеціалізованого модуля виявлення аномалій для системи виявлення вторгнення Snort на базі нейромережевого методу аналізу активності в мережі.

Проаналізована предметна область дослідження, Докладно описано структуру, методи роботи та архітектуру системи Snort. Для вирішення завдань, котрі пов'язані із аналізом мережного трафіку та задач запобігання виявлення вторгнень в мережу, варто використати алгоритми машинного навчання чи застосувати можливості нейромереж.

Запропоновано змінити наявну архітектуру Snort шляхом інтегрування в неї додаткового адаптивного нейромережевого модуля (на основі само організаційних карт Кохонена). Він здатний виявити невідомий чи шкідливий трафік. Нейромережевий модуль та стандартний набір правил Snort функціонуватимуть паралельно з метою точнішого виявлення такого трафіку. Розглянуто кластеризаційний алгоритм, із застосуванням якого представлено метод функціонування модуля.

Експериментальні дослідження із застосуванням моделі розробленого модуля підтверджують можливість виявлення вторгнень у мережу.

ANNOTATION

Creating an anomaly detection module for the Snort intrusion detection system // Faberskyi Andrii // Ternopil Ivan Pul'uj National Technical University, Faculty of Computer Information Systems and Software Engineering, Department of Cyber Security // Ternopil, 2024 // P. - 48, Fig. - 15, Table - 5, Slides - 12, References - 20.

Keywords: anomalies, IDS, snort, network, security, analysis, traffic, signatures, attack, filtering.

Thesis deals with the development of a specialized anomaly detection module for the Snort intrusion detection system based on the neural network method of network activity analysis.

The subject area of research is analyzed, the structure, work methods and architecture of the Snort system are described in detail. To solve tasks related to the analysis of network traffic and the tasks of preventing detection of network intrusions, it is worth using machine learning algorithms or applying the capabilities of neural networks.

It is proposed to modify the existing Snort architecture by integrating an additional adaptive neural network module (based on Kohonen self-organizing maps) into it. It is capable of detecting unknown or malicious traffic. The neural network module and the standard Snort ruleset will operate in parallel to more accurately detect such traffic. A detailed clustering algorithm, with the application of which the method of functioning of the module is presented.

Experimental studies using the model of the developed module confirm the possibility of detecting network intrusions.

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, ОДИНИЦЬ СКОРОЧЕНЬ І ТЕРМІНІВ

IDS (Intrusion Detection System) – система виявлення вторгнень.

IPS (Intrusion Prevention System) - система запобігання вторгнень.

SOM (Self-organizing maps) – карти Кохонена, що самоорганізуються.

БД – база даних.

Виявлення вторгнення - це процес моніторингу подій, що відбуваються в комп'ютерній системі або мережі, та їх аналіз на наявність ознак можливих загроз та порушень правил комп'ютерної безпеки, політик допустимого використання або стандартних політик безпеки.

МН – машинне навчання.

МПД – мережа передачі даних.

НД – набір даних.

ПЗ – програмне забезпечення.

ЗМІСТ

ВСТУП

Для виявлення вторгнень використовують два суміжних поняття – IDS та IPS, котрі необхідно розрізняти. IDS є програмним або апаратним компонентом, котрий автоматизує процес виявлення вторгнень. Тоді як IPS є технологією виявлення вторгнень або загроз, а також вжиття заходів для їх захоплення. Вона поєднує знання IDS в автоматичному режимі.

IDS мають велику вагу у захисті безпеки комп'ютерних систем і Internet-мережі, адже вона є активним засобом захисту. IDS безпосередньо і миттєво має змогу виявляти стан мережі, перевіряти потоки і мережі та контролювати її діяльність, а також подавати попередження, проводити запис інформації у БД, на її основі і аналізується вторгнення та створюється журнали вторгнень.

IDS online можна інсталювати на різному вузлі мережі, вказавши будь-які місця. Такі системи можна адаптувати до різних мереж, котрі здатні виробляти складну систему захисту. В даний час багато фірм інсталюють IDS у точках доступу внутрішньої приватної чи, навіть, загальнодоступної мережі. Прикладом можуть бути точки доступу мобільного шлюз або корпоративного комутатора.

Метою роботи є розробка модуля виявлення аномалій для IDS Snort, на основі нейромережевого методу аналізу мережної активності.

В процесі виконання роботи потрібно вирішити наступні завдання:

– дослідження існуючих систем виявлення мережевих атак та розробка структури інтелектуального нейромережевого модуля;

– зіставлення технологій виявлення мережевих атак у трафіку в умовах апріорної інформації;

– розробка структури та алгоритму інтелектуального нейромережевого модуля;

– експериментальна перевірка програмної реалізації нейромережевого алгоритму виявлення мережевих атак.

1 АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ

1.1 Особливості IDS / IPS

IDS/IPS є пристроєм чи програмним додатком, що відстежує мережу [1] або систему щодо шкідливих впливів і порушень політик безпеки.

Про будь-який зловмисний трафік або порушення політик безпеки, зазвичай, повідомляється адміністратору або вся інформація збирається централізовано з використанням системи захисту інформації та управління подіями. IDS/IPS розрізняються за типами подій, які можуть бути розпізнані, та за методами, які використовуються для виявлення інцидентів [2].

За винятком моніторингу та аналізу аномальної активності, всі типи технологій IDS/IPS можуть виконувати такі три функції:

– запис інформації, пов'язаної з виявленими подіями - інформація про події записується локально, після чого вона може бути відправлена в інші системи, наприклад, сервери реєстрації, рішення для безпеки інформації та управління подіями, а також системи управління підприємством;

– повідомлення адміністратора безпеки про виявлені критичні події. IDS/IPS використовує кілька методів для надсилання повідомлень адміністратору таких, як електронна пошта, повідомлення в інтерфейсі користувача IDPS, переривання протоколу SNMP, повідомлення системного журналу, програми користувача або сценарії. Повідомлення про подію включає лише основну інформацію. Адміністратор повинен отримати доступ до IDS/IPS для отримання додаткової інформації;

– звітність - інформація про події підсумовується та може надати додаткову інформацію.

1.2 Варіанти IDS / IPS

На даний момент є значне число систем IDS/IPS, серед інших [3]: Snort,

Suricata, SolarWinds Security Event Manager, OSSEC, Bro, Sagan, Security Onion, Samhain.

Між ними є багато спільного, багато з них фактично є переробленими механізмами виявлення, котрі тісно співпрацюють із базами сигнатур системи Snort. Проте функціональне наповнення перероблених правил є невідомим, здебільного вони тільки відмічають трафіку атрибут «підозрілий» для подальшого його блокування. А вже сам адміністратор мережі повинен на свій розсуд змінити налаштування системи, причому в окремих випадках постає потреба у зміні топології МПД Варто додати, що такі дії із трафіком на вході є для процесора достатньо трудомісткими.

Альтернатива, яку ж систему використовувати, перебуває у залежності від різноманітних факторів та параметрів, наприклад топологія мережі, функціональна складова захисних систем, закладена фінансова складова і, зазвичай, ризиків. Необхідно згадати, що комерційні системи володіють рядом переваг, як то вони постійні оновлення версій, технічна підтримка, наявність сертифікатів, котрі дають їм змогу і право роботи в установах, що забезпечують обробку персональних даних.

Зараз найбільше використовуються дві головні IDS/IPS - це «Snort» і «Suricata», обидві володіють безкоштовною OpenSource – ліцензію для поширення [4]. Система Snort містить значну базу норм і правил та забезпечує її регулярне і постійне оновлення, наявна великий об’єм якісної документації. Вона здатна функціонувати колективно із іншим ПЗ (в т.ч. із Suricata) та здатна забезпечити захист МПД із значними обсягами трафіку.

Основний принцип реалізації завдань обробки інформації та виявлення аномалій розглянемо на системі виявлення «Snort» [5]. Дана система - це безкоштовне і легке ПЗ з відкритим вихідним кодом для виявлення мережевих вторгнень для Linux і Windows, що дозволяє виявляти загрози, що виникають. Створена в 1998 році Мартіном Рошем, засновником і колишнім технічним директором Sourcefire, Snort в даний час розробляється Cisco, яка придбала Sourcefire в 2013 році. В даний час мова опису сигнатур "Snort" фактично є

стандартом абсолютної більшості IDS.

1.3 Структура та методи роботи системи Snort

Система «Snort» дуже гнучка в налаштуванні, залежно від параметрів котрого система може працювати в режимі моніторингу як вузлова чи мережева. Систему Snort можливо встановити безпосередньо в топологію МПД для захисту від атак ззовні [6].

Система Snort призначена виконує такі основні завдання:

– аналіз;
– розбір вмісту пакета;
– виконання протоколювання;
– відображення повідомлення про атаки адміністратору мережі;
– здатність блокувати аномальний трафік.

Система «Snort» спроможна з’ясовувати:

– застосування експлойтів (знаходити Shellcode);
– чи сканується система (юзери, порти, ОС і т.п.);
– різні Backdoors;
– Web -фільтри;
– віруси;
– атаки на служби Telnet, FTP, DNS і т.п.;
– атаки DoS/DDoS;
– атаки, пов'язані з Web- серверами (cgi, php, frontpage, iss і т.д.);
– атаки на БД SQL, Oracle і т.п.;
– атаки з протоколів SNMP, NetBios, ICMP;
– атаки на SMTP, imap, pop2, pop3.

1.4 Архітектура системи Snort

На рисунку 1.1 показана спрощена функціональна схема Snort, котра

містить кілька модулів.

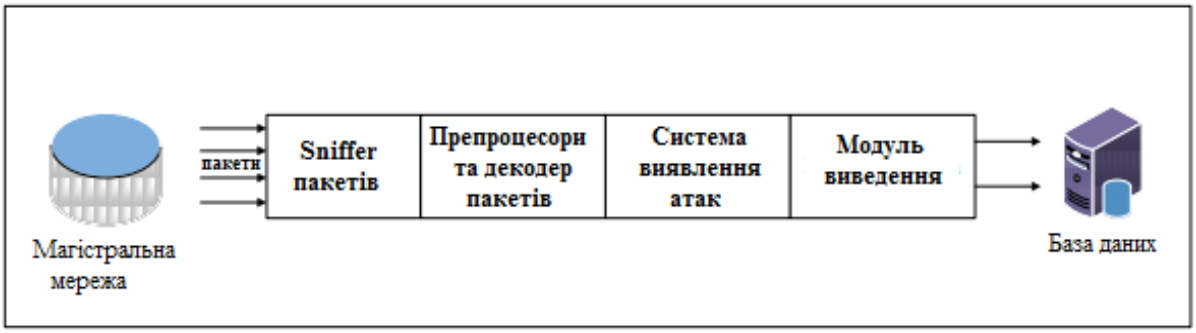

Рисунок 1.1 – Функціональна схема системи IDS Snort

У таблиці 1.1 наведено опис та функціональні можливості кожного з них.

Таблиця 1.1 – Опис модулів системи Snort

Назва модуля	Опис, можливості
Sniffer пакетів	Відповідає за захоплення даних, що передаються по мережі, для подальшої їх передачі на препроцесор і декодер пакетів. Робить це він за допомогою бібліотеки DAQ (Data AcQuisition). Працювати даний сніфер може "в розрив" (inline), пасивному режимі (passive) або читати мережеві дані із заздалегідь підготовленого файлу.
Препроцесори та декодер пакетів	Займається розбором заголовків захоплених пакетів, пошуком аномалій та відхилень від RFC, аналізом TCP- прапорців, виключенням окремих протоколів з подальшого аналізу та іншою аналогічною роботою. Фокусується на стеку TCP/IP. Якщо декодер розбирав трафік на 2-му та 3-му рівні еталонної моделі, то препроцесори призначені для більш детального аналізу та нормалізації протоколів на 3-му, 4-му та 7-му рівнях. Серед найпопулярніших препроцесорів можна назвати frag3 (робота з фрагментованим трафіком), stream5 (реконструкція TCP- потоків), http_inspect_ (нормалізація HTTP -трафіку), DCE/RPC2, sfPortscan (застосовується для виявлення сканування портів) та різні декодери для протоколів Telnet, FTP, SMTP, SIP, SSL, SSH, IMAP тощо. Деякі розробники пишуть свої препроцесори (наприклад, для промислових протоколів) і додають у власні IDS, побудовані з урахуванням «Snort».

Продовження таблиці 1.1

Система виявлення атак	Складається з двох частин. Конструктор правил збирає множину різних вирішальних правил (сигнатур атак) в єдиний набір, оптимізований для наступного застосування підсистемою інспекції захопленого та обробленого трафіку у пошуках тих чи інших порушень
Модуль виведення	Після виявлення атаки Snort може видати (записати або відобразити) відповідне повідомлення у різних форматах - файл, Syslog, ASCII, PCAP, Unified2 (двійковий формат для прискореної та полегшеної обробки). Система Snort прямо підтримує чотири варіанти виводу до БД із використанням своїх модулів виведення. Підтримуються такі формати, як MySQL, PostgreSQL, Oracle iunixODBC. Це дає змогу виконати вимогу більшості користувачів БД. І, закономірно, якщо якась БД не підтримується, тоді варто почати проект створення необхідного модуля розширення. Модуль виведення до БД вимагає як параметрів часу компіляції, і настройок в конфігураційному файлі.

На рисунку 1.2 показано принцип проходження даних через систему Snort

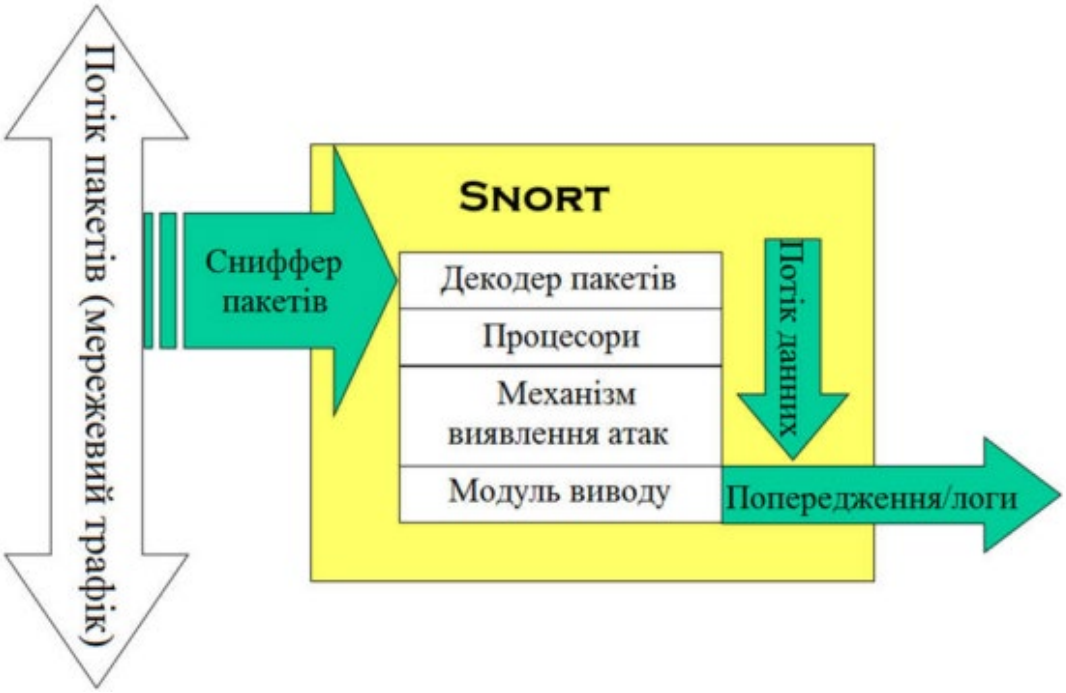

Рисунок 1.2 – Графічне відображення процесу проходження даних через IDS Snort

1.5 Принцип функціонування системи Snort

Варто наголосити, попри те, що наявні регулярні оновлення системи Snort, її функціонал роботи є незмінним [7]. В основі лежить два основні принципи функціонування:

– елементарний аналіз, співставлення атрибутів мережевих пакетів на вході із атрибутами, які визначені базою правил самої системи;

– елементарний аналіз, який базується на розбиранні та перевірянні заголовка мережних пакетів на вході, аналіз інформації, яку містить цей пакет, не передбачається, цей метод є менш ресурсозатратним для системи. Згідно з даними, отриманими при розбиранні заголовка пакету з мережі і порівняння їх із своєю базою норм Snort, система спроможна визначити чи була спроба мережевої атаки. Сигнатурний ж аналіз, зі свого боку, потребує аналізу даних мережного пакету, даний метод є ефективнішим методом виявлення загроз, але водночас і більш трудомістким за кількістю обчислювальних ресурсів.

1.6 Висновки до першого розділу

Сучасні IDS / IPS зайняли надійну нішу в методах виявлення шкідливих атак, але здебільшого навіть при їх використанні інформаційні системи залишаються вразливими.

IDS / IPS однаково реагують на події, що повторюються, які не відповідають їхнім правилам у БД. При цьому такі системи діють точно до заданих наборів правил у їхній БД, їм необхідний проміжок часу для оновлення системи, а також час на обробку нової мережевої атаки експертами та додавання інформації про неї до існуючих БД.

Ґрунтуючись на перерахованих вище проблемах, рішенням може послужити використання алгоритмів МН або застосування нейронних мереж при вирішенні завдань, пов'язаних з аналізом мережного трафіку і завдань запобігання виявлення вторгнень в мережу.

2 ТЕОРЕТИЧНА ЧАСТИНА

2.1 Опис принципів вирішення завдань виявлення вторгнень

На рисунку 2.1 наведено процес функціонування IDS, який в загальному складається із таких етапів [8]:

- контроль;
- аналіз;
- оповіщення;
- реагування на виявлення.

Рисунок 2.1 – Етапи процесу виявлення вторгнень

Насправді, для повнішого механізму виявлення нейромережеві технології застосовуються разом із IDS / IPS [9], рішення, із застосуванням нейронних мереж спираються на такі технології аналізу.

Методи виявлення відомих порушень:

- сигнатурний аналіз. Порівняння сигнатури даних, чи сигнатури поведінки з сигнатурою в оновлюваній базі. Сигнатура може бути представлена також шаблоном або регулярним виразом;

– експертні системи з урахуванням правил.

Методи виявлення аномалій:

– детектори, що реагують на наведення порогового значення заданого

параметра;

– статистичні системи (системи класифікаторів навчання, наївний класифікатор Байєса);

– поведінковий аналіз (можливо віднести до нечітких правил);

– використання моделі "штучна імунна система".

Ці рішення мають свої недоліки:

– сигнатурний аналіз працює строго за заданими правилами системи Snort, він не зможе виявити невідому атаку;

– експертні системи також працюють за заданими правилами, як і сигнатурний аналіз;

– для сигнатурного та експертного аналізу необхідна підтримка актуальних баз правил;

– системи правил часто здатні виявити атаку за невеликого відхилення послідовності дій, за принципом порівняння діяльність-правило. Такі системи, що ґрунтуються на правилах, частково вирішують проблеми з пошуком аномалій, але такі системи можуть сильно збільшити кількість хибних спрацьовувань;

– системи на основі правил часто не мають достатньої гнучкості в структурі правил;

– статистичні системи не чутливі до порядку прямування подій (що правильно не для всіх існуючих систем);

– для них і для порогових детекторів важко задати порогові значення відстежуваних систем виявлення атак характеристик;

– при тривалому використанні статистичних систем виникає небезпека перенавчання, коли атакуючі дії розглядаються як нормальні.

Нейронні мережі не варто розглядати як повну заміну IDS / IPS, а як доповнення до статистичних систем, але в деякій мірі вони можуть замінити сигнатурний пошук та інші методи [9].

Нейронні мережі мають свої переваги:

– можливість аналізу неповних вхідних даних або зашумленого

сигналу;

– відсутність необхідності формалізації знань (замінюється навчанням);

– відмовостійкість системи;

– можливість простого розпаралелювання роботи;

– потребують меншого втручання оператора;

– існує можливість виявлення невідомих атак;

– мережа здатна навчатися автоматично та в процесі роботи;

– обробка величезної кількості даних.

Нейронні мережі також мають недоліки:

– основні методи нейронних мереж часто не призводять до однозначних рішень;

– для роботи нейронних мереж необхідно пройти процес навчання, що потребує значних часових витрат;

– для того, щоб навчити мережу, треба підготувати навчальну та тестову вибірки, що не завжди просто;

– навчання мережі часто відбувається неоднозначно і в деяких випадках мережу необхідно перенавчати;

– роботу нейронної мережі трактувати однозначно неможливо;

– важко пояснити, чому мережа прийняла те чи інше рішення (проблема вербалізації);

Отже, неможливо гарантувати повторюваність та однозначність отримання результатів.

В таблиці 2.1 наведені описані вище методи виявлення вторгнень.

Таблиця 2.1 – Характеристики методів виявлення вторгнень

Методи	Рівень спостереження	Верифі-кація	Адаптив-ність	Стійкість	Обчислювальна складність
Аналіз сигнатур	Хост, мережа, додатки	Так	Ні	глобальна	*O(log n)*
Статистичний аналіз	Хост, мережа	Ні	Так	локальна	*O(n)*
Аналіз систем станів	Хост, мережа, додатки	Так	Ні	локальна	*O(n)*
Графи сценаріїв атак	Хост, мережа, додатки	Так	Так	локальна	*NP*
Експертні системи	Хост, мережа	Так	Так	глобальна	*NP*
Методи засновані на специфікаціях	Мережа	Так	Ні	локальна	*O(log n)*
Нейронні мережі	Хост, мережа, додатки	Так	Так	локальна	*O(n)*
Імунні мережі	Хост, мережа	Ні	Так	локальна	*O(n)*
Кластерний аналіз	Хост, мережа, додатки	Ні	Так	локальна	*O(n)*
Поведінкова біометрія	Хост	Ні	Так	локальна	*O(n)*

Так як кількість нових атак, котрі не були відомі раніше, постійно збільшується, тому саме адаптивні методи виявлення вторгнень і будуть найкращими для застосування [10]. Саме тому в кваліфікаційній роботі для створення адаптивного модуля і будуть використовуватися адаптивні методи виявлення аномалій (вторгнень).

2.2 Функціонал роботи адаптивного модуля у системі Snort

Ключове рішення – необхідно виконати заміну наявної архітектури Snort, відображеної на рисунку 1.1, і інтегрувати в неї ще один модуль. Це дозволить створити нову архітектуру системи, наведеної на рисунку 2.2.

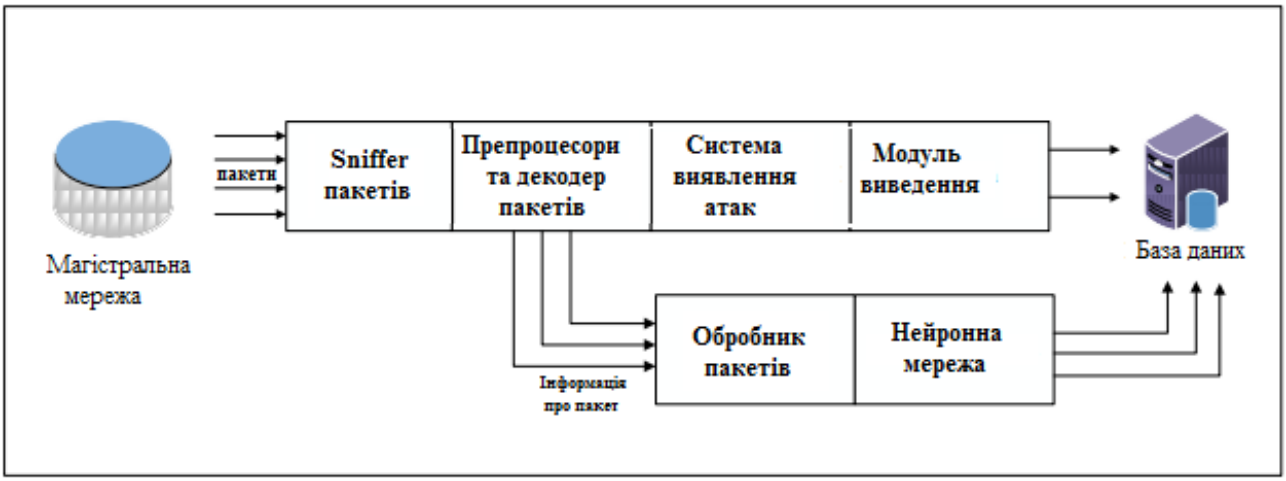

Рисунок 2.2 – Функціональна схема системи виявлення вторгнення IDS Snort із створеним адаптивним модулем

Розроблений нейромережевий модуль функціонує спільно із набором норм системи Snort. Для чого ж паралельна робота модулі із цим набором? Це потрібно, бо правила Snort дають змогу виявляти тільки відомий шкідливий трафік. А наш адаптивний модуль здатний виявити також і невідомий чи змінений варіант шкідливого трафіку, котрий, і собі, здатен зменшити число помилкових визначень. Це дозволить значно зменшити число хибнопозитивних сигналів виявлення ти покращити точність їх детектування.

Мережевий трафік через препроцесор буде передаватися у нейромережевий модуль та у набір норм Snort. Ці дві складові функціонуватимуть паралельно з метою точнішого детектування шкідливого трафіку.

2.3 Передобробка та подання вхідних даних

Одержання надійних даних є самою важливою частиною процесу роботи з моделями МН і є серйозним питанням, так як відкритість НД надзвичайно мала. З однієї сторони, значна кількість таких наборів є внутрішніми і не можуть бути застосовані у відкритому доступі з огляду на конфіденційність інформації, а з іншої - НД дуже безіменні і не відтворюють наявні представлення удосконалення мережі, чи не містять визначених статистичних параметрів, саме тому ідеального НД ще не отримано [11]. Його просто немає. Отож, пошукачі повинні використовувати наявні НД, котрі, як наслідок, досить часто не відповідають критеріям оптимальності. По ходу змінення поведінки мережі і її моделей, а також удосконалення систем вторгнень, з'явилася потреба переходу від незмінних та одноразових НД до значно динамічніших, котрі не лише показують структуру трафіку, але й змінюються, розширюються і відтворюються.

Для визначення вхідних параметрів в роботі буде застосовуватися НД CSE-CIC-IDS2018 Канадського інституту кібербезпеки, котрий був запропонований якраз для систем аналізу, тестування та оцінки IDS із нахилом на мережні виявлення аномалій, у цьому НД розмежовані дані вторгнення в мережу, атаки та цінні зв'язки. Ця БД містить стандартний НД, який включає великий вибір втручань, котрі симулюються в середовищі військової мережі. Він містить 4898431 екземпляр із 41 атрибутом.

Будь-яке з'єднання має позначення нормального чи атаки із визначеним типом. Запис такого з'єднання містить біля 100 байт.

Тут виділяють для атак три групи:

- DOS: відмова у обслуговуванні;
- R2L: несанкціонований доступ із віддаленої машини;
- U2R: несанкціонований доступ до локальних кореневих привілеїв.

Дані необхідно певним чином підготувати перед їх застосуванням в алгоритмі МН. Якісь компоненти можна достатньо легко віднайти, тоді як інші можуть бути визначені тільки із використанням певних експериментів чи

шляхом проведення спеціальних тестів. Застосування усіх функцій НД не обов'язково і не завжди є гарантією максимально хороших параметрів IDS. Це може призвести до збільшення обчислювальних затрат, та навіть частоти появи різних помилок в системі. Пов’язано це з тим, що окремі функції є занадто складними чи непотрібними для визначення відмінностей у різних класах.

Основним внеском цього НД є запровадження атрибутів, котрі пропоновані експертом. Вони дозволяють зрозуміти поводження різноманітних типів атак, в т.ч. головні параметри виявлення DOS, PROBE, R2L і U2R. Нижче буде наведено перелік вхідних характеристик, котрі використані для навчання.

Таблиця 2.2 – Вхідні параметри для навчання

Назва	Опис
duration	Тривалість з'єднання в секундах
protocol_type	Тип використовуваного протоколу, тобто TCP, UDP та ін.
service	Тип використовуваних послуг, тобто http, ftp, telnet та ін
flag	Прапор з'єднання: норма чи помилка
scr_bytes	Число байт даних від джерела до одержувача
dst_bytes	Число байт даних від одержувача до джерела
land	1 якщо з'єднання з/на такому ж хості/порті
wrong_fragments	Кількість "невірних" фрагментів
urgent	Кількість термінових (urgent) пакетів
count	Кількість підключень до цього хоста за останні 2 секунди
srv_count	Число підключень до цього сервісу за останні 2 сек.
serror_rate	Відсоток підключень із SYN помилками
diff_srv_rate	Відсоток підключень до різних сервісів
srv_diff_hast_rate	Відсоток підключень до різних хостів
dst_host_srv_count	Кількість з'єднань до локального хоста, встановлених віддаленою стороною і використовують одну і ту ж службу

2.4 Опис принципу побудови нейромережевого аналізатора

Всі детектори аномалій, безвідносно від тієї моделі, котра застосована при їх формуванні, працюють на засадах зіставлення даних зразка із наявною сукупністю факторів та сигналізації, якщо трапиться вихід за визначене порогове значення. Будь-яка модель зобов’язана володіти таким стандартом. Що стосується нейромережі, еталон створюється під час навчання упродовж нормального функціонування МПД, після цього така мережа, працюючи, детектує відхилення, які виходять за зазначені границі [12].

Складності цього підходу:

– мережу потрібно навчити за нормальної роботи МПД, проте відсутня впевненість, що в даний час атака не проводиться, тобто мережа не набуде навчена при так званій "аномальній активності";

– припинити роботу МПД або від’єднати її від зовнішніх мереж також не вдасться (оскільки при перериванні її роботи, у подальшому активність під’єднаної МПД вже визначатиметься аномальною);

– обсяг даних, отже, і період навчання, може бути досить великим (робота МПД практично точно різниться у різні періоди діб і може відрізнятися поквартально);

– уникнути проблему навчання за нормальної роботи не є можливим (її розв’язують через побудову правил, котрі окреслюють роботу МПД за нормального режиму, проте в цьому випадку це не застосовується).

Після проведеного аналізу існуючих системи, детектувати аномалії найкраще за допомогою кластеризації із застосуванням "карт активності мережі". Зазвичай, задача кластеризації в межах застосування нейромережевого критерію розв’язується із застосуванням SOM.

2.5 Алгоритм кластеризації, SOM

Основною ідеєю SOM є перетворення потоку вхідного сигналу будь-

якого виміру в одно- або двовимірне дискретне відображення і самоналаштовано провести його впорядковане перетворення. Через те і формується карта, поміщаючи нейрони у вузлах одно- чи двовимірної решітки. Також є можливим застосування карт із більшим розміром також можливе.

При навчанні у змагання нейрони випадковим чином налаштовуються на різноманітні вхідні паттерни або цілі класи таких паттернів. Розміщення нейронів (інакше нейронів, котрі виграють) буде упорядковуватися, і на решітці формується усвідомлена система координат [13]. Тобто, SOM виробляє потрібну топографічну карту паттернів на вході.

В роботі варто зробити фокус на визначеному виді SOM, котрий носить назву мережа Кохонена. Такий SOM володіє структурою прямого зв'язку із цілісним обчислювальним шаром, котрий міститься у рядках і стовпцях. Будь-який із цілком зв'язаний із всіма властиво вихідними вузлами власне вхідного шару (див. рисунок 2.3).

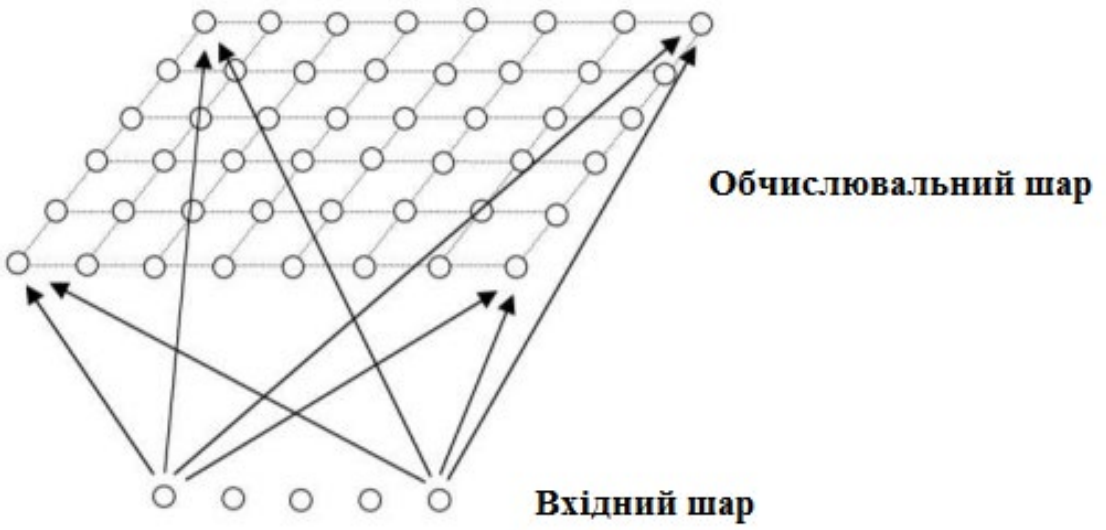

Рисунок 2.3 – Представлення карти Кохонена

Як видно одновимірна карта володіє тільки одним рядком (чи одним стовпцем) у визначеному обчислювальному шарі.

Перебіг самоорганізації складається із чотирьох основних складових [13]:

– ініціалізація: усі ваги з’єднань отримують початкові значення малими випадковими величинами;

– конкуренція: розраховують визначені ваги дискримінантної функції для кожного шаблону на вході, котра забезпечує властиво базу для власне конкуренції. Переможцем буде визначено той нейрон, який володітиме мінімальним значенням такої функції;

– кооперація: переможний нейрон встановлює розташування у просторі топологічного околу нейронів, які є збудженими, водночас гарантуючи базу для кооперації між нейронами, що є суміжними;

– адаптація: ті нейрони, які є збудженими, знижують свої особисті величини дискримінантної функції щодо до вхідного паттерну, при допомозі корегування значення з'єднання, таким відгук нейрона-переможця надалі для використання такого ж вхідного паттерну зростає.

2.5.1 Конкурентний процес

Якщо вхідний простір є D-вимірним (тобто D вхідних одиниць), ми можемо записати вхідні шаблони як $x = \{x_i : i = 1, D\}$, ваги з'єднання між вхідними одиницями і та нейронами j у обчислювальному шарі можуть бути записані $w_j = \{w_{ji} : j = 1, \ldots, N; i = 1, \ldots, D\}$, де N – загальна кількість нейронів.

Потім ми можемо визначити нашу дискримінантну функцію, що дорівнює квадрату евклідової відстані між вхідним вектором x і вектором ваги w_j для кожного нейрона j

$$d_j(x) = \sum_{i=1}^{D}(x_i + w_{ji})^2 \qquad (2.1)$$

Іншими словами, нейрон, вектор ваг якого найближче до вхідного вектора (тобто найбільш схожий на нього), оголошується переможцем.

Таким чином, безперервний вхідний простір можна порівняти з

дискретним вихідним простором нейронів за допомогою простого процесу конкуренції між нейронами.

2.5.2 Кооперативний процес

У нейробіологічних дослідженнях було виявлено, що у наборі порушених нейронів спостерігається «колективна» взаємодія. Коли спрацьовує один нейрон, його найближчі сусіди, зазвичай, збуджуються сильніше, ніж ті, хто знаходиться далі. Існує топологічна околиця, яка розпадається з відстанню.

Визначимо схожу топологічний окіл для нейронів у нашій SOM. Якщо S_{ij} - бічна відстань між нейронами *i* та *j* на сітці нейронів, ми приймаємо в якості нашого топологічного околу, де *I(x)* - індекс нейрона, що виграв

$$T_{j,I(x)} = \exp\left(-S_{j,I(x)}^2/2\sigma^{2}\cdot\right) \tag{2.2}$$

Топологічний окіл має кілька важливих властивостей: є максимальним у нейрона, що виграв, симетричним щодо цього нейрона, монотонно зменшується до нуля в міру збільшення відстані, і він є трансляційно інваріантним (тобто не залежить від місця розташування переможного нейрона).

Особливістю SOM є те, що розмір σ околу має зменшуватися з часом. Популярна залежність від часу - експоненційний спад:

$$\sigma(t) = \sigma_0 \exp\left(-t/\boldsymbol{\tau}_{\sigma}\right) \tag{2.3}$$

2.5.3 Адаптивний процес

Очевидно, що SOM повинна включати певний адаптивний або навчальний процес, за допомогою якого виходи самоорганізуються і формується карта характеристик між входами і виходами.

Сенс топографічного околу полягає в тому, що не тільки нейрон, що переміг, оновлює свої ваги, але і його сусіди будуть оновлювати свої ваги, хоча і не так сильно, як сам переможець. На практиці відповідне рівняння оновлення ваги полягає в тому, що ми маємо залежну від часу (епохи) *t* швидкість навчання $\eta(t) = \eta_0 \exp(-t/\tau_\eta)$ та оновлення застосовуються до всіх схем навчання х упродовж багатьох епох.

$$\Delta w_{ji} = \eta(t) T_{j,I(x)}(t)(x_i - w_{ji}) \tag{2.4}$$

Ефект кожного оновлення ваги навчання полягає в переміщенні векторів ваги w_i нейрона, що виграв, і його сусідів до вхідного вектора х. Повторне подання даних навчання, таким чином, призводить до топологічного впорядкування.

2.5.4 Упорядкування і конвергенція

За умови, що параметри $(\sigma_0, \tau_\sigma, \eta_0, \tau_\eta)$ обрані правильно, ми можемо почати з початкового стану повного безпорядку, і алгоритм SOM поступово призведе до організованого представлення патернів активації, взятих із вхідного простору. Однак можна опинитися у метастабільному стані, у якому карта об'єктів має топологічний дефект.

Є два ідентифіковані етапи цього адаптивного процесу:

– фаза впорядкування чи самоорганізації – під час якої відбувається топологічне впорядкування векторів ваги. Як правило, це займе цілих 1000 ітерацій алгоритму SOM і необхідно ретельно розглянути вибір параметрів сусідства та швидкості навчання;

– фаза конвергенції – під час якої карта об'єктів точно налаштовується та забезпечує точну статистичну кількісну оцінку вхідного простору. Як правило, число ітерацій на цьому етапі буде принаймні в 500 разів більше, ніж

число нейронів у мережі, і знову параметри повинні бути ретельно обрані.

2.6 Висновки до другого розділу

В даний час модерні IDS / IPS успішно розвиваються до задач комплексного вирішення і відчувають потребу не тільки успішно справлятися із своїм завданням із детектування вторгнень, але і давати на розгляд системним адміністраторам набір ПЗ для діагностики мережі. Водночас вони не забувають впроваджувати в життя своє головне завдання, застосовуючи методи детектування аномалій, як і методи детектування тих порушень, котрі їм відомі.

При аналізі предметної області було виявлено, що на сьогоднішній день спостерігається нестача готових рішень з цієї проблеми. У теоретичній частині для вирішення задач нейромережевого детектування застосвуються методи на базі персептронів із багатьма шарами, рециркуляційних власне нейронних мереж, на базі SOM.

Наведені методи для роботи із нейронними мережами демонструють значне зростання точності детектування аномалій при проведені аналізу МПД. Тим не менше вони успішно допомагають при формуванні топології мережі та віднаходженні помилок у тих мережах, що функціонують.

Основна відмінність описаного вище методу карт, котрі самоорганізуються, у тому, що метод позбавляє необхідності навчати нейромережу на наперед доведеному до готовності НД, що навчається. Такий метод корисний для вирішення поставлених завдань у роботі та нейромережевий модуль навчатиметься безпосередньо в момент нормальної роботи МПД.

3 ПРАКТИЧНА ЧАСТИНА

3.1 Алгоритми роботи нейромережевого модуля

В основі навчання розробленого модуля лежить той факт, що мережа піддається навчання упродовж визначеного часу при нормальній роботі МПД. За таких обставин формується певне число кластерів, котрі за правильного налаштування мережі цілком описують нормальну роботу МПД.

Використовують два варіанти детектування аномалій.

За першого мережа, котра є попередньо навченою на тестових НД, володіє кількома кластерами, у випадку поступлення аномальних даних структура та чисельність кластерів буде мінятися.

Для другого варіанту застосовується передавання нових даних для перевіряння попередньо навченої мережі та зіставлення того, як точно вони пасують до одного із побудованих кластерів.

У цих випадках аномалії можливо виявити через зміну кількості існуючих кластерів у моделі для навчання, рухові швидкості змінення формування та знищення нейронів і їх зв'язків. І тим не менше ймовірно детектувати аномалії при відхиленні величин нових вимірів від середньої величини наявних нейронів, на котрих мережа властиво і робила навчання.

У методах діяльностей в мережі динамічні обчислення проводити достатньо складно, нейромережа здатна регулювати чисельність нейронів і зв'язків між ними при кожному повторенні, хоча певні алгоритми формують та знищують встановлену чисельність нейронів при кожному повторенні. В той самий час пертурбація числа кластерів не є вірним параметром детектування загрози. Одним із наюільш вірних варіантів є пертурбація відхилень від величини наявних нейронів, власне на яких мережа і проводила навчання.

Діаграма роботи нейромережевого модуля наведена на рисунку 3.1.

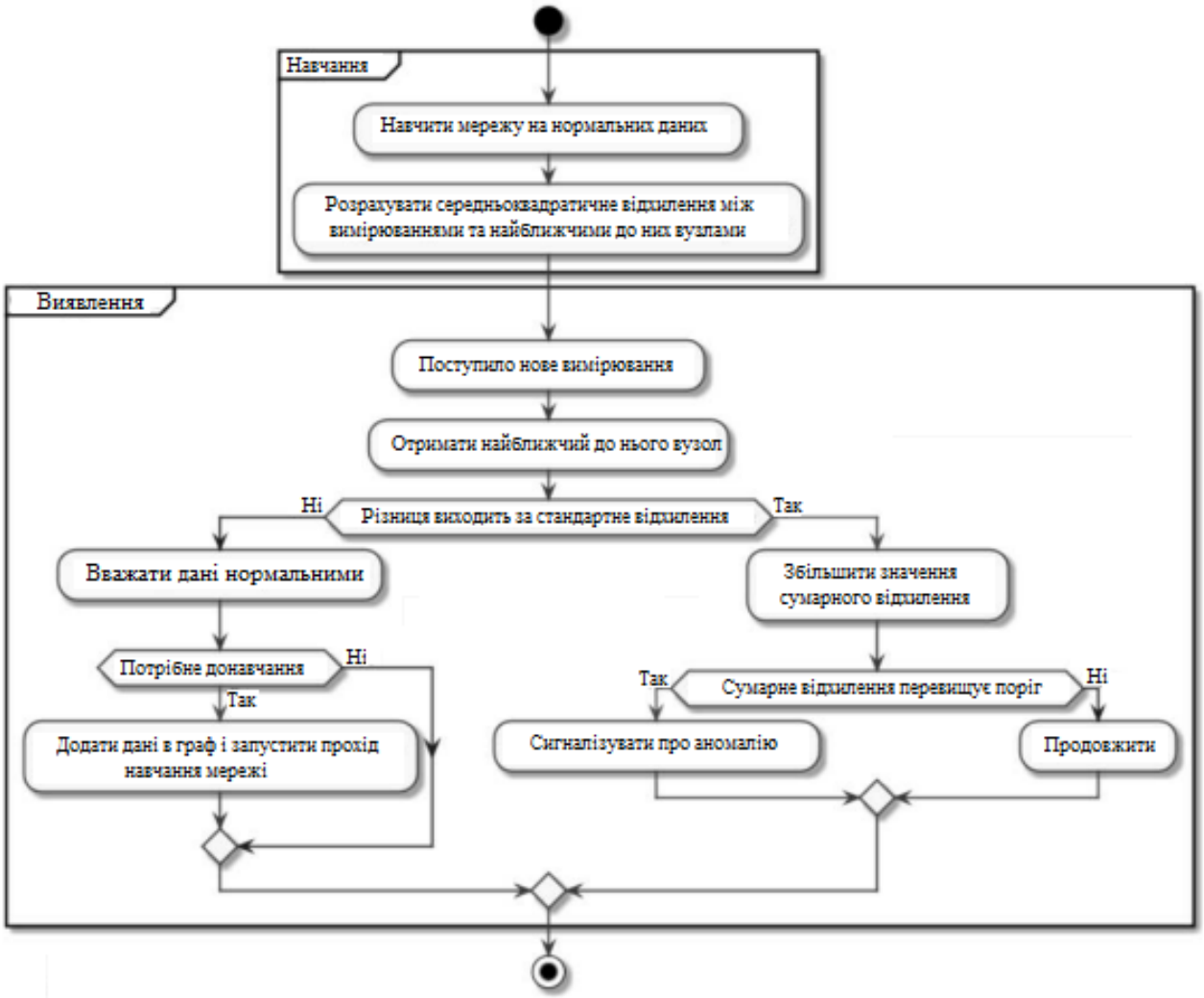

Рисунок 3.1 – Схема алгоритму роботи нейромережевого модуля

З рисунка видно, що в цьому методі можна застосувати варіант адаптивного підлаштування, котре подібне до напівавтоматичного навчання зі вчителем.

Для повнішого і точнішого детектування алгоритм, відображений на рисунку 3.1, знаходить не одне відхилення, а обчислює середнє між найкращими нейронами. Варто наголосити, що для кожного кластера нейронів рахується середнє арифметичне значення усіх відхилень, котрі містяться у ньому.

Для досі невідомих НД, які подані для перевіряння нейромережевим модулем, шукається самий ближчий нейрон з тих, які скомплектовані на стадії навчання та обчислюється відстань між ними, у подальшому саме ця відстань і зіставляється із середньоквадратичним відхиленням для кластера, в якому

знаходиться цей нейрон.

Коли така одержана відстань між нейроном і тими даними, котрі надані для перевірки, є більшою, ніж середнє відхилення за кластером, такий НД і буде вважатися з аномаліями.

3.2 Огляд та підсумки функціонування із тестовими даними

З метою тестування роботоздатності запропонованого модуля його було перевірено на НД CSE-CIC-IDS2018, який розробив Канадський інститут кібербезпеки. Нейромережевий модуль здійснює навчання на такій навчальній вибірці з цього НД, котра містить 1011 запис, потім той навчальний НД, знову відправляється на модуль з метою перевірки однозначності такого навчання.

Модуль, який вже попередньо навчали, було використано до повних даних з тої самої множини, з котрої і була сформована навчальна вибірка, підсумки функціонування наведено в таблиці 3.1.

Таблиця 3.1 – Результат роботи нейромережевого модуля на реальному НД

Змінна	Значення	Опис
l_time	1869.00	Час навчання за секунди
te_l_time	39.5	Час перевірки в секундах на наборі, з якого було сформовано навчальну вибірку
te_t_time	125.61	Час перевірки в с на повному наборі даних тестової вибірки
g_l_perc	69.5	Відсоток знайдених аномалій для повного набору, з якого було сформовано навчальну вибірку
g_t_perc	78.4	Відсоток знайдених аномалій для повного набору даних тестової вибірки
f_l_perc	0.00	Відсоток хибних спрацьовувань для повного набору, з якого була сформована навчальна вибірка
f_t_perc	36.9	Відсоток хибних спрацьовувань для повного набору даних тестової вибірки

Мовою розробки модуля був обраний Python із бібліотекою нейромереж Scikit-Leam [14]. Вона містить реалізації значного числа алгоритмів МН, котрі базуються на алгоритмах навчання (як з учителем, так і без нього). Задача завантаження вхідної інформації не входить до Scikit-Leam., тому з цією метою використана бібліотека NumPy.

З метою забезпечення візуалізації даних застосовується спеціалізована бібліотека NetworkX. Основне її призначення - опрацювання мережевих структур і графів.

Для роботи модуля необхідні дані, котрі поступають від препроцесора системи Snort на обробку, застосовується стандартний модуль csv [15]. Вони пересилаються у виді ненормалізованого масиву NumPy, із використанням бібліотеки Scikit-Learn такий масив даних стає нормалізованим, відтак після цього робота вже проводиться тільки з даними, які вже є нормалізованими. На базі зчитаних даних формується ненаправлений граф, котрий і буде застосовуватися для візуалізації, додатково буде побудовано порожній граф, для додавання до нього нейронів і зв'язків між ними. Згодом багато разів відбувається виклик кластеризуючого методу, куди передається одне з значень, що є набором координат точки в багатовимірному просторі. Одержуючи дані цієї точки, метод, котрий кластеризує, здійснює вибір самих ближчих нейронів внутрішнього графа і, за умови, що вони не є задовільними, формує нові нейрони та зв'язки.

На рисунку 3.2 наведено фрагмент лістингу коду нейромережевого модуля.

```
    def train(self, max_iterations=10000, save_step=50, stop_on_chi=False):

        self._dev_params = None
        self._save_img(self._fignum, 0)
        graph = self._graph
        max_nodes = self._max_nodes
        d = self._d
        ld = self._lambda
        alpha = self._alpha
        update_winner = self.__update_winner
        data = self._data
        start_time = self._start_time = time.time()
        train_step = self.__train_step

        for i in range(1, max_iterations):
            tm = time.time() - start_time
            print(f'Training time = {round(tm, 2)} s, ' +
                  f'Time per record = {tm / len(data)} s ' +
                  f'Training step = {i} / {max_iterations}, Clusters count =
{self.number_of_clusters}, ' +
                  f'Neurons = {len(self._graph)}'
                  )
            for x in data:
                update_winner(x)
```

Рисунок 3.2 – Частина коду розробленого модуля

Кластеризуючий метод раз за разом записує відображення з візуалізацією, він викликається два рази: спершу для вимальовки даних, другий раз для відображення нейромережі зверху даних. По закінченню навчання усі записані зображення заливаються у gif.

Встановлені параметри для тестування даних наведені у таблиці 3.2.

Таблиця 3.2 – Параметри тестових даних

Назва параметру	Його величина
Кількість кроків навчання для SOM	7000
Кількість нормальних записів у навчальній вибірці	516
Кількість аномальних записів у навчальній вибірці	495
Кількість нормальних записів у тестовій вибірці	2152
Кількість аномальних записів у тестовій вибірці	9698
Повний розмір тестової вибірки	11850

3.3 Огляд функціонування візуалізаційного модуля

З метою відкритого і явного відтворення даних окремо було написано візуалізаційний модуль, на рисунку 3.3 – 3.5 наведені візуалізації карт активності мережі, які сформовані на тестовому НД.

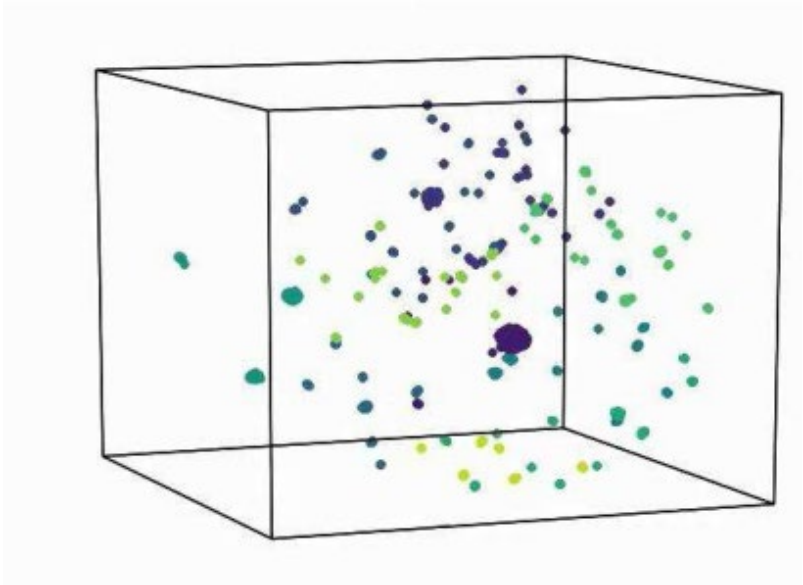

Рисунок 3.3 – Карти мережевої активності, побудовані на навчальній вибірці, що містить лише «аномальний» трафік

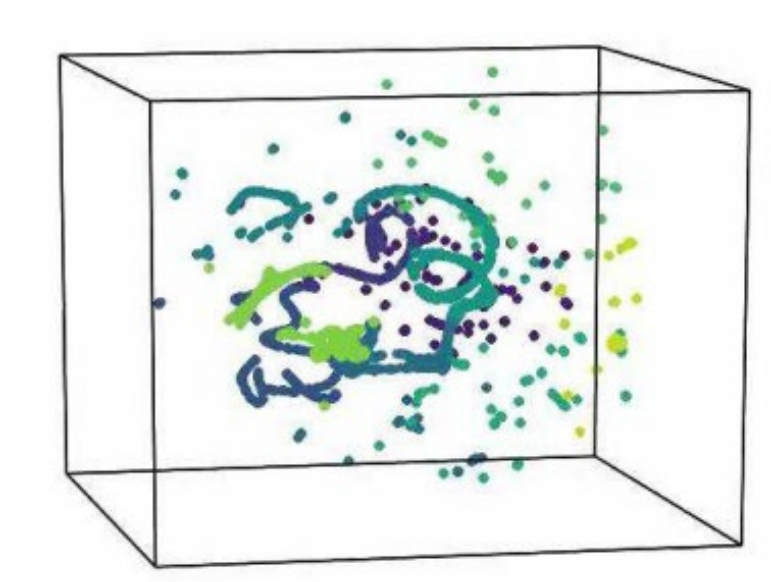

Рисунок 3.4 – Карти мережевої активності, побудовані на навчальній вибірці, що містить лише «нормальний» трафік

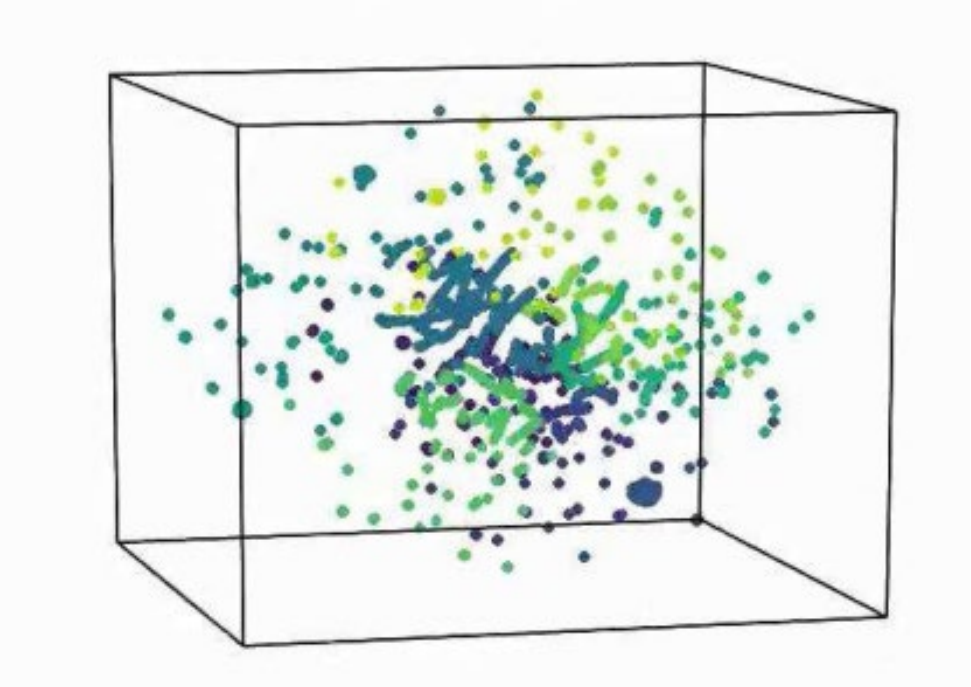

Рисунок 3.5 – Карти мережевої активності, побудовані на повній навчальній вибірці

Ці карти були створені із використанням бібліотеки HyperTools [16], знову ж таки мовою Python.

Бібліотека, яка візуалізує та керує багатовимірними даними, фактично є множиною засобів Python для одержання геометричного представлення про них. В її базі лежить Matplotlib та Seaborn (для формування графіків), а також Scikit-Leam (для операцій із даними).

Інформація була представлена при допомозі звичайного модуля CSV, дані зчитуються із НД CSE-CIC-IDS2018 (як вже було згадано раніше).

На рисунку 3.6 представлено фрагмент коду модуля візуалізації.

```
import hypertools as hyp
from collections import OrderedDict
import csv
import numpy as np

def read_ids_data(data_file, labels_file='CSE-CIC-IDS2018 Names.csv'):
    selected_parameters = ['duration', 'protocol_type', 'service', 'flag', 'src_bytes',
'dst_bytes', 'land', 'wrong_fragment', 'urgent', 'count', 'srv_count', 'serror_rate',
'diff_srv_rate', 'srv_diff_hast_rate', 'dst_host_srv_count']

    hyp.plot(np.array(result), '.', reduce='UMAP', ndims=3,
             n_clusters=10, animate='spin', palette='viridis',
             duration=3, rotations=1, legend=False, explore=False, show=True,
             save_path='./video.mp4')

read_ids_data('CSE-CIC-IDS2018.csv')
```

Рисунок 3.6 – Частина лістингу коду модуля візуалізації

В основі функцій кластеризацій лежить методу K-Means, для тестового набору даних даний метод прийнято використовувати, так як заздалегідь відоме число рядків та число груп трафіку різного роду. Цей алгоритм аналізує всі вхідних елементи на наперед відоме певне число кластерів. Його функціонування потребує завчасного задання кількості потрібних кластерів, що виключає можливість застосування даного алгоритму при роботі з реальними даними на МПД.

У візуалізаційних картах кластери властиво є рядом векторів з відстанню між ними в такій групі меншою, аніж відстань до груп-сусідів.

Тут застосовується U-аггау, стандартизована матриця відстаней для формування карт із тривимірними ґратками. Візуалізація даних проходить із застосуванням двовимірних карт, проте при візуалізації великорозмірних даних у просторі з меншою розмірністю втрачаються топологічні зв’язки. При використанні цього методу проходить знаходження відстані між векторами ваг нейрона у нейромережі та його сусідами, які знаходяться найближче. У подальшому ці значення використовуватимуться з метою визначення кольору, котрим властиво вузол буде відрисовано.

3.4 Огляд та підсумки функціонування на реальних даних

З метою перевіряння роботоздатності нейромережевого модуля було виконано записування дампу трафіку при допомозі Wireshark (спеціального ПЗ для аналізування трафіку у комп'ютерних мережах) [17]. Записування відбувалося із використанням реальних даних упродовж 60 секунд з мережі інтернету університету, мультисервісної корпоративної інформаційно-обчислювальної мережі та склала тестову вибірку, кількість рядків становила 936 840.

На рисунку 3.7, 3.8 відображені звіти про виконану роботу системи Snort.

Рисунок 3.7 – Звіт про роботу системи Snort (частина 1)

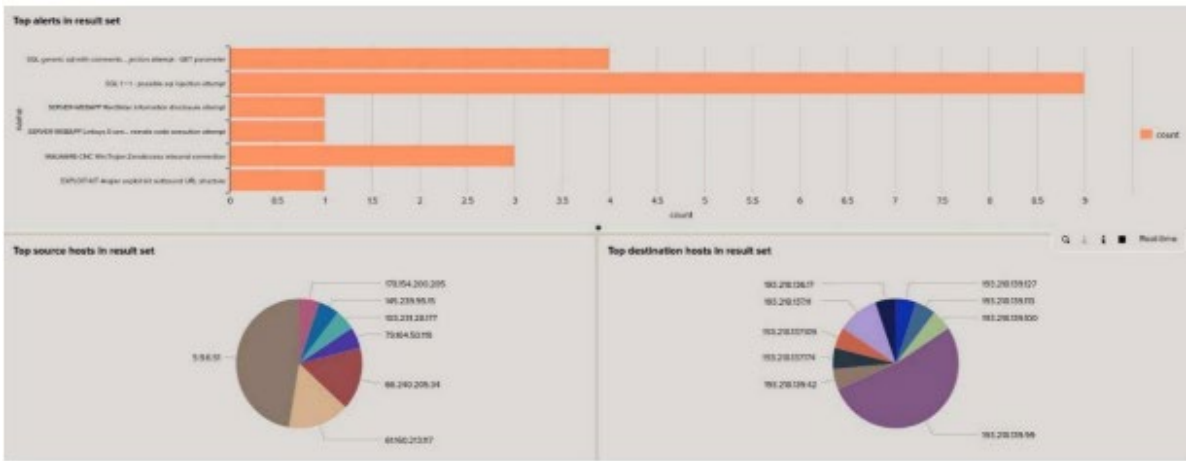

Рисунок 3.8 – Звіт про роботу системи Snort (частина 2)

Цей тестовий набір розділено на 2 частини наступним чином: для навчання взято перші 200 тисяч його рядків; а вже для тестування модуля взято решту записів, котрі залишилися (736 840).

При допомозі візуалізаційного модуля вдалося створити графічні відображення самоорганізованої карти, після проходження 46 тисяч кроків та після 109 тися кроків навчання, які показані на рисунку 3.9.

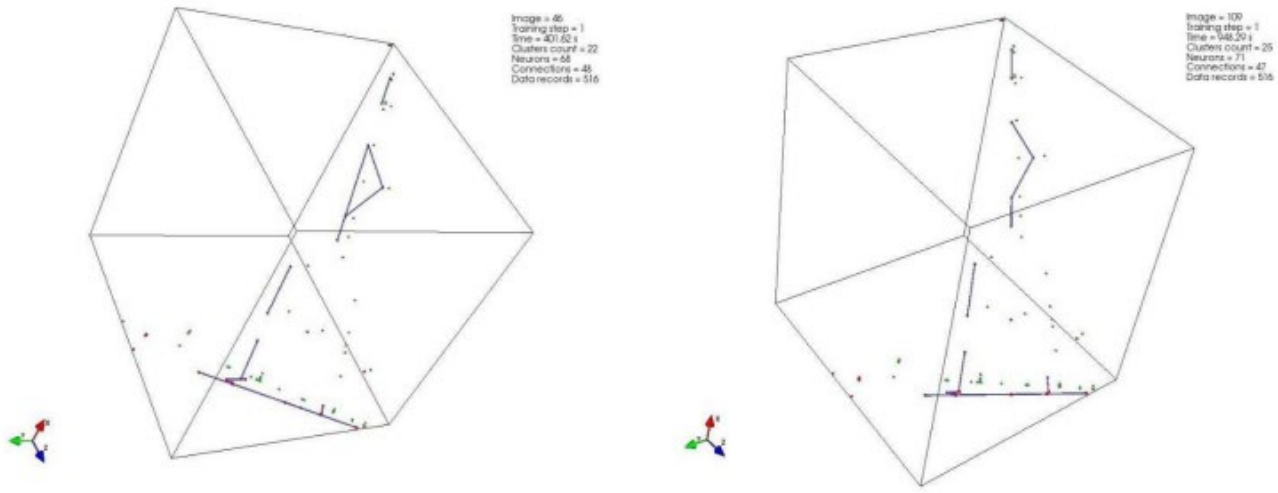

Рисунок 3.9 – Карти, що самоорганізуються після 46000 кроків і після 109000 кроків навчання

Самоорганізовані карти були сформовані із 200 тисяч рядків трафіку мережі, котрі були надіслані в нейромережевий модуль. Основна структура карти, що самоорганізується, згідно із графічним представленням сформувалася уже на 46 тисячному кроці. І різниці після 46 тисяч і після 109 тисяч повторів не помічається.

Подання даних для тестування здійснюється через модуль препроцесора системи Snort. Після цього в порядку нормальної роботи здійснюється зіставлення висновків отриманих нейромережею величин із тими, котрі отримані в результаті роботи модуля детектування атак у системі Snort.

У таблиці 3.3 наведено число детектованих аномалій нейромережевого модуля і число віднайдених аномалій модулем детектування атак системи

«Snort» після формування самоорганізаційної карти. У цьому випадку звірка відбувалася на тестовому наборі, котрий містить 736 840 записів.

Таблиця 3.3 – Підсумки роботи на реальному НД

Опис	IDS Snort	Нейромережевий модуль
Кількість оброблених записів	736 840	736 840
Кількість виявлених аномалій	19	14

На рисунку 3.10 продемонстровано звіт роботи нейромережевого модуля на тестових даних. Згідно звіту модуль виявив 14 відхилень, що не відповідають його навченій моделі.

```
Застосування детектора до тестового набору даних із використанням навченої моделі

Читання тестового набору . . .
Кількість записів: 736840
Аномальні записи    ╞ 0, Нормальні записи = 748,    Час   виявлення  = 1.13 с, Час на запис = 0.00112664413452  с
Аномальні записи    = 0, Нормальні записи = 1522,   Час   виявлення  = 2.26 с, Час на запис = 0.00112577366829  с
Аномальні записи    = 0, Нормальні записи = 2279,   Час   виявлення  = 3.38 с, Час на запис = 0.00112768054008  с
Аномальні записи    = 0, Нормальні записи = 3776,   Час   виявлення  = 4,51 с, Час на запис = 0.00112878522873  с
Аномальні записи    = 0, Нормальні записи = 5294,   Час   виявлення  = 5,64 с, Час на запис = 0.00112647589048  с
Аномальні записи    = 0, Нормальні записи = 6045,   Час   виявлення  = 6.76 с. Час на запис = 0.00112660801411  с
Аномальні записи    = ®, Нормальні записи = 6796,   Час   виявлення  = 7.89 с, Час на запис = 0.00112878522873  с
Аномальні записи    = 1, Нормальні записи = 8316,   Час   виявлення  = 9.01 с, Час на запис = 0.00112577366829  с


Аномальні записи    = 14, Нормальні записи = 735298, Час  виявлення = 188.87 с, Час на  запис = 0,00112742733955 с
Аномальні записи    = 14, Нормальні записи = 736078, Час  виявлення = 189.45 с, Час на  запис = 0.00112647589048 с
Аномальні записи    = 14, Нормальні записи = 736826, Час   виявлення = 190.91 с, Час на  запис = 0.00112592681971 с

Виявлено аномалії (кількість = 14) [нормальні записи = 736826, час виявлення = 193.45 с, час на запис = 0.00112551166035 с]
```

Рисунок 3.10 – Звіт роботи нейромережевого модуля на тестових даних

На рисунку 3.7, 3.8 подано звіти системи «Snort», у звіті вказані всі типи атак їх кількісне відношення до загального числа, загальна кількість атак виявлених системою «Snort» склала 19.

3.5 Висновки до третього розділу

В результаті виконаної симуляції нейромережевого модуля детектування аномалій з’явилася наявність детектування вторгнення у мережу, перевіряння роботоздатності модуля на дійсному трафіку проводиться шляхом співставлення одержаних результатів нейромережевого модуля із стандартним модулем детектування загроз системи Snort.

За умови, при якій результати, одержані за допомогою нейромережевого модуля, співпадають із тими, котрі отримані з системи Snort, тоді можна стверджувати, що навчання було успішним, та функціонування системи є коректним. Якщо ж ці результати відрізняються, в такому випадку можна констатувати, що нейромережевий модуль детектував аномалію, котра наразі невідома для системі Snort і цей НД потрібно дослідити додатково, або навчання нейромережевого модуля було некоректним і є потреба у перенавчанні нейромережі із використанням іншому НД для тестування.

Одержані результати свідчать про достатню перспективність концепції для покращення якості функціонування IDS / IPS для роботи МДП.

4 БЕЗПЕКА ЖИТТЄДІЯЛЬНОСТІ, ОСНОВИ ОХОРОНИ ПРАЦІ

4.1 Стихійні лиха та їх класифікація

Стихійні дії сил природи, поки що не повною мірою підвладні людині та щорічно завдають державі і населенню величезних збитків. Стихійні лиха - це такі явища природи, що викликають екстремальні ситуації, порушують нормальну життєдіяльність населення, роботу безлічі об'єктів. Стихійні лиха є трагедією для будь-якої держави. Через стихійні лиха страждає економіка країни, бо при цьому руйнуються виробничі підприємства, знищуються матеріальні цінності, гинуть люди.

Стихійні лиха - небезпечні природні явища, як правило раптового походження, хоча іноді і прогнозовані за допомогою метеорогогії, але на інтенсивність яких люди впливати не можуть. Їх можна класифікувати: за швидкістю переміщення - землетруси, зсуви, цунамі, снігопади, ожеледі - швидкі; підвищення рівня води в ріках через інтенсивні опади або танення снігу, льоду (повіні), звільнення внутрішньої енергії Землі, виверження вулканів - повільні. Часто виникають потужні, високошвидкісні потоки повітря через швидкий перепад значень атмосферного тиску (урагани, смерчі і т.п.). Стихійні лиха речовинного характеру можуть ініціювати виникнення різноманітних полів, які негативно впливають на здоров'я, самопочуття людини [18].

Стихійні явища часто виникають в комплексі, що значно посилює їх негативний вплив. Небезпечні природні явища визначаються трьома основними групами процесів - ендогенні, екзогенні та гідрометеорологічні.

Стихійні лиха, які характерні для України, за структурою можна поділити на прості, що включають один елемент - наприклад, сильний вітер, зсув або землетрус та складні. Вони складаються з декількох процесів однієї групи або кількох груп. Найбільші збитки спричиняють повені - 40%, на другому місці - циклони (20%), на третьому - посухи та землетруси (15%). Деякі стихійні лиха (пожежі, обвали, зсуви і навіть землетруси) можуть виникати в результаті дій

самих людей, тобто мають антропогенне походження, але наслідки їх завжди є діями сил природи. Для кожного стихійного лиха характерна наявність властивих йому вражаючих чинників, що несприятливо впливають на стан здоров'я, життя людини [18].

Причинами стихійних лих можуть бути:

– швидке переміщення речовини (землетрусу, зсуви);

– вивільнення внутриземної енергії (вулканічна діяльність, землетруси);

– підвищення рівня вод річок, ставків і морів (повені, цунамі);

– вплив надзвичайно сильного вітру (урагани, торнадо, циклони).

Важливо своєчасно провести роботи, спрямовані на локалізацію природного лиха, щоб зменшити зони руйнувань, звести до мінімуму кількість загиблих та постраждалих.

В Україні найчастіше спостерігаються такі надзвичайні ситуації природного характеру:

– небезпечні геологічні явища (зсуви, обвали, осипки, просадки земної поверхні);

– небезпечні метеорологічні явища (зливи, урагани, сильні снігопади, сильний град, ожеледь);

– небезпечні гідрологічні явища (повені, паводки);

– природні пожежі лісових та торф’яних масивів;

– масові інфекції та хвороби людей, тварин, рослин.

В останні роки кількість стихійних лих в Україні та в світі в цілому значно збільшилася. Найчастіше в Україні виникають такі природні катастрофи як землетруси, повені, посухи (на Півдні України), лісові пожежі в літню пору року, снігові замети, зсуви поверхні.

Є серйозні підстави вважати, що масштабність впливу лиха й катастроф на соціальні, економічні, політичні та інших процесів сучасного нашого суспільства та їх драматизм вже перевищили такий рівень, який дозволяв ставитися до них як до локальних збоїв у розміреному функціонуванні державних та громадських структур [19].

Отже, перед людиною та громадськістю в XXI в. вимальовується нова мета - глобальна безпека. Досягти цього можна, в першу чергу, за допомогою зміни світогляду людини, а також покращення системи профілактичних заходів у боротьбі зі стихійними лихами, а саме: вдосконалення рятувальних служб та рятувальної техніки, проведення попереджувальних заходів та пропагандистської роботи з громадянами щодо правил поведінки та дій під час стихійних лих. Це допоможе в майбутньому зменшити кількість загиблих та постраждалих від природних катастроф, а також зменшить матеріальні збитки, що були завдані стихійним лихом.

Природні лиха з часом нікуди не зникнуть. Будуть виникати землетруси в геологічно активних районах, будуть виникати повені, а штормові припливи стануть, раз у раз затопляти морські узбережжя, не обійдеться і пожеж. Людина безсила запобігти природним процесам, але тільки в наших силах зменшити кількість жертв і матеріальних втрат.

4.2 Соціальне значення охорони праці

Соціальне значення охорони праці полягає в сприянні росту ефективності суспільного виробництва шляхом безперервного вдосконалення і поліпшення умов праці, підвищення їх безпеки, зниження виробничого травматизму і профзахворювань [19]. Соціальне значення охорони праці проявляється в зростанні продуктивності праці, збереженні трудових ресурсів і збільшенні сукупного національного продукту.

Охорона праці полягає в сприянні росту ефективності виробництва, яке досягається шляхом безперервного вдосконалення і поліпшення умов праці, підвищення їх безпеки, зниження виробничого травматизму і профзахворювань.

Зростання продуктивності праці відбувається в результаті збільшення фонду робочого часу завдяки скороченню внутрішньо-змінних простоїв шляхом ліквідації мікротравм або зниження їх кількості, а також завдяки запобіганню передчасного стомлення шляхом раціоналізації і покращення умов праці та

введенню оптимальних режимів праці і відпочинку та інших заходів, які сприяють підвищенню ефективності використання робочого часу.

Важливим питанням є зростання продуктивності праці, яка відбувається в результаті збільшення фонду робочого часу завдяки скороченню внутрішньозмінних простоїв шляхом ліквідації мікротравм або зниження їх кількості, а також завдяки запобіганню передчасного стомлення шляхом раціоналізації і покращення умов праці та введенню оптимальних режимів праці і відпочинку та інших заходів, які сприяють підвищенню ефективності використання робочого часу [20].

Особливої уваги заслуговує те, що збереження трудових ресурсів і підвищення професійної активності працюючих відбувається завдяки покращенню стану здоров’я і подовженню середньої тривалості життя шляхом покращення умов праці, що супроводжується високою трудовою активністю і підвищенням виробничого стажу. Підвищується професійний рівень також завдяки зростанню кваліфікації і майстерності. Відповідно і збільшення сукупного національного продукту відбувається завдяки покращенню вищеперелічених показників та їх складових компонентів [19]. Збереження трудових ресурсів і підвищення професійної активності працюючих відбувається завдяки покращенню стану здоров'я і подовженню середньої тривалості життя шляхом покращення умов праці, що супроводжується високою трудовою активністю і підвищенням виробничого стажу. Підвищується професійний рівень також завдяки зростанню кваліфікації і майстерності. Збільшення сукупного національного продукту відбувається завдяки покращенню вищеперелічених показників та їх складових компонентів. Крім того, соціальне значення охорони праці проявляється в зростанні продуктивності праці, збереженні трудових ресурсів.

Комплекс заходів з поліпшення умов праці може забезпечити приріст продуктивності праці на 15-20%. Так, нормалізація освітлення робочих місць збільшує продуктивність на 6-13% та скорочує брак на 25%. Раціональна організація робочого місця підвищує продуктивність праці на 21%, раціональне

фарбування робочих приміщень – на 25% [20]. Збільшення ефективного фонду робочого часу може бути досягнуто за рахунок скорочення тимчасової непрацездатності працівників внаслідок хвороб та виробничого травматизму.

ВИСНОВКИ

У проведеній роботі проаналізовано IDS/IPS, загальну методику опрацювання подій в мережі досліджено на основі системи Snort. Головною ідеєю проведеного аналізу була потреба у з'ясуванні основних проблем таких систем та ймовірності їх розв’язання при допомозі використання технологій нейромереж.

Пропонується до використання адаптивний модуль, котрий функціонує паралельно (спільно) із нормами системи Snort. Така спільна робота обумовлена тим, що система Snort детектує лише відомий шкідливий трафік, котрий внесений до системної бази знань. На додачу нейромережевий ж модуль володіє здатністю детектувати невідомий чи змінені випадки шкідливого трафіку, котрий, зі своєї сторони, значно вплине на точність детектування (покращить її).

Дослідивши наявні нейромережеві технології, було прийнято рішення детектувати аномалії при допомозі методів кластеризації, так званих «карт активності мережі». У межах сучасних нейромережевих технологій задача кластеризації розв’язується з використанням SOM. Був описаний алгоритм кластеризації, карти Кохонена, що самоорганізуються, і запропонований метод роботи нейромережевого модуля, котрий функціонує відповідно до цього алгоритму.

Результати виконаних експериментальних досліджень із застосуванням запропонованої моделі нейромережевого модуля підтвердили можливість детектування спроб вторгнення у мережу. З метою перевіряння роботоздатності модуля на реальному трафіку здійснюється порівняння результатів, одержаних за допомогою нейромережевого модуля, із результатами, котрі отримані модулем детектування атак системи Snort.

ПЕРЕЛІК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Микитишин А. Г., Митник М. М., Стухляк П. Д., Пасічник В. В. Комп'ютерні мережі. Книга 1 [навчальний посібник]. Львів : «Магнолія 2006», 2013. 256 с.

2. Nataliya Zagorodna, Iryna Kramar (2020). Economics, Business and Security: Review of Relations. Business Risk in Changing Dynamics of Global Village BRCDGV-2020: Monograph / Edited by Pradeep Kumar, Mahammad Sharif. India, Patna: Novelty & Co., Ashok Rajpath,. 446 p., pp.25-39.

3. 10 найкращих систем виявлення вторгнень. [Електронний ресурс. - Режим доступу: https://uk.myservername.com/top-10-best-intrusion-detection-systems (дата звернення: 30.04.2024).

4. Бекер, І., Тимощук, В., Маслянка, Т., & Тимощук, Д. (2023). МЕТОДИКА ЗАХИСТУ ВІД ПОВІЛЬНИХ ТА ШВИДКИХ BRUTE-FORCE АТАК НА IMAP СЕРВЕР. Матеріали конференцій МНЛ, (17 листопада 2023 р., м. Львів), 275-276.

5. What is Snort? [Електронний ресурс]. - Режим доступу: https://www.fortinet.com/resources/cyberglossary/snort (дата звернення: 30.04.2024).

6. Kharchenko, O., Raichev, I., Bodnarchuk, I., & Zagorodna, N. (2018, February). Optimization of software architecture selection for the system under design and reengineering. In 2018 14th International Conference on Advanced Trends in Radioelecrtronics, Telecommunications and Computer Engineering (TCSET) (pp. 1245-1248). IEEE.

7. SNORT IDS. [Електронний ресурс]. – Режим доступу: http://www.snort.org/ (дата звернення: 06.05.2024).

8. Тимощук, В., & Стебельський, М. (2023). Шифрування даних в операційних системах. Матеріали VI Міжнародної студентської науково-технічної конференції „Природничі та гуманітарні науки. Актуальні питання “, 183-184.

9. Skorenkyy, Y., Zolotyy, R., Fedak, S., Kramar, O., & Kozak, R. (2023, June). Digital Twin Implementation in Transition of Smart Manufacturing to Industry 5.0 Practices. In CITI (pp. 12-23).

10. Fryz, M., Mlynko, B., Mul, O., & Zagorodna, N. (2010). Conditional Linear Periodical Random Process as a Mathematical Model of Photoplethysmographic Signal. Rigas Tehniskas Universitates Zinatniskie Raksti, 45, 82.

11. Tymoshchuk, V., Karnaukhov, A., & Tymoshchuk, D. (2024). USING VPN TECHNOLOGY TO CREATE SECURE CORPORATE NETWORKS. Collection of scientific papers «ΛΌГОΣ», (June 21, 2024; Seoul, South Korea), 166-170. https://doi.org/10.36074/logos-21.06.2024.034.

12. Басюк Т.М. Машинне навчання: Навчальний посібник. Львів: Видавництво «Новий Світ - 2000», 2021. 315 с.

13. Kharchenko, A., Halay, I., Zagorodna, N., & Bodnarchuk, I. (2015, September). Trade-off optimal decision of the problem of software system architecture choice. In 2015 Xth International Scientific and Technical Conference" Computer Sciences and Information Technologies"(CSIT) (pp. 198-205). IEEE.

14. A Python Toolbox for Scalable Outlier Detection (Anomaly Detection) [Електроний ресурс]. – Режим доступу: https://GitHubyzhao062/pyod (дата звернення: 26.05.2024).

15. Тимощук, В., Долінський, А., & Тимощук, Д. (2024). СИСТЕМА ЗМЕНШЕННЯ ВПЛИВУ DOS-АТАК НА ОСНОВІ MIKROTIK. Матеріали конференцій МЦНД, (17.05. 2024; Ужгород, Україна), 198-200. https://doi.org/10.62731/mcnd-17.05.2024.008.

16. Ревнюк, О. А., Загородна, Н. В., Козак, Р. О., Карпінський, М. П., & Флуд, Л. О. (2024). The improvement of web-application SDL process to prevent Insecure Design vulnerabilities. Прикладні аспекти інформаційних технологій, 7(2), 162-174.

17. Tymoshchuk, V., Dolinskyi, A., & Tymoshchuk, D. (2024). MESSENGER BOTS IN SMART HOMES: COGNITIVE AGENTS AT THE FOREFRONT OF THE INTEGRATION OF CYBER-PHYSICAL SYSTEMS AND THE INTERNET OF THINGS. Матеріали конференцій МЦНД, (07.06. 2024; Луцьк, Україна), 266-267. https://doi.org/10.62731/mcnd-07.06.2024.004

18. Толок А.О. Крюковська О.А. Безпека життєдіяльності: Навч. посібник. – 2011. – 215 с.

19. Безпека в надзвичайних ситуаціях. Методичний посібник для здобувачів освітнього ступеня «магістр» всіх спеціальностей денної та заочної (дистанційної) форм навчання / укл.: Стручок В. С. Тернопіль: ФОП Паляниця В. А., 2022. 156 с.

20. Основи охорони праці: Підручник.; 3-те видання / За ред. Ткачука К. Н. – К.: Основа, 2011. – 480 с.