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: Створення системи виявлення вторгнень на основі архітектури штучної імунної системи

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

(підпис) Черник О.А. (прізвище та ініціали)

Керівник (підпис) Баран І.О. (прізвище та ініціали)

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

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

Рецензент (підпис) Литвиненко Я.В. (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)
«__» _____ 2021 р.

ЗАВДАННЯ
НА КВАЛІФІКАЦІЙНУ РОБОТУ

на здобуття освітнього ступеня Магістр
(назва освітнього ступеня)

за спеціальністю 125 Кібербезпека
(шифр і назва спеціальності)

Студенту Чернику Олегу Андрійовичу
(прізвище, ім'я, по батькові)

1. Тема роботи Створення системи виявлення вторгнень на основі архітектури штучної імунної системи

Керівник роботи Баран Ігор Олегович, к.т.н., доц., декан ФІС
(прізвище, ім'я, по батькові, науковий ступінь, вчене звання)

Затверджені наказом ректора від « 16 » 11 2023 року № 4/7-1061

2. Термін подання студентом завершеної роботи 26.12.2023р.

3. Вихідні дані до роботи наукові літературні джерела

4. Зміст роботи (перелік питань, які потрібно розробити)
1. Аналіз предметної області.
2. Розробка методів виявлення аномальних запитів.
3. Аналіз ефективності системи.
4. Охорона праці та безпека в надзвичайних ситуаціях

5. Перелік графічного матеріалу (з точним зазначенням обов'язкових креслень, слайдів)
1. Тема роботи. 2. Актуальність. 3. Мета, задачі, об'єкт, предмет дослідження. 4. Наукова новизна, практичне значення роботи. 5. Типи мережевих атак, Послідовність дій зловмисника. 6. Основні компоненти СВВ. 7. Архітектура пропонованого підходу аналізу трафіку. 8. Імунні системи розпізнавання. 9. Реалізація СВВ. 10. Вікно клієнтського додатку. 11. Обчислення результату комбінації методів 12. Вікна налаштування та тестування системи. 13. Аналіз ефективності ФІС 14. Аналіз ефективності ІСР. 15. Аналіз ефективності комбінації методів 16. Основні результати дослідження

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., проректор з адміністративно-господарської роботи та будівництва		

7. Дата видачі завдання 2023 р.

КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назва етапів роботи	Термін виконання етапів роботи	Примітка
1.	Ознайомлення з завданням до кваліфікаційної роботи	16.11 – 17.11	*Виконано*
2.	Підбір джерел про системи виявлення вторгнень	18.11 – 26.11	*Виконано*
3.	Опрацювання джерел про дослідження системи		
	виявлення вторгнень на основі штучних імунних систем	27.11 – 30.11	*Виконано*
4.	Виконання дослідження щодо розробки системи		
	виявлення вторгнень на основі штучних імунних систем	01.12 – 06.12	*Виконано*
5	Розробка алгоритмів функціонування системи	07.12 – 10.12	
6.	Оформлення розділу «Аналіз предметної області»	11.12 – 13.12	*Виконано*
7.	Оформлення розділу «Розробка методів виявлення		
	аномальних запитів»	14.12 – 15.12	*Виконано*
8.	Оформлення розділу «Аналіз ефективності системи»	16.12 – 18.12	*Виконано*
9.	Виконання завдання до підрозділу «Охорона праці та		
	безпека в надзвичайних ситуаціях»	06.12 – 16.12	*Виконано*
10.	Оформлення кваліфікаційної роботи	14.12 – 19.12	*Виконано*
11.	Нормоконтроль	18.12 – 20.12	*Виконано*
12.	Перевірка на плагіат	16.12 – 19.12	*Виконано*
13.	Попередній захист кваліфікаційної роботи	17.12 – 20.12	*Виконано*
14.	Захист кваліфікаційної роботи	27.12	*Виконано*

Студент _____ Черник О.А.
(підпис) (прізвище та ініціали)

Керівник роботи _____ Баран І.О.
(підпис) (прізвище та ініціали)

АНОТАЦІЯ

Створення системи виявлення вторгнень на основі архітектури штучної імунної системи // Черник Олег Андрійович // Тернопільський національний технічний університет імені Івана Пулюя, факультет комп'ютерно-інформаційних систем та програмної інженерії, кафедра кібербезпеки, група СБм-61 //Тернопіль, 2023// с. – 78, рис. – 25, табл. – 14 , слайд. – 16, бібліогр. –35.

Ключові слова: АЛГОРИТМ КЛОНАЛЬНОЇ СЕЛЕКЦІЇ, АЛГОРИТМ НЕГАТИВНОГО ВІДБОРУ, ІМУННІ СИСТЕМИ РОЗПІЗНАВАННЯ, СИСТЕМА ВИЯВЛЕННЯ ВТОРГНЕНЬ, ФОРМАЛЬНА ІМУННА СИСТЕМА, ШТУЧНА ІМУННА СИСТЕМА

Здійснено критичний огляд мережевих загроз, описані різні мережеві атаки. Виконано аналіз сучасних систем виявлення вторгнень, приведено їх класифікацію за місцем збору інформації, швидкістю реагування та механізмом виявлення. Наведено типову архітектуру системи виявлення вторгнень. Також описано штучні імунні системи, особливості використання алгоритмів клональної селекції і негативного відбору.

Пропонується для детектування аномальних запитів застосувати формальні імунні системи та імунні системи розпізнавання. Два ці підходи інтегруються в складову аналізу системи виявлення вторгнень. При використанні імунокомп'ютингу сформовано алгоритм побудови формальної імунної системи. В основі розробленої системи виявлення вторгнень лежить клієнт-серверна архітектура. Навчання системи виявлення вторгнень та генерація трафіку проходили із використанням набору даних CICIDS 2017.

Проаналізовано ефективність функціонування пропонованої системи виявлення вторгнень. Отримані результати експериментів свідчать, що штучна імунна система може успішно бути застосована для виявлення різних мережевих вторгнень.

ANNOTATION

Creation of an intrusion detection system based on the architecture of an artificial immune system // Chernyk Oleh // Ternopil Ivan Pul'uj National Technical University, Faculty of Computer Information Systems and Software Engineering, Department of Cyber Security // Ternopil, 2023 // p. - 78, Fig. - 25, Table - 14, Slides - 16, References - 35.

Keywords: CLONAL SELECTION ALGORITHM, NEGATIVE SELECTION ALGORITHM, IMMUNE RECOGNITION SYSTEMS, INTRUSION DETECTION SYSTEM, FORMAL IMMUNE SYSTEM, ARTIFICIAL IMMUNE SYSTEM

A critical review of network threats is carried out, various network attacks are described. The analysis of modern intrusion detection systems was performed, and their classification based on the location of information collection, response speed, and detection mechanism was provided. A typical architecture of an intrusion detection system is given. Artificial immune systems, features of using clonal selection algorithms and negative selection are also described.

It is proposed to use formal immune systems and immune recognition systems to detect abnormal requests. These two approaches are integrated into the analysis component of the intrusion detection system. When using immunocomputing, an algorithm for building a formal immune system was formed. The basis of the developed intrusion detection system is the client-server architecture. The intrusion detection system was trained and traffic generated using the CICIDS 2017 dataset.

The effectiveness of the proposed intrusion detection system was analyzed. The experimental results show that the artificial immune system can be successfully applied to detect various network intrusions.

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, ОДИНИЦЬ СКОРОЧЕНЬ І ТЕРМІНІВ

АЗ – аномальний запит.

АКС – алгоритм клональної селекції.

АНВ - алгоритм негативного відбору.

БД – база даних.

ВА – виявлення аномалій.

ІСР – імунні системи розпізнавання.

ІТ – інформаційні технології.

МН – машинне навчання.

НД – набір даних.

НСД – несанкціонований доступ.

ОС – операційна система.

ПЗ – програмне забезпечення.

СВВ (IDS - Intrusion Detection System) – система виявлення вторгнень.

СРМ – сингулярне розкладанням матриці.

ФІС – формальна імунна система.

ШІС (штучна імунна система) – адаптивна обчислювальна система, що використовує моделі, принципи, механізми та функції, описані в теоретичній імунології, які застосовуються для вирішення прикладних завдань.

ЗМІСТ

ДОДАТКИ

Додаток А. Тези конференції

Додаток Б. Фрагмент програмного коду

ВСТУП

Актуальність теми. У світі ІТ відіграють значну роль у житті людини. З кожним роком у світі збільшується кількість користувачів комп'ютерів, більшість яких з'єднуються в комп'ютерні мережі. За даними [1] у світі понад 5,1 мільярди людей користуються інтернетом. Сучасний розвиток комп'ютерних мереж допомагає швидко обмінюватися інформацією, однак повсюдна інтеграція мереж збільшує кількість можливих загроз.

Мережеві атаки є будь-якими діями та засобами, що використовуються для зловмисного порушення роботи мережі, в результаті яких під загрозою знаходяться дані та активи організацій та установ, персональні дані фізичних осіб та інша важлива інформація. Для контролю безпеки роботи в мережі використовуються комбінації різних програмних рішень, як то антивіруси, міжмережеві екрани, алгоритми шифрування, приватні віртуальні мережі (VPN), системи аутентифікації. Додатковим інструментом для виявлення спроб НСД до системи є СВВ.

СВВ з'явилися понад тридцять років тому, коли зі збільшенням кількості користувачів у мережах виникла потреба контролювати їхні дії для забезпечення безпечної та надійної роботи мережі. З того часу СВВ допомагають виявляти аномальні дії в мережі та попереджають адміністраторів систем про можливі загрози.

ШІС - це адаптивні системи, засновані на теорії імунних систем хребетних [2]. Вони відносяться до методів МН, що належать галузі штучного інтелекту. Основна ідея застосування даного методу полягає в тому, що імунна система забезпечує потужний захист роботи організму від зовнішніх шкідників та різних патогенів, що може бути прототипом захисту комп'ютерних систем від кібератак [3]. ШІС є відносно новим напрямом наукових досліджень, мають великий потенціал застосування в галузі захисту інформації в мережах, тому цей дослідний напрямок є актуальним.

Мета дослідження: дослідити та створити СВВ з урахуванням архітектури ШІС.

В роботі поставлено та розв’язано наступні задачі:

– вивчити існуючі мережеві атаки та визначити набір атак, які будуть розглядатися під час створення системи;

– дослідити методи ВА, що застосовуються у СВВ;

– розробити алгоритми виявлення АЗ з урахуванням архітектури ШІС;

– реалізувати компоненти СВВ;

– оцінити ефективність отриманої СВВ.

Об’єкт дослідження: СВВ.

Предмет дослідження: проектування СВВ із застосуванням ШІС.

Методи дослідження: наукові праці закордонних та вітчизняних учених за тематикою дослідження, фундаментальні положення кібербезпеки та ШІС; методи - аналітичний, порівняльний, системного аналізу, проектування.

Наукова новизна отриманих результатів:

– сформовано алгоритм для послідовного опису створення ФІС;

– запропоновано аналіз запитів на предмет їх аномальності виконувати із застосуванням комбінації методів ФІС та ІСР, що дозволило підвищити якість детектування атак до 96,47%.;

– для тестування розробленої СВВ було використано різні експериментальні методики, спрямовані на оцінку різних характеристик функціонування системи.

Практичне значення одержаних результатів. Результати проведеного дослідження можуть бути ефективно застосовані для детектування мережевих вторгнень різного роду. Із використання розробки адміністратори комп’ютерних мереж підприємств, установ та організацій зможуть успішно виявляти аномальні дії в мережі та отримувати повідомлення про ймовірні загрози.

Апробація. Результати дослідження апробовано на XI науково-технічній конференції «Інформаційні моделі, системи та технології» у вигляді опублікованих тез [4].

Структура роботи. Робота складається з пояснювальної записки та графічної частини. Пояснювальна записка складається з вступу, 4 розділів,

висновків, списку використаної літератури та додатків. Обсяг роботи: пояснювальна записка – 78 арк. формату А4, графічна частина – 16 слайдів.

1 АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ

1.1 Огляд мережевих загроз

Мережева атака - це атака, спрямовану на мережну інфраструктуру віддаленого об'єкта, основна мета якої є отримання інформації чи збій у роботі мережі [5].

Існує два основні види мережевих атак:

– пасивні, створені задля отримання чи перегляд даних у комп'ютерних мережах без зміни чи втручання у роботу системи;

– активні, створені задля зміна даних чи порушення нормальної роботи системи.

Під час роботи мереж, зокрема більшості організацій, все web- запити, передачі файлів і автентифікація користувачів обробляються мережними пристроями. Тому основним джерелом витоку інформації та проникнення в систему є неконтрольовані мережеві пристрої. Також збільшенню кількості мережевих атак і їх складності сприяють доступність інструментів атаки і хакерські знання, що розвиваються. Наприклад, сучасні DDoS-атаки можуть бути виконані на прикладному рівні, на відміну від попередніх років, коли атаки могли проводитися лише на мережевому та транспортному рівнях.

Основною метою зловмисників є організації та корпоративні мережі. Причиною здійснення атаки може бути:

— бажання отримання матеріальної вигоди;

— промислове шпигунство;

— терористичні ідеї;

— політичні розбіжності чи замовлення;

— пошук слави та визнання хакерської спільноти;

— виявлення існуючих уразливостей системи;

— заподіяння шкоди даним, фінансам та репутації організації.

Для досягнення своїх цілей мережні хакери виконують різні шкідливі дії, серед яких незаконне використання облікових записів користувачів, запуск коду

пошкодження систем або пошкодження даних, крадіжка обладнання або ПЗ.

Існують різні типи мережевих атак, варто навести найбільш популярні [5].

DoS та DDoS- атаки. DoS -атака - це сукупність дій, спрямована на припинення роботи Інтернет-ресурсу. У DDoS- атаках використовується безліч комп'ютерів, які раніше були заражені вірусом і тепер можуть виконувати шкідливі запити мережі за сигналом хакера, без відома і згоди користувачів (рис.1.1). Метою зловмисників можуть бути будь-які інтернет-ресурси: онлайн-магазини, ігрові сервери, сайти компаній та державні сайти. DoS -атаки в основному здійснюються за двома сценаріями: відправка на віддалений комп'ютер спеціальних мережевих пакетів, які порушують роботу системи або відправлення великої кількості мережевих пакетів, які заповнюють всі ресурси системи, що атакується.

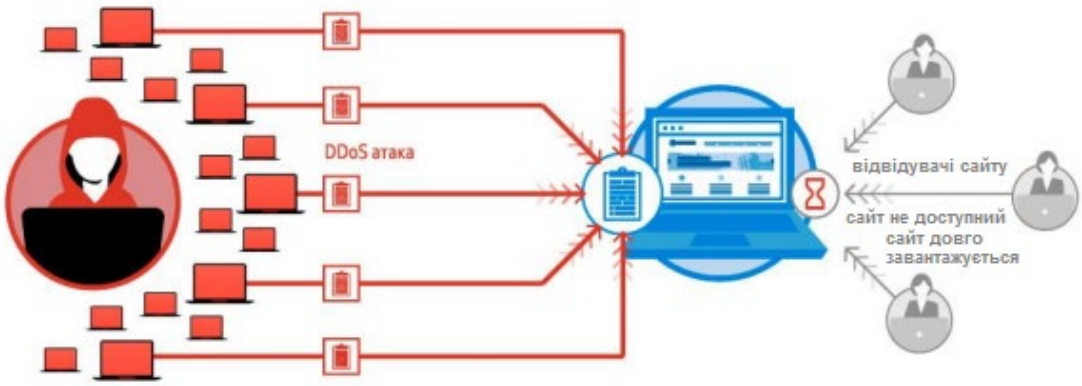

Рисунок 1.1 – Схема проведення DDoS- атаки

Спуфінг (IP Spoofing). Кожен пристрій, підключений до інтернету, відправляє IP -пакети в мережу, в яких зберігається ip -адреса відправника та дані прикладного рівня. Зловмисник, маючи можливість керувати мережевим пристроєм, може помістити довільну ip -адресу в даних відправника. Таким чином, пакет у відповідь надійде на вказану ip- адресу, а справжня адреса атакуючого виявиться прихованою (рис. 1.2). Даний підхід застосовується з метою обману систем безпеки та у складі деяких DoS -атак (SYN -флуд, DNS посилення).

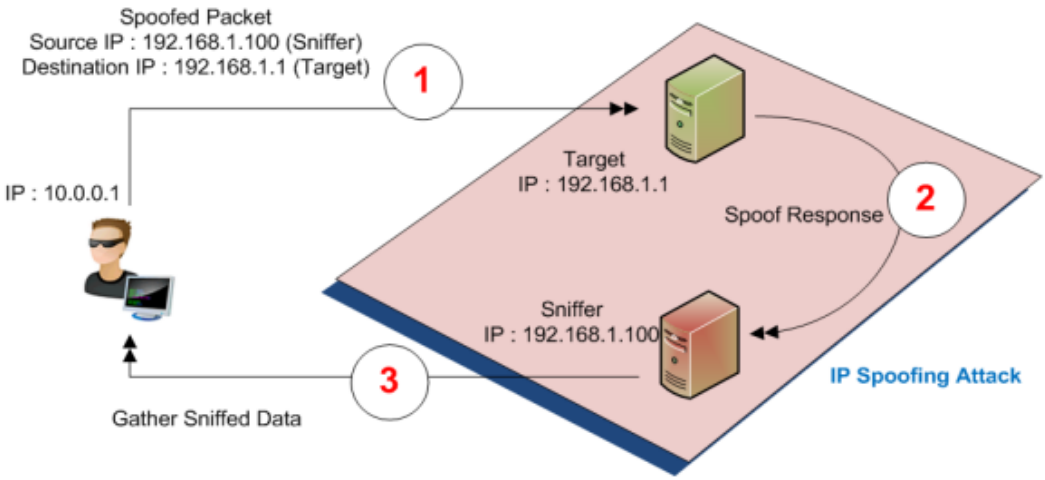

Рисунок 1.2 – Схема проведення IP-спуфінгу

Мережеві атаки вторгнення. Цей вид атак використовується за необхідності отримання даних з віддаленого комп'ютера. Для цього зловмисник вживає дій для отримання управління над ОС жертви. При успішному проведенні атаки жертва може зазнати великих втрат, оскільки атакована система може повністю управляти мережевим хакером. Атаки вторгнення найчастіше використовують уразливості мережевих сервісів ОС.

Сніфер пакетів. Є програмою, призначеною для перехоплення пакетів даних, що проходять через мережу. Даний вид прикладних програм працює на рівні Ethemet з мережевими інтерфейсними картами (NIC) і використовується для легального аналізу трафіку, а також діагностики несправності в мережі. Але можливості цих програм можуть бути використані зловмисниками для отримання конфіденційної інформації. Особливою небезпекою є передача пакетів, що містять імена користувачів та паролі.

Man-in-the-middle attack. Цей метод заснований на вразливості в стеку протоколів TCP/IP та способі побудови заголовків. Прикладом атаки є встановлення спілкування між користувачами не безпосередньо, а через «третю особу», яка перехоплює всю інформацію і потім перенаправляє одержувачу. Захистом від цього типу атак є автентифікація користувачів, контроль цілісності повідомлень та використання криптографічних методів.

Трояни. Це програми, які маскуються під легальні програми з метою

потрапити до системи жертви. Троянські програми складно відрізнити від легальних програм, оскільки вони виглядають і працюють так само, як і скомпрометована програма чи файл. Більшість програм дистанційного керування належить до цього типу. Троянські програми мають великий набір можливостей завдання шкоди атакованим системам або за допомогою них.

Соціальна інженерія. Це використання людського фактору шляхом переконання або обману для отримання доступу до системи. Зловмисник намагається будь-якими способами впливати на людину, яка має необхідний доступ або дані, для їх незаконного використання у своїх цілях.

В основі здійснення атаки зловмисником зазвичай лежить наступна послідовність дій (рис. 1.3):

Рисунок 1.3 – Послідовність дій зловмисника

– створення картки мережі. Зламування мережі починається зі збору різної інформації, метою якого є створення картки мережі. Для цього необхідно дізнатися тип ОС, встановлені програми, діапазон адрес і доступні порти. Джерелами інформації можуть бути загальнодоступні ресурси, сайти компанії та спеціальні інструменти, такі як Nslookup, Nmap, Tracert;

– сканування портів. Застосовується для збору інформації про використовувані мережеві служби та можливі відкриті порти системи. Відкриті порти системи можна дізнатися за допомогою загальнодоступних інструментів,

таких як Nmap та XSpider. Знаючи відкриті порти, зловмисник може спробувати проникнути в систему, підбираючи логіни або паролі або за допомогою програми-експлойта;

– отримання доступу. Зловмисник шукає вразливості системи та застосовує програмні засоби для отримання доступу. Зазвичай для цього використовують троянські програми та програми підбору пароля. Небезпека для системи також є наявність мережевих служб, які працюють, але не використовуються, облікові записи користувачів, які не мають паролів та активні гостьові облікові записи;

– розширення привілеїв. Після отримання доступу до системи зазвичай зловмисник має обмежені права. Для реалізації поставлених цілей зловмиснику необхідно підвищити свій рівень доступу та отримати права адміністратора. До методів розширення прав відносять пошук паролів у реєстрі, програми злому паролів, пошук адміністративних документів з інформацією;

– встановлення бекдорів та видалення слідів діяльності. Після отримання доступу до системи мережні хакери намагаються залишити програму-бекдор на системі для отримання легкого доступу до системи. Для закріплення вдалого результату необхідно видалити всі докази діяльності хакера.

1.2 Системи виявлення вторгнень

Зі зростанням вимог безпеки для її забезпечення під час роботи в мережі застосовуються різноманітні програмно-апаратні рішення. Одним із таких рішень є СВВ (IDS) [6]. Це система, яка відстежує мережевий трафік на предмет аномального впливу і повідомляє користувачів про виявлення аномального трафіку. Перші СВВ з'явилися понад три десятиліття тому [8]. Перші великі розробки у цій галузі було зроблено ВПС США. У 1980 році Джеймс Андерсон, спеціаліст з комп'ютерної безпеки та член Ради з науки в галузі оборони при ВПС США, представив доповідь про моніторинг та спостереження загроз комп'ютерній безпеці [7]. У цій доповіді були представлені основні ідеї СВВ, за якими було побудовано першу модель. Робота даної моделі ґрунтувалася на

постійному аналізі мережного трафіку та порівнянні його зі списком відомих загроз.

З кінця 1980-х років адміністратори корпоративних мереж стали включати СВВ до інструментів, що забезпечують безпеку мережі, і застосовували їх у роботі. Однак при роботі були виявлені недоліки використання систем, такі як неможливість виявити атаки нульового дня, оскільки трафік порівнювався тільки зі списком відомих сигнатур, та використання великої кількості ресурсів для постійного сканування.

Вирішення проблеми виявлення нових атак прийшло з впровадженням нового методу – ВА. Він був заснований на виявленні нетипових поведінок у мережі та забезпечував виявлення аномальних ситуацій.

В даний час найбільший інтерес становлять СВВ, що працюють на хмарних обчисленнях. На підставі даних компанії Thread Stack, IDS є одними з технологій, що найбільше продаються в галузі забезпечення безпеки [7].

СВВ відрізняються одна від одної, в основному, за місцем збору інформації, механізмом виявлення і за швидкістю реагування [4]. На основі даних критеріїв наведемо класифікацію СВВ [5].

За місцем збору інформації IDS поділяються на:

— хостові (HIDS), контролюють безпеку комп'ютера чи системи від внутрішніх та зовнішніх атак. До внутрішніх атак відносяться аномальні види поведінки програм, спроба доступу до заборонених ресурсів. До зовнішніх атак належать аномальна активність або вміст пакетів, виявлені під час аналізу мережевих взаємодій. До мінусів таких систем можна віднести збір даних на рівні вузла та пасивність системи;

— мережеві (NIDS), відстежують мережевий трафік, і при ВА, які повідомляють про це адміністраторів. Мережеві СВВ контролюють велику кількість комп'ютерів, тим самим мають більше інформації про те, що відбувається в мережі.

За швидкістю реагування IDS поділяються на [9]:

— динамічні системи, працюють у реальному часі. Дані IDS як реального часу відстежують трафік на наявність аномальних ознак. Вони

дозволяють здійснювати своєчасний моніторинг мережі та виявляти аномальні дії максимально швидко. Однак, за такого режиму роботи споживається велика кількість ресурсів системи;

— статичні системи, здійснюють аналіз мережі, перевіряють зміни мережевих сервісів. Вони надають відомості про атаку та допомагають усунути заподіяну шкоду. Також їх можна застосувати для одержання відомостей про механізм атаки та для запобігання наступним атакам такого типу.

За механізмом ВА IDS поділяються на:

— сигнатурні. У цьому типі систем кожен пакет мережного трафіку порівнюється з шаблонами атак, які у базі даних атак. До плюсів даного методу відносяться простота використання та низький показник хибних спрацьовувань при хорошому підборі сигнатури атак. Основними недоліками є необхідність збирання якісної бази сигнатур атак та складність детектування раніше невідомих або не внесених до бази атак;

— системи ВА. IDS, засновані на ВА, аналізують дії, що відбуваються в мережі. Стандартна поведінка в мережі визначається адміністратором або за допомогою навчального НД при розробці системи. Дії, які не вписуються у рамки стандартної поведінки, вважаються аномальними. Аналіз трафіку щодо аномалії дає можливість виявляти атаки, невідомі системі раніше. До недоліків такого типу систем належать завдання правил визначення аномалій та сильна залежність ефективності роботи від них [10].

Приклади СВВ, що належать до різних типів, представлені у таблиці 1.1.

Таблиця 1.1 – Приклади СВВ різного типу

За місцем збору інформації	Хостові	Мережеві
	OSEEC	Snort IIS, Cisco Secure IDS, Dragon Enterasys
За швидкістю реагування	Статичні	Динамічні
	Hummingbird	ІКС UTM+, Ребус СОВ, Рубікон
За методом виявлення	Сигнатурні	Пошук аномалій
	Suricata	Snort, Bro-IDS

1.3 Архітектура систем виявлення вторгнень

Усі СВВ можна поділити на дві категорії: автономні та клієнт-серверні системи. Перші збирають інформацію, проводять її аналіз та видають попередження, працюючи на одному хості. Системи, що належать до другої категорії, мають більш складну структуру. Давачі IDS встановлюються в найбільш уразливих місцях корпоративної мережі, здійснюють аналіз та передають оповіщення на загальну консоль управління.

Архітектуру СВВ поділяють на локальну та глобальну [6]. Локальна частина складається з елементарних складових, які у глобальній архітектурі зв'язуються та застосовуються для обслуговування систем.

Основні компоненти СВВ представлені на рис. 1.4.

Рисунок 1.4 – Компоненти СВВ

Сенсори відповідають за взаємодію з системою та за отримання даних, необхідних для виявлення атаки. Сенсори IDS бувають різних типів, кожен з яких взаємодіє з конкретною частиною системи, що захищається. До основних типів сенсорів відносять мережні, системні, прикладні та сенсори сервісу безпеки. Отримання даних здійснюється із системних журналів, мережевих

адаптерів, ядра ОС.

Сенсори передають зібрану інформацію у підсистему збору та обробки інформації. На даному етапі інформація уніфікується, фільтрується та зберігається у БД зареєстрованих подій. Оброблені дані передаються у підсистему аналізу даних.

Залежно від передбачених методів аналізу, отримана інформація перевіряється на наявність аномальних подій. Для здійснення аналізу система звертається до бази знань. Як було згадано раніше, СВВ розрізняються за методом виявлення атак, і залежно від методу, що застосовується, в базі знань можуть зберігатися сигнатури атак, профілі користувачів, правила відповідності нормальної роботи системи та інші дані.

Результати аналізу направляють у підсистему реагування. Якщо під час перевірки інформації виявляється підозріла активність, у консоль управління надсилається сигнал про можливу небезпеку. Далі адміністратор, який контролює роботу системи безпеки, приймає рішення про те, чи була реальна небезпека та вживає заходів щодо її запобігання.

Для ефективного та зручного використання СВВ виробники додають графічний інтерфейс користувача. Значним попитом на ринку користуються продукти, які вміють пояснити причину виявлення підозрілої активності і запропонувати можливі дії у відповідь.

У СВВ використовують різні методи виявлення підозрілих подій. Залежно від застосовуваного підходу їх поділяють на адаптивні методи, неадаптивні методи та методи, що ґрунтуються на аналізі аномалій.

До адаптивних методів належать методи, які можуть змінювати свою поведінку залежно від інформації, що надходить, або подій, що відбуваються. Як така інформація зазвичай виступає статистика отриманих даних або інша інформація, яка була отримана під час роботи системи. Такі алгоритми можуть бути корисними, коли є необхідність динамічного перевизначення вихідних налаштувань системи. Найпопулярнішими адаптивними методами виявлення вторгнень є алгоритми МН. Параметри таких алгоритмів здатні автоматично коригуватися відповідно до цієї статистики. Насправді найчастіше

використовуються алгоритми нейронних мереж.

Варто навести найпопулярніші адаптивні методи [11].

Нейронні мережі – це клас алгоритмів МН, натхненних біологічними ідеями роботи мозку. Основним компонентом мережі є нейрон, який отримує на вхід деяку інформацію, здійснює обчислення і на виході повертає результат. Нейронна мережа складається із шарів зв'язаних нейронів. Нейронні мережі довели свою ефективність у вирішенні багатьох завдань.

Генетичні алгоритми – це клас евристичних алгоритмів, заснованих на теорії еволюції Чарльза Дарвіна. За такого підходу до створення наступного покоління популяції відбираються кращі екземпляри з одержання потомства. При цьому використовуються оператори мутації, відбору та схрещування. Дані алгоритми знайшли застосування у завданнях оптимізації та моделювання.

ШІС - це клас алгоритмів, які ґрунтуються на процесах та механізмах, що відбуваються в біологічних імунні системи. Сучасні імунні системи базуються на теоріях клональної селекції, негативного відбору та ФІС. Вони застосовуються у вирішенні завдань МН, таких як розпізнавання образів, класифікація, кластеризація та оптимізація.

Системи на байєсовскій логіці - це системи, засновані на імовірнісних моделях та теоремі Байєса. На основі цієї теореми вважається, що між ймовірностями подій є певні співвідношення. Байєсівське програмування знаходить застосування в байєсовських мережах, прихованих марківських моделях та фільтрах Калмана. Дані системи застосовуються за умов, коли необхідна інформація частково недоступна.

До неадаптивних методів, згідно з [12], належать «методи, засновані на використанні апарату математичної статистики, де нормальна поведінка користувача визначається на основі статистичної обробки заданих параметрів». Наведемо приклади таких методів.

Експертна система. Це комп'ютерна система, котра імітує людське ухвалення рішень. У СВВ вони працюють за наперед визначеним набором правил, які описують атаку. Весь механізм обробки подій представляється як if-then-else. Приклад таких систем є Wisdom&Sense, ComputerWatch.

Сигнатурні методи. Сигнатурний аналіз заснований на семантичному аналізі атаки, представленої у певному форматі (сигнатурі). Цей метод виявляє атаки шляхом порівняння даних, що надходять з відомими сигнатурами атак. За підсумками сигнатурного аналізу працюють системи Real Secure, NetRanger.

Графи сценаріїв атак. Даних підхід ґрунтується на поданні кінцевих станів системи та переходів між ними у вигляді орієнтованих графів [13]. При побудові графа атак будуються всі можливі шляхи здійснення атаки шляхом пошуку неприпустимих шляхів.

1.4 Імунні системи

ШІС належать до методів МН та належать до галузі науки, що займається штучним інтелектом. Основою ШІС є процеси, котрі відбуваються у імунних системах хребетних організмів. Біологічні імунні системи є складними системами, які складаються з антигенів, антитіл та багатьох інших компонентів. У процесі життєдіяльності організму імунна система еволюціонує, щоб забезпечити необхідний рівень захисту організму від зовнішніх впливів.

Імунітет людини складається з вродженого імунітету та набутого [14]. Вроджений імунітет захищає життєдіяльність організму від народження та супроводжує його протягом усього життя. Також протягом життя у людини формується набутий імунітет, який покликаний підтримувати захист, створюваний вродженим імунітетом і адаптуватися до змін, що відбуваються. В ІТ найбільший інтерес представляє набутий імунітет у зв'язку з більшою специфічністю та володінням властивостями адаптивності та імунної пам'яті. Основними властивостями імунних систем, які використовуються при проектуванні моделей та для вирішення різних завдань, є:

– специфічність. Ця властивість розглядається в процесі імунної відповіді, коли при попаданні антигену в організм формується захист проти даного антигену або його найближчого представника;

– саморегуляція. Властивість проявляється у здатності до формування імунної відповіді та, за необхідності, її придушення;

– імунна пам'ять. Грунтується на властивості специфічності системи, коли після імунної відповіді на антигени параметри антигену зберігаються. При повторному попаданні в організм це допомагає здійснювати більш швидку та сильну імунну відповідь. В умовах безперервних збурень, що виникають внаслідок процесів, що відбуваються всередині організму або при взаємодії з довкіллям, функціонування імунної системи забезпечується властивостями подвійної пластичності. Подвійна пластичність є властивістю багатьох біоінспіритивних (заснованих на біологічних ідеях) методів і складається з параметричної пластичності та структурної пластичності. Параметрична пластичність передбачає адаптацію системи під час роботи шляхом зміни параметрів системи. Структурна пластичність передбачає нові умови адаптації шляхом додавання чи виключення елементів у системі. В імунних системах параметрична пластичність забезпечується за рахунок зв'язків між вузлами мережі та їх інтенсивностей, а структурна пластичність виникає за рахунок додавання нових клітин під час операцій клонування та мутацій, та видалення вже існуючих при апоптозі та імунізації.

Імунна система є складноорганізованою структурою, процеси якої не до кінця вивчені та описані. Однак є чотири основні теорії, на основі яких будуються методи ШІС [15-17]. В основі даних теорій лежать такі визначення:

— лімфоцити –клітини імунної системи, які виробляють антитіла для придушення дії зовнішніх патогенів;

— антигени –зовнішні компоненти, проти яких організм здійснює імунну відповідь, вважаючи їх небезпечними;

— генна бібліотека –набір ДНК всього організму. В ШІС під генною бібліотекою мається на увазі набір найбільш пристосованих клітин системи;

— афінність - значення ступеня подібності клітин, що є основою процесу розпізнавання клітинами одне одного;

— апоптоз - операція порівняння двох клітин, у ході якої при виявленні показника афінності клітин, що перевищує значення порога афінності, залишається один з екземплярів;

— імунізація - операція порівняння двох клітин, при якій

відновлюються клітини з різними мітками приналежності до класів, видалених в ході операції апоптозу;

— клонування –процес реплікації батьківських клітин, які мають найкращі показники пристосованості;

— мутація - процес випадкової зміни частин клітини збільшення різноманітності популяції. Найчастіше застосовується до клітин, отриманих під час клонування.

АНВ був створений Форрестом у 1994 році та використаний для виявлення вірусу у комп'ютерах. АНВ ґрунтується на механізмі створення зрілих Т-лімфоцитів у тимусі. Цікавим аспектом цього процесу є те, що він відповідає за склад популяції імунних клітин, тим не має права створювати клітини, схожі з клітинами організму, які можуть викликати аутоімунну реакцію. Ця проблема відома як "self-nonself discrimination" [18]. Для вирішення цієї проблеми необхідно передбачити зміну популяції системи лише в тому випадку, коли створюються клітини, які не відповідають вихідним клітинам системи.

Обробка інформації в систему здійснюється за допомогою АНВ. У результаті відбору ВА ґрунтується на зіставленні їх із очікуваними відхиленнями [19]. Моделювання відхилень передбачає побудову можливих моделей аномальних даних (детекторів) і досягається шляхом генерації патернів, які не відповідають жодному патерну клітин системи. Далі модель аномальних даних використовується для моніторингу даних, що надходять, при якому вони порівнюються з детекторами. Процес створення детекторів показано на рис. 1.5.

АНВ використовується у вирішенні задач ВА, у розпізнаванні образів та бінарної класифікації. Залежно від області завдання вибирається подання даних, якість виявлення та кількість необхідних детекторів.

Теорія клонального відбору було запропоновано Бернетом 1959 року. Ця теорія передбачає, що у процесі життєдіяльності система може самостійно змінюватися з умов взаємодії з довкіллям. Передбачається, що система повинна передбачати патогени, які впливатимуть на неї.

У біологічних процесах це відбувається за принципами дарвінівської теорії еволюції, коли розмножуються лімфоцити при зв'язку з антигеном [20].

Рисунок 1.5 – Процес генерації детекторів

Важливою особливістю є те, що в процесі розмноження клітини її копії піддаються помилкам (процес соматичної гіпермутації). У ШІС клонування піддаються кращі клітини системи, далі в копії привносяться зміни та отриманий набір клітин конкурує з рештою клітин системи за входження в популяцію. Схема алгоритму показано на рис. 1.6.

Рисунок 1.6 – АКС

АКС використовується у завданнях розпізнавання образів, оптимізації

функцій та є складовою інших імунних алгоритмів. Найбільш відомою версією алгоритму є CLONALG (і його версії CLONALG1, CLONALG2), адаптивна версія алгоритму ACS та версія для класифікації CLONCLAS.

Теорія імунних мереж була запропонована в 1974 нобелівським лауреатом Нільсом Йерном, який намагався пояснити властивості адаптивної навчальності та імунної пам'яті. В основі цієї теорії лежить поєднання рецепторів лімфоцитів в ідіотипічну мережу. Навіть за відсутності зовнішніх патогенів у мережі відбувається взаємодія між клітинами системи. В імунних системах хребетних організмів зв'язок імунних клітин та антигенних пептидів описується поведінкою поверхневих протеїнів та їх рецепторів. ШІС, засновані на даних ідеях, називаються ФІС та належать до моделі імунокомп'ютингу. У моделі імунокомп'ютингу зв'язок між клітинами описується з допомогою механізму CPM. Дана теорія застосовується під час вирішення завдань кластеризації, розпізнавання образів та оптимізації функцій.

Алгоритм дендритних клітин [21] пов'язаний із АКС. У процесі клональної селекції імунна відповідь ініціюється специфічними антигенами лише проти чужорідних клітин. Однак було виявлено, що ці клітини самі потребують стимуляції для проведення адаптивної імунної відповіді.

1.5 Висновки до першого розділу

В цьому розділі проведено огляд мережевих загроз, наведені різні типи мережевих атак. Докладно проаналізовані існуючі СВВ, наведено їх класифікацію та типову архітектуру.

Також описано ШІС, застосування АНВ та АКС.

Застосування ШІС в роботі спричинено тим фактом, що вони містять позитивні сторони як генетичних алгоритмів, так і нейронних мереж.

2 РОЗРОБКА МЕТОДІВ ВИЯВЛЕННЯ АНОМАЛЬНИХ ЗАПИТІВ

2.1 Завдання виявлення аномальних запитів

Як методику виявлення вторгнень у систему було обрано ШІС [15]. У рамках цієї роботи було запропоновано використовувати два методи - імунокомп'ютинг та ICP. Дані методи будуть використовуватись для вирішення задачі класифікації запитів у підмодулі аналізу даних прототипу СВВ.

Методи, засновані на ШІС, використовуються для вирішення задачі класифікації даних, яка полягає у віднесенні запиту до одного з класів: клас АЗ, клас нормальних запитів. До нормальних запитів належать запити, які потенційно небезпечні кінцевої системи. АЗ включають запити, які можуть призвести до збою або до некоректної роботи системи [16].

Трафік, що надходить $Req = \{ Req_1, Req_2, Req_3, \dots , Req_n \}$ перевіряється за допомогою обраних методів, після чого кожному із запитів надається мітка приналежності до одного із класів. В результаті вихідний НД може бути представлений у вигляді $Req_c=\{<Req_i, K_i>\}(i = \overline{1,n})$, де $K_i \in \{0,1\}$ Мітка класу, що дорівнює 0, означає клас АЗ, 1 - клас нормальних запитів.

2.2 Імунокомп'ютинг

Модель імунокомп'ютингу ґрунтується на принципах $C_i = < P, K, Fit >$ обробки інформації формальними протеїнами та імунними мережами. Ключовим поняттям у цьому підході є ФІС, котрі моделюють процес розпізнавання антигенів та проведення імунної відповіді. Імунні клітини розглядаються у вигляді формальних В-клітин та формальних Т-клітин [16]. В-клітини формуються з навчальної вибірки на основі чисельних перетворень та СРМ [17]. Т-клітини, які є детекторами, генеруються на основу роботи АНВ. Таким чином ФІС можна уявити у вигляді набору клітин $FIN = \{C_1, C_2, \dots C_n\},$ в якому клітина складається з вектора значень атрибутів клітини $P = \{P_1, P_2, \dots P_m\},$ класу

належності клітини *(K)* та значення функції пристосованості *(Fit)*. Значення функції пристосованості пропонується додати для того, щоб забезпечити еволюцію та адаптацію мережі. Функція пристосованості відображатиме якість класифікації клітиною антигенів. При початковій ініціалізації даний параметр дорівнюватиме 0, а при тестуванні системи збільшення значення буде відбувається за формулою (2.1):

$$Fit(C_i)+=\begin{cases} 0.5\ ,\ d(C_i, C_{test}) = min \\ 0.3,\ d(C_i, C_{test}) \leq h \\ 0,\ d(C_i, C_{test}) > h \end{cases} \quad (2.1)$$

Для аналізу трафіку за допомогою імунокомп'ютингу необхідно створити імунну мережу на основі навчальної вибірки, згенерувати детектори, перетворити запити у простір ФІС та класифікувати їх. Архітектура пропонованого підходу аналізу трафіку представлена на рис. 2.1.

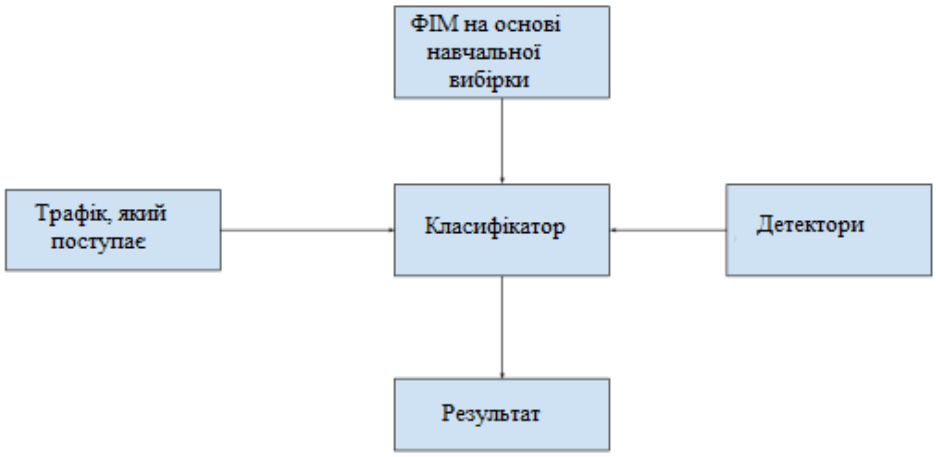

Рисунок 2.1 – Архітектура пропонованого підходу

Послідовний опис створення ФІС може бути поданий у вигляді наступного алгоритму:

1) Навчання системи. Формування ФІС починається з обробки навчальної вибірки, яка зчитується з файлу і перетворюється на рядкову матрицю. Рядки матриці містять екземпляри вибірки, які розподіляються по стовпцях матриці

відповідно до значень атрибутів. Кожному значенню атрибутів ставиться у відповідність числове значення, після чого навчальна матриця набуває вигляду $A(k x m)$. Відповідності атрибутів та числових значень зберігаються для можливості подальшого використання.

2) Генерація детекторів. Цей процес відбувається з урахуванням АНВ, метою якого є знайти комбінації значень атрибутів, які не відповідають жодній вже існуючій комбінації. Процес відбору комбінацій починається з генерації патернів $Pt = \{Pt_1, Pt_2, \dots, Pt_l\}$, після чого перевіряється подібність патернів з екземплярами навчальної матриці $A = \{A_1, A_2, \dots A_k\}$. Якщо ідентичний екземпляр не було знайдено, цей патерн додається до набору детекторів $D = \{D_1, D_2, \dots D_t\}$. Кількість детекторів було визначено як $t = 0{,}5 \; x \; k$.

Після формування необхідної кількості детекторів формується загальна навчальна матриця $Matrix = \begin{bmatrix} A \\ D \end{bmatrix}$, де $A(k x m)$ – матриця, сформована з урахуванням навчальної вибірки, $D(t x m)$ - матриця детекторів, сформована з урахуванням АНВ.

3) СРМ. В основі імунокомп'ютингу лежить припущення, що поведінка білків у системі може бути описана за допомогою СРМ.

Припустимо, що M - це матриця розміру $m x n$, елементи якої належать до поля K, яке є або полем дійсних чисел, або полем комплексних чисел. Тоді існує декомпозиція матриці M, звана СРМ, що зумовлює приведення до канонічного виду. Після розкладання матрицю M можна подати у вигляді добутку $M = USV$, де U - матриця лівих сингулярних векторів розмірністю $m x m$, S - діагональна матриця з невід'ємними речовими значеннями розмірністю $m x n$; V - матриця правих сингулярних векторів розмірністю $n x n$.

Тоді проекція образу в простір ФІС, побудованої на основі клітин сингулярного розкладання, обчислюватиметься за формулою (2.2):

$$w_i = \frac{1}{s_i} Z^T V_i \qquad (2.2)$$

де w_i - i -те значення вихідного образу, s_i - i -те сингулярне значення матриці S, Z^T - транспонований вектор вхідного образу, V_i - правий i -ий сингулярний вектор.

Для опису простору ФІС було обрано $i = 3$, при якому клітини ФІС можна подати у вигляді матриці $C = \begin{pmatrix} C_{11} & \cdots & C_{13} \\ \vdots & \ddots & \vdots \\ C_{n1} & \cdots & C_{n3} \end{pmatrix}$, де кожному рядку відповідає значення класу клітини та функції $(K \in \mathbb{N})$ пристосованості $(Fit \in \mathbb{N})$.

Внаслідок СРМ *Matrix* і проектування вхідних образів формуються імунні клітини системи $FIN = \{C_1, C_2, \dots, C_n\}$. Для реалізації СРМ використовувалася бібліотека Accord.Net.

4) Апоптоз та імунізація. Дані процеси є частиною забезпечення розвитку та функціонування імунної системи. Для забезпечення процесу розпізнавання використовується метрика $d(C_i, C_j) = \|P_i - P_j\|$, де ||P|| означає евклідову норму. У основі операції апоптозу лежить ідея, що якщо $d(C_i, C_j) \leq h$, тобто. клітина C_i впізнає клітину C_j, то C_j видаляється з поточної популяції клітин. При імунізації, якщо $d(C_i, C_j) \leq h$, але вони належать різним класам, C_j додається до популяції FIN.

При спільному застосуванні даних операцій кількість клітин системи зменшується за рахунок видалення повторюваних або дуже схожих однотипних клітин.

5) Вибір ФІС. Через застосування вищеописаних правил і при виборі різного порога афінності можуть формуватися різні мережі. Щоб вибрати оптимальне значення порога афінності, потрібно провести серію формувань мережі з різними значеннями, підрахувати індекс нероздільності і вибрати мережу з мінімальним значенням даного показника. Підрахунок індексу нероздільності виконується за такою формулою (2.3).

$$i = \ln(n) - \ln(s) - \ln(h) \tag{2.3}$$

де n - початкова кількість клітин у системі, s - число клітин в результаті апоптозу та імунізації, h - порогове значення афінності.

У процесі етапу навчання формується ФІС, що складається з клітин навчальної вибірки та згенерованих детекторів, які пройшли операції апоптозу та імунізації.

Для підтримки різноманітності у системі та підвищення якості виявлення аномалій пропонується використовувати АКС. Ключовими операціями даного методу є клонування клітин та мутація. Надихаючись ідеями генетичних алгоритмів, пропонується здійснювати клонування клітин системи з найкращими значеннями функції пристосованості. Кожна клонована клітина зазнає мутації. Кількість клонів клітини визначається як $CL = d(C_i, C_j) \times YK,$ де YK дорівнює значенню параметра рівня клонування. Кількість мутацій обчислюється як $Mut = (1 - d(C_i, C_j)) \times m \times YM$ і залежить від рівня мутації *(YM)*. Клітини, отримані в результаті АКС $CM = \{CM_1, CM_2, \dots, CM_q\},$ проходять відбір, при якому, $Fit(CM_i) > Fit(C_j)$, то $FIN_j = CM_i.$. Таким чином, в ході відбору визначаються кращі клітини, які займуть місце найменш пристосованих клітин мережі.

2.3 Імунні системи розпізнавання

Для вирішення задачі класифікації запитів пропонується використовувати імунну систему, що складається з лімфоцитів (B-клітин) та клітин пам'яті (B^m - клітин) [19]. Тоді імунна система представляється у вигляді $AIS = B \cup B^m.$ Як метрика у цьому підході використовується поняття афінності. Афінність двох елементів показує відношення між кількістю збігів компонент до норми, що є відстанню між елементами.

Навчання системи відбувається за алгоритмом, наведеним нижче.

1) Формування клітин системи на основі навчальної вибірки. У цьому підході в якості навчальної вибірки пропонується використовувати лише АЗ.

Клітини системи формуються з навчальної матриці, де кожному атрибуту у відповідність ставиться числове значення. Таким чином, клітина системи $C_i =< P, K, Fit >$, складатиметься зі значень атрибутів, класу приналежності клітини та функції пристосованості клітини, яка аналогічна до функції пристосованості, що застосовується в імунокомп'ютингу. Таким чином, тут навчальна матриця матиме вигляд $B = \begin{pmatrix} C_{11} & \cdots & C_{1m} \\ \vdots & \ddots & \vdots \\ C_{n1} & \cdots & C_{nm} \end{pmatrix}$, де n - кількість клітин у системі, m - кількість атрибутів клітини. Також кожному значенню відповідають мітка приналежності класу $(K \in \mathbb{N})$, і значення функції пристосованості $(Fit \in \mathbb{N})$.

Для функціонування системи необхідно задати такі параметри: рівень клонування *(YK)*, рівень мутації *(YM)*, поріг афінності *(AL)*.

2) Формування генної бібліотеки. Генна бібліотека складається з B^m - клітин, які відбираються на основі показників, що характеризують ефективність застосування клітини під час роботи системи. Як процедуру відбору клітин було запропоновано використовувати метод Уорда, в якому здійснюватиметься кластеризація АЗ. Клітками генної бібліотеки стануть клітини, розташовані у центрі кластерів. Якщо кластер складається з одного елемента, він і буде центром. Якщо кластер складається з кількох елементів, то центром буде вважатися елемент, який найближче розташований до теоретичного центру, формула (2.5):

$$\varpi^0 = \left\{ \frac{\sum_j \varpi_1^j}{M}, \frac{\sum_j \varpi_2^j}{M}, \dots, \frac{\sum_j \varpi_m^j}{M} \right\} \quad (2.5)$$

де $\frac{\sum_j \varpi_i^j}{M}$ дорівнює середньому значенню i-ого атрибуту кластера.

3) Апоптоз та імунізація. Операції апоптозу та імунізації здійснюються на основі ідей, викладених при описі навчання ФІС.

4) Клонування та мутація. Задля більшої динамічності у функціонуванні

системи використовується АКС. В якому процедура клонування та мутації клітини виглядає наступним чином (лістинг 2.1):

```
Мутація та Клонування (Cell1, Cell2):
Affin = AffinDist (Cell1, Ce112)
Clone=[ ]
CloneCol = affin* YK // Рівень клонування
Mutantcol = (1-affin) * vec1en * YM // Рівень мутації Від 1 до
CloneCol :
Cell3 = copy(Cell1)
        Від 1 до Mutantcol:
        NUM = Random(0, Veclen)
        word = Random(0,1)
        Cell3 [num] = word;
        Clone.Add(Cell3);
```

Лістинг 2.1 – Процедура клонування та мутації клітини

Тут AffinDist – афінна відстань між клітинами, veclen – довжина вектора атрибутів, CloneCol – кількість клонів клітини, Mutantcol – кількість мутованих клітин.

Класифікація запитів, що надходять, відбувається за наступним принципом:

1) запит $Req_i = \{Req_{i1}, Req_{i2}, Req_{i3}, \dots, Req_{im}\}$, який надходить, транслюється в простір ICP на основі перетворення значення Req_{ij} та C_{kl}, шляхом зіставлення значень.

2) Пошук найближчої клітини системи

$$f(C^*) = d_a(C_i, Req_i) \to max, (i = 1, \dots, n),$$

має на увазі знаходження клітини системи, для якої показник афінної відстані має максимальне значення. Обчислення афінної відстані між клітинами виконується так $d(C_i, C_j) = \frac{\Sigma_k s_k}{m}$, де $S_k = \begin{cases} 0, & P_{ki}(C_i) - P_{kj}(C_j) = 0 \\ 1, & \text{в іншому випадку} \end{cases}$

3) Якщо клітина системи, відстань до якої більша за поріг афінності не була знайдена, отже, запиту присвоюється клас $Fit(Req_j) = 1.$ Інакше, коли $d(C_i, Req_j) \geq AL$, запит вважається аномальним та $Fit(Req_j) = 0.$

2.4 Порівняння методик

Обидва методи ґрунтуються на теоріях імунних систем. Система містить клітини, що складаються з масиву значень атрибутів, класу приналежності клітини та функції пристосованості. Процес навчання систем є схожим.

Відмінності полягають у навчальних вибірках та у принципах формування навчальної матриці. У ФІС як навчальну вибірку використовуються дані нормальних запитів та АЗ, перетворення здійснюється за допомогою сингулярного розкладання матриць. В ІСР навчальна вибірка складається з даних АЗ, навчальна матриця утворюється за допомогою числового представлення значення атрибутів. Клітини в системах відсортовані за зменшенням значення функції пристосованості.

Для збільшення різноманітності клітин та еволюціонування в обох системах використовується АКС. Також для генерації детекторів ФІС використовується АНВ. У ІСР використовується поняття генної бібліотеки, у якій зберігаються клітини системи з найкращими показниками функції пристосованості.

У ході процедури класифікації запитів у ФІС здійснюється пошук найближчої клітини системи і результатом стає клас її приналежності. Класифікація ІСР ґрунтується на знаходженні клітини системи, що «дізнається» вхідний запит. Якщо така клітина була знайдена, запит вважається ймовірно аномальним [20].

2.5 Набір даних CICIDS 2017

Як НД, що використовується для навчання та виявлення аномалій прототипу СВВ, був обраний CICIDS 2017. Цей НД був запропонований співробітниками Канадського інституту кібербезпеки (CIC) [22]. CICIDS 2017 має великий інтерес серед дослідників, оскільки в ньому містяться дані, які дуже схожі з реальними даними мережевого трафіку. Створення реалістичного фонового трафіку було основним пріоритетом при побудові НД. У наборі є

доброякісний фоновий трафік і трафік, що містить сучасні мережеві загрози. Для формування високоякісного фонового трафіку були використані абстрактні поведінки 25 юзерів з використанням протоколів HTTP, HTTPS, FTP, SSH, та електронної пошти. Поведінка користувачів будувалася на основі B-профільної системи.

Для побудови надійного еталонного НД висунули критерії відповідності [22]:

– повна конфігурація мережі;

– повний трафік: наявність агента профілювання користувача та 12 різних машин у мережі жертви та мережі атак;

– позначений НД: мітки безпеки та відомості про атаки кожного дня;

– повна взаємодія: дві різні мережі, які з'єднані за допомогою інтернету;

– повне захоплення: весь трафік захоплюється та зберігається на сервері;

– доступні протоколи (зокрема HTTP, HTTPS, FTP, SSH та протоколів електронної пошти);

– розмаїтість атак: включені найбільш поширені атаки, котрі базуються на звіті McAfee, такі як Web based, Brute force, DoS, DDoS, Infiltration, Heart-bleed, Bot і Scan, описані в цьому НД;

– неоднорідність: захоплення мережевого трафіку від усіх машин-жертв упродовж здійснення атак;

– набір об'єктів: більше 80 об'єктів мережного потоку, взяті з мережевого трафіку, згенерованого із використанням CICFlowMeter, та представлення НД мережевого потоку у вигляді CSV -файлу;

– метадані: представлення НД, що містить опис часу, атак, потоків та міток [23].

CICIDS 2017 відповідає всім перерахованим критеріям. Дані розташовані у двох папках: MachineLeamingCVE та TrafficLabeümg. MachineLeamingCVE містить 8 файлів у форматі .cve, призначених для МН. Дані поділені по днях тижня та доби. Збір даних розпочався 3 липня 2017 року у понеділок та закінчився 7 липня 2017 року у п'ятницю. Кожен файл містить різні типи трафіку. Опис змісту даних представлені у табл. 2.1.

Таблиця 2.1 - Опис файлів

Дата	Назва файлу	Зміст
Понеділок, 3 липня 2017	Monday-WorkingHours.pcap_ISCX.csv	Benign (Normal human activities)
Вівторок, 4 липня 2017	TuesdayWorkingHours.pcap_ISCX.csv	Benign, FTP-Patator, SSH-Patator
Середа, 5 липня 2017	Wednesday-workingHours.pcap_ISCX.csv	Benign, DoS GoldenEye, DoS Hulk, DoS Slowhttptest, DoS slowloris
Четвер, 6 липня 2017	Thursday-WorkingHours-Morning-WebAttacks.pcap_ISCX.csv	Benign, Web Attack - Brute Force, Web Attack - Sql Injection, Web Attack - XSS
Четвер, 6 липня 2017	Thursday-W orkingHours-Afternoon-Infilteration.pcap_ISCX.csv	Bening, Infiltration
П'ятниця, 7 липня 2017	Friday-WorkingHours-Morning.pcap_ISCX.csv	Bening, Bot
П'ятниця, 7 липня 2017	Friday-WorkingHours-Afternoon-DDos.pcap_ISCX.csv	Bening, PortScan
П'ятниця, 7 липня 2017	Friday-WorkingHours-Afternoon-PortScan.pcap_ISCX.csv	Bening, DDoS

Оскільки СВВ повинна виявляти всі види атак, під час навчання системи файли CICIDS2017 було запропоновано поєднати в один файл. При аналізі цього файлу було підраховано кількість запитів, що належать кожному з класів (табл. 2.2).

Таблиця 2.2 - Кількість екземплярів

Клас	Кількість запитів
BENIGN	3217337
DoS Hulk	231073
PortScan	317860
DDoS	256054
DoS GoldenEye	10293
FTP-Patator	7938

Продовження таблиці 2.2

SSH-Patator	5897
DoS slowloris	5796
DoS Slowhttptest	5499
Bot	3932
Web Attack - Brute Force	1507
Web Attack - XSS	652
HeartBleed	11
Infiltration	36
Web Attack - Sql Injection	21

Кожен клас характеризує можливий тип поведінки користувачів, які здійснюють певний тип атак. Мітки класів показують належність запиту до одного з типів атак, а також яким чином цей запит було сформовано. У НД CICIDS2017 містяться запити з мітками таких класів:

1) BENIGN - безпечні фонові запити від користувачів;

2) DoS Hulk – це інструмент відмови в обслуговуванні веб-сервера, написаний для дослідних цілей. Він призначений для створення обсягів унікального і заплутаного трафіку на веб-сервері, минаючи механізми кешування і, отже, потрапляючи в прямий пул ресурсів сервера;

3) PortScan - атаки, засновані на вразливості відкритих портів у системі;

4) DDoS - розподілена атака типу відмова в обслуговуванні здійснюється з різних джерел шляхом заповнення цільової машини або ресурсу надмірними запитами в спробі перевантажити систему і запобігти виконання деяких або всіх законних запитів;

5) DoS GoldenEye - дослідний інструмент для тестування безпеки системи, спрямований на навантаження HTTP серверів з метою паралізувати їхню роботу;

6) FTP-Patator - багатопотоковий інструмент, спрямований на зламування FTP сервера за допомогою підбору паролів;

7) SSH-Patator - багатопотоковий інструмент, спрямований на зламування SSH сервера за допомогою підбору паролів;

8) DoS slowloris - змушує сервер атакувати обслуговувати велику кількість відкритих з'єднань за допомогою відправки незавершених HTTP - запитів. Сервер очікує завершення з'єднань та перестає бути доступним для інших користувачів;

9) DoS Slowhttptest - інструмент, що реалізує атаки типу відмова в обслуговуванні на рівні додатків, спрямованих на заповнення пулу можливих підключень до сервера;

10) Bot – мережа комп'ютерів, заражених шкідливою програмою, за допомогою якої зловмисники можуть керувати комп'ютером;

11) Web Attack – Brute Force – атаки на сервіси, при яких зловмисник намагається підібрати логін або пароль для отримання доступу;

12) Web Attack – XSS – атаки на веб-системи з метою отримання необхідної інформації або доступу за допомогою впровадження на веб-сторінки шкідливого коду;

13) HeartBleed – атака, спрямована на отримання інформації з сервера або клієнта, що використовують OpenSSL за допомогою некоректного HeartBeat запиту;

14) Infiltration – атаки, спрямовані на проникнення в систему;

15) Web Attack - Sql Injection – атака, спрямована на злам сайтів та програм, що працюють з БД, за допомогою впровадження SQL запитів.

Опис кожного профілю здійснюється на підставі 80 характеристик. Розробники використовували різні статистичні показники пакетів інкапсуляції мережевих подій.

TrafficLabelling містить повний пакет навантажень у форматі захоплення пакетів (PCAP). Вибірка містить усі атрибути профілів системи, а також значення Flow ID, Source IP, Source Port, Destination IP, Protocol, Timestamp. Ця вибірка також поділена на кілька файлів, що відповідають дням тижня, часу доби та характеру запитів до об'єкта. Вибірка може використовуватись для тестування системи.

Для роботи з вибраним НД було реалізовано клас CICIDS. Цей клас надає можливості зчитування вибірки, нормалізації даних, а також функції для аналізу

складу та розбиття загальної вибірки відповідно до поставлених критеріїв.

2.6 Реалізація системи виявлення вторгнень

Для реалізації СВВ було обрано середовище розробки Microsoft Visual Studio 2017 та мову програмування C#. Як програмне рішення реалізації проекту було обрано клієнт-серверну архітектуру. Таким чином, СВВ складатиметься з двох частин – клієнтського та серверного додатків. Обидві частини були реалізовані у вигляді додатків Windows Forms, тестування яких виконувалося на основній машині під ОС Windows 10 та віртуальній машині під ОС Windows 8.1. Архітектура пропонованого підходу показана на рис. 2.2.

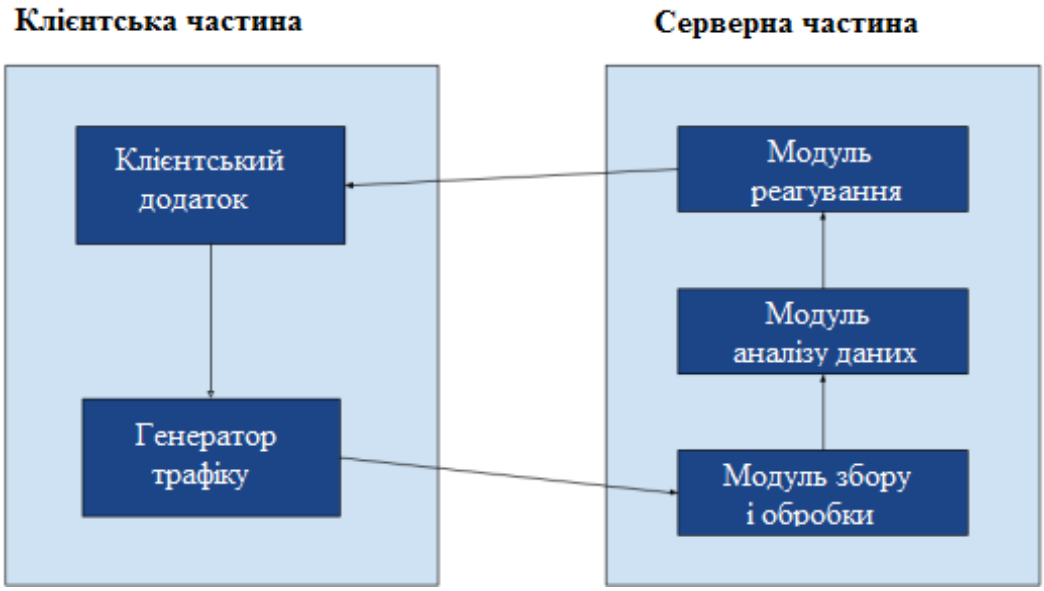

Рисунок 2.2 – Архітектура розробленої СВВ

Клієнтська програма містить інтерфейс користувача, який дозволяє користувачеві проводити перевірку трафіку за допомогою СВВ (рис. 2.3). Після натискання користувачем кнопки «Почати перевірку» починається генерація трафіку на основі НД CICIDS 2017, який відправляється на сервер для подальшого аналізу. Зв'язок між клієнтським та серверним додатком здійснюється на основі TCP сокетів за допомогою можливостей бібліотеки System.Net.

Рисунок 2.3 – Вікно клієнтського додатку

Після отримання трафіку на сервері відбувається виділення ключових компонентів запиту та подання їх у вигляді, необхідному для подальшого аналізу. Аналіз запитів здійснюється на основі комбінації двох методів: ФІС та ІСР. Опис роботи обох методів було описано раніше. При тестуванні системи було встановлено, що комбінація методів дає кращий результат, тому як компонент аналізу був обраний даний підхід.

Комбінація методів ґрунтується на показниках рівня виявлення систем. Показники рівня виявлення системи зберігаються у файлі і після кожного тестування змінюються залежно від ефективності роботи системи. Як новий показник рівня виявлення у файлі зберігається середнє арифметичне попереднього показника та показника, отриманого під час тестування. Формування результату з урахуванням класифікації двох систем виконується відповідно до рис. 2.4.

За допомогою модуля реагування отриманий результат відправляється клієнту, на основі чого можуть бути вжиті дії щодо безпеки системи. Усі оброблені запити зберігаються у БД, як у серверної, і на клієнтській стороні. Для управління системою було створено дослідницьку компоненту. Ця компонента містить можливості класифікації даних та налаштування параметрів системи.

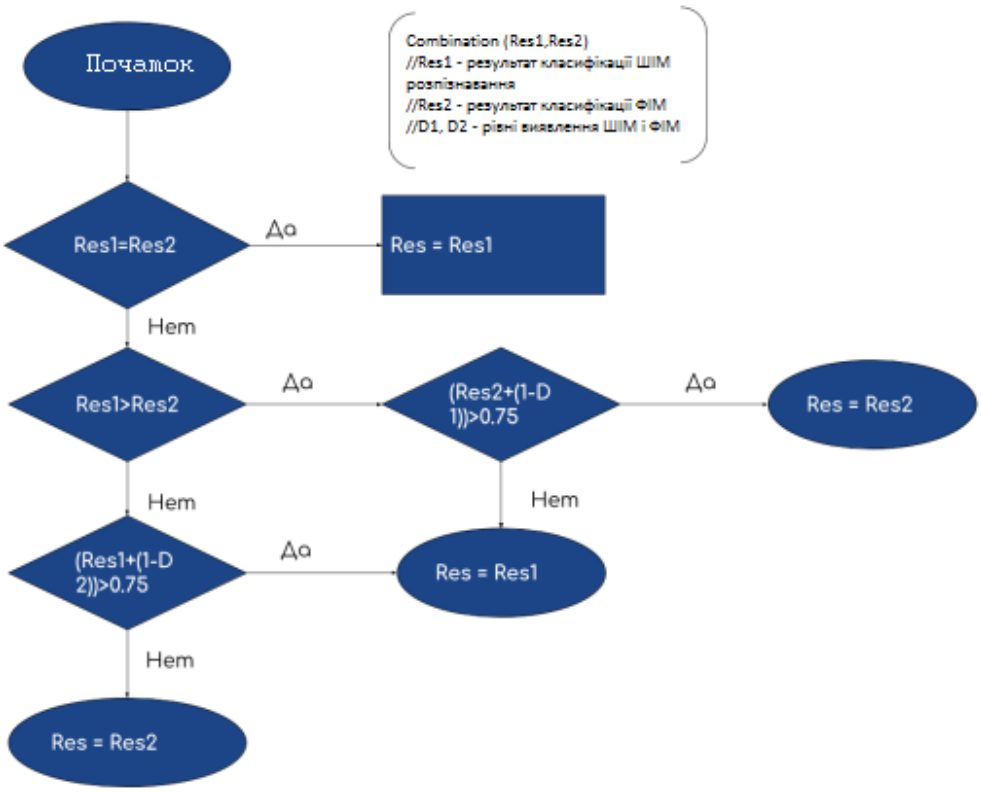

Рисунок 2.4 – Обчислення результату комбінації методів

Вікно налаштування системи показано на рис. 2.5.

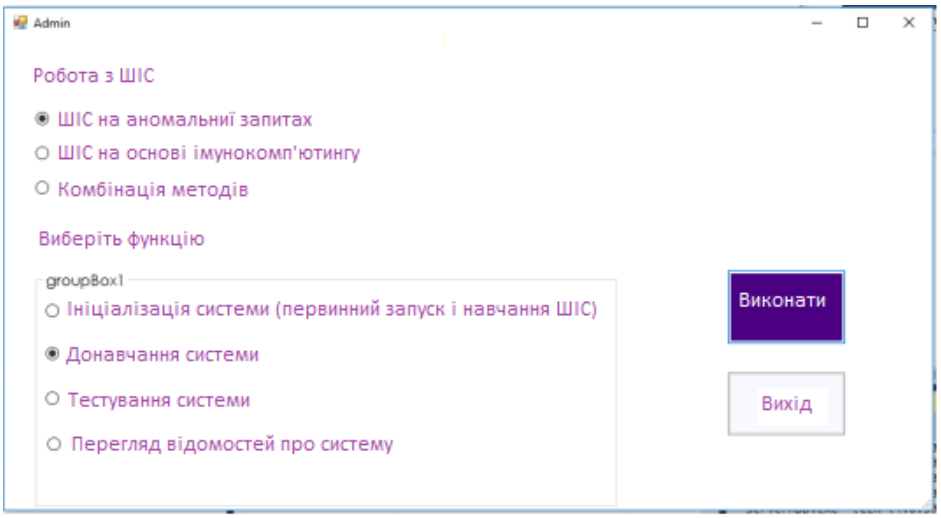

Рисунок 2.5 – Вікно налаштування параметрів системи

Реалізований функціонал дозволяє навчити систему, провести додаткове навчання системи, протестувати її роботу та переглянути відомості про систему. Основні функції, що виконують дані дії, представлені в табл. 2.3.

Таблиця 2.3 – Методи роботи зі ШІС

Опис	Назва	Ключові функції
Створення та функціонування штучної імунної системи розпізнавання	Alternative AIS .cs	InitialisationSystem(string path); TeachingSystem(AltemativeAIS ais); AdditTrainingAIS(string path, AlternativeAIS ais, List<string>[] AttributeValues); TrainingAlternativeAIS(string path); ClassificationMessage(Cell newmessage, AlternativeAIS ais); RecoveryAlternativeAIS(); SaveAlternativeAIS(AlternativeAIS ais, List<string>[] AttributeVal);
Створення формальної імунної мережі	AISImComp.cs	ApoptozImmunization(List<Cell> cells, double affdist); IndexOfInseparability(double ml, double m2, double h); InitializationSystem(string teachdata); ProjectionToFIS(SVD svdmatr, double[,] binmatr); CloningMutation(Cell cl, AISImComp ais); ClonalSelection(AISImComp fis);
Функціонування формальної імунної мережі	AISImClassif.cs	TestNormalization(string[,] test, List<string>[] AttributeValues); AdditTraining(string[,] data, AISImComp ais, SVD svd, List<string> [] AttributeValues); Classification(string[,] test); Fitnessing(double[] distance , AISImComp ais, double bestdist); EuclidDistance(Cell cl, Cell c2);
Робота з вибірками	CICIDS.cs	ReadFromFile(string path); MatrixNormalization(string[,] comp_matr, List<string>[] AttributeValues); AnalyzeData();

2.7 Висновки до другого розділу

У другому розділі роботи було запропоновано для детектування АЗ використати ФІС та ІСР. Обидва ці підходи інтегровані в складову аналізу СВВ. Із застосуванням іммунокомп’ютингу сформовано алгоритму створення ФІС. Дослідження визначили, що показники детектування ІСР кращі, ніж в ФІС.

Розроблена СВВ має клієнт-серверну архітектуру. Навчання СВВ та генерація трафіку проходили із використанням НД CICIDS 2017.

3 АНАЛІЗ ЕФЕКТИВНОСТІ СИСТЕМИ

3.1 Опис оцінки ефективності системи

Для розрахунку оцінки ефективності функціонування системи та вибору найкращої методики аналізу запитів потрібно провести тестові випробування. Для тестування системи був розроблений клас Testing, в якому є можливість вибору системи та тестового файлу (рис. 3.1).

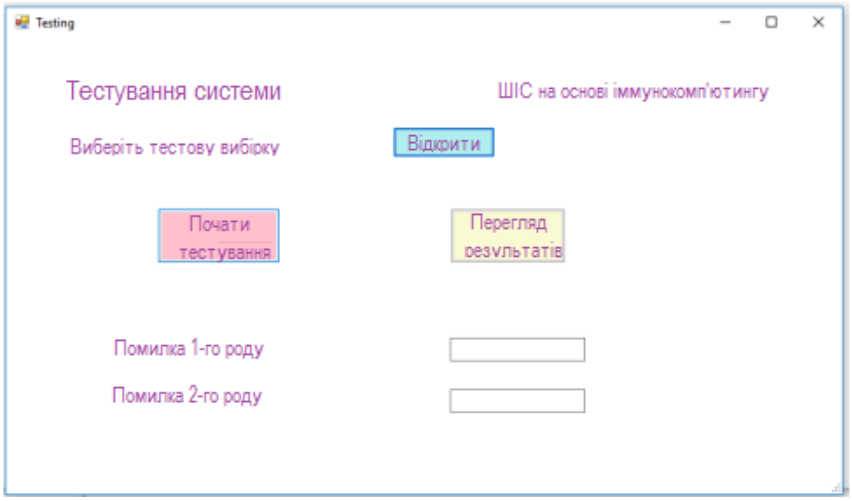

Рисунок 3.1 – Вікно тестування системи

Після натискання кнопки «Почати тестування» починається класифікація запитів, вибраних як тестова вибірка. Результати класифікації зберігаються у файл і для подальшого аналізу роботи викликається метод CompareResult(), який порівнює результати, отримані в ході класифікації та початкові мітки належності запитів. На основі отриманих результатів обчислюються параметри оцінки системи: кількість правильних спрацьовувань, кількість хибних спрацьовувань, рівень виявлення, рівень хибних спрацьовувань, помилки першого та другого роду. Всі дані зберігаються в текстовий файл для подальшого аналізу.

Для тестування кожної методики аналізу запитів було проведено експерименти, спрямовані на оцінку різних характеристик роботи системи. Експерименти мали деякі відмінності, які залежали від особливостей побудови системи.

Усі результати класифікації запитів можна розділити на чотири категорії [24] (рис. 3.2):

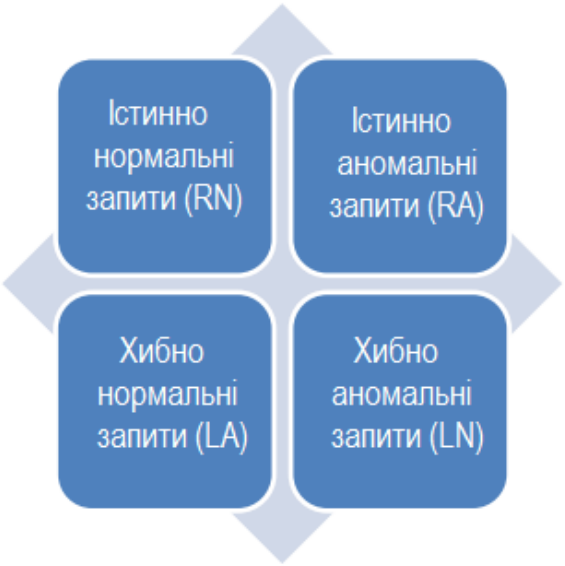

Рисунок 3.2 – Класифікація результатів системи

Кількість правильних спрацьовувань системи складається з нормальних запитів, які були класифіковані правильно (істинно нормальні запити) та АЗ, які були правильно класифіковані (істинно АЗ). Кількість помилкових спрацьовувань системи можна показати як суму кількості нормальних запитів, прийнятих за аномальні (хибно АЗ) і кількості АЗ, прийнятих нормальні (хибно нормальні запити).

Для аналізу ефективності були використані поняття помилки першого та другого роду. Дані поняття використовуються під час перевірки гіпотез у статистиці, а також у системах, де результат подається в бінарному вигляді на основі деякого критерію, з можливістю отримання помилкового результату [25]. Обчислення помилок проводиться за формулами (3.1, 3.2).

$$Err1 = \frac{LA}{M} \tag{3.1}$$

$$Err2 = \frac{LN}{M} \tag{3.2}$$

Помилка першого роду показує відношення кількості нормальних запитів, класифікованих як аномальні, до кількості тестових запитів. Помилка другого роду показує відношення кількості АЗ, класифікованих як нормальні, до кількості тестових запитів.

Для аналізу ефективності системи можуть бути використані значення рівня виявлення та хибних спрацьовувань. Рівень виявлення та рівень хибних спрацьовувань обчислювалися в ході роботи системи на основі формул (3.3) та (3.4).

$$Detect = \frac{RN+RA}{M} = \frac{R}{M} \quad (3.3)$$

$$FDetect = \frac{LN+LA}{M} = \frac{L}{M} \quad (3.4)$$

Рівень виявлення відображає сумарну кількість правильних класифікацій нормальних запитів та АЗ. Рівень хибних спрацьовувань, навпаки, показує загальну величину хибних спрацьовувань і дорівнює сумі помилок першого і другого роду.

Для оцінки компоненти аналізу, що вирішує завдання класифікації, підрахуємо характеристики точності та повноти. Для проведених досліджень, точність показує скільки класифікованих АЗ, дійсно були аномальними (формула (3.5)). Повнота системи показує частку правильно класифікованих АЗ до кількості аномальних запитів у тестової вибірці (формула (3.6)).

$$Precision = \frac{RA}{RA+LA} \quad (3.5)$$

$$Recall = \frac{RA}{RA+LN} \quad (3.6)$$

3.2 Аналіз ефективності формальних імунних мереж

Для оцінки ефективності роботи прототипу СВВ, у якому аналіз трафіку

виконується на основі методів ФІС, було проведено такі експерименти:

– визначення впливу складу навчальної вибірки на результат;
– аналіз залежності результату від розміру навчальної вибірки;
– дослідження перехресної перевірки;
– оцінка ефективності застосування АКС.

Особливістю даного підходу є вибір складу навчальної вибірки, так як для створення детекторів необхідно, щоб усі запити належали до класу нормальних. Однак із реалізацією функції додаткового навчання з'явилася можливість додати до навчальної вибірки АЗ. Для вибору оптимального складу навчальної вибірки були проведені експерименти на системі з нормальними запитами та детекторами, нормальними запитами та АЗ, та на системі, що включає всі три позиції. Результати тестувань представлені у табл. 3.1.

Таблиця 3.1 – Результати тестування ФІС при різних вибірках

Параметри системи	Нормальні запити +детектори	Нормальні запити + аномальні	Нормальні запити + аномальні + детектори
Кількість клітин у системі (N)	20112	20034	20677
Кількість тестових запитів (M)	10561	10561	10561
Кількість правильних спрацьовувань (R)	7737	8433	9004
Кількість помилкових спрацьовувань (L)	2824	2128	1557
Рівень виявлення (Detect)	0,7326	0,7985	0,8525
Рівень хибних спрацьовувань (FDetect)	0,2674	0,2015	0,1475

На основі отриманих результатів можна зробити висновок, що оптимальним підходом є використання як навчальної вибірки запитів обох типів та детекторів, створених на основі АНВ. Показник виявлення запитів такого підходу вищий за показники застосування вибірки без АЗ або вибірки без детекторів на 11,99 % та 5,4 % відповідно (рис. 3.3).

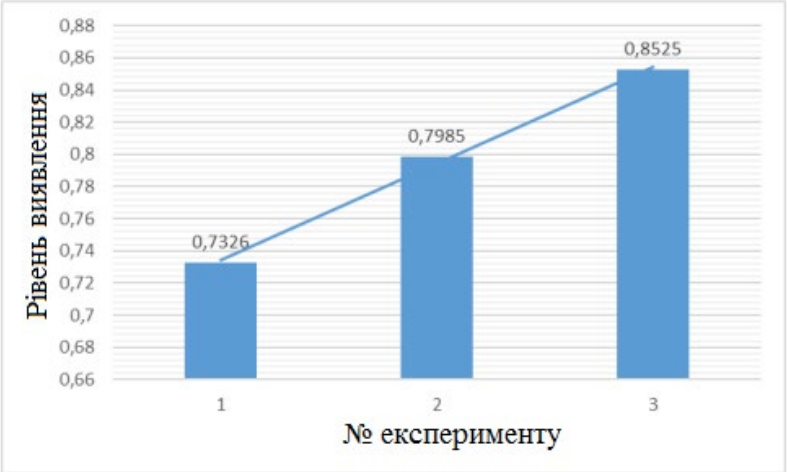

Рисунок 3.3 – Залежність рівня виявлення від складу навчальної вибірки

Іншим важливим аспектом при побудові системи є розмір навчальної вибірки. Для визначення залежності якості результату від кількості клітин у системі було проведено відповідні дослідження. У ході експериментів було виявлено обмеження на розмір навчальної вибірки, яке пояснюється складністю виконання СРМ великого розміру (понад 4 млн елементів). Однак збільшення кількості клітин може бути виконано за допомогою процедури донавчання. Результати досліджень роботи систем, що відрізняються кількістю клітин, представлені в табл. 3.2.

Таблиця 3.2 – Залежність результату від розміру навчальної вибірки

Параметри	Експ. №1	Експ. №2	Експ. №3
Кількість клітин у системі (N)	20023	100507	300619
Кількість тестових запитів (M)	20014	20014	20014
Кількість правильних спрацьовувань (R)	16211	17254	18407
Кількість помилкових спрацьовувань (L)	3803	2760	1607
Істинно нормальні запити (RN)	7823	8139	8541
Аномальні запити (RA)	8388	9115	9866
Помилково нормальні запити (LN)	2419	1692	941

Продовження таблиці 3.2

Параметри	Експ. №1	Експ. №2	Експ. №3
Невірно аномальні запити (LA)	1384	1068	666
Помилка 1 роду (Err1)	0,069	0,0535	0,033
Помилка 2 роду (Err2)	0,121	0,0845	0,047
Рівень виявлення (Detect)	0,81	0,862	0,9197
Рівень хибних спрацювань (FDetect)	0,19	0,138	0,0803
Точность (Precision)	0,8583	0,8951	0,9367
Полнота (Recall)	0,7761	0,8434	0,9129

На основі даних результатів зауважимо, що значення рівня виявлення збільшується зі збільшенням кількості клітин у системі. Також збільшуються показники точності та повноти системи, максимальний результат у 93,67% та 91,29%, яки1 досягається при найбільшому розмірі системи. Цей факт можна пояснити тим, що зі збільшенням кількості клітин у системі збільшується їхня різноманітність, і зростає ймовірність знаходження найближчої клітини системи правильного класу. Порівняння показників помилок 1-го та 2-го роду, а також рівня хибних спрацьовувань представлено за допомогою гістограми на рис. 3.4.

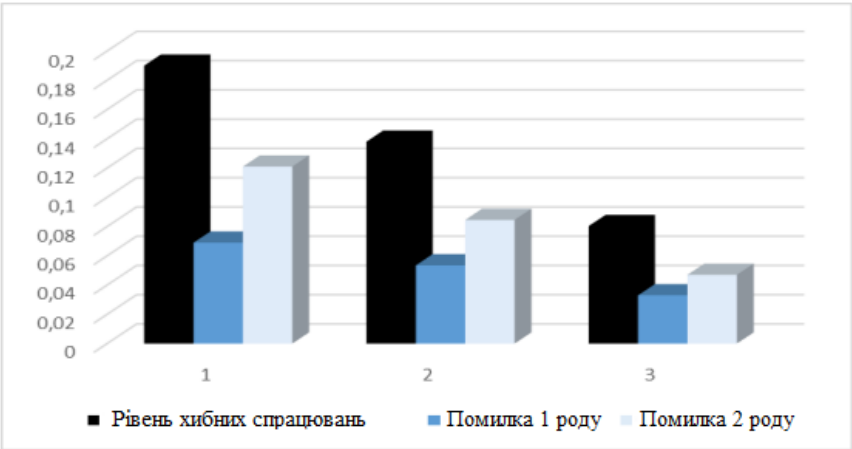

Рисунок 3.4 – Залежність показників ФІС від величини вибірки

Дослідження перехресної перевірки було спрямоване на емпіричну оцінку

узагальнюючої здатності методик, що застосовуються, а також на перевірку працездатності системи в умовах атаки «нульового дня». У цьому дослідженні наявна вибірка була поділена на шість частин відповідно до класів запитів. Потім на п'яти частинах виконувалося навчання системи, а частина даних, що залишилася, застосовувалася для тестування працездатності системи. Загальна вибірка була поділена на такі частини: Benign, Dos, Web Attack, Patator, PortScan, DDoS. Для складання вибірки з необхідних класів та заданої кількості екземплярів було реалізовано клас AutoTesting.

Як навчальну вибірку було використано 100000 запитів, як тестову вибірку 50000 запитів. Під час проведення даних досліджень було отримано результати, подані у табл. 3.3.

Таблиця 3.3 – Результати кросвалідації

Перевірка атаки	DDoS	Dos	Web attack	Patator	PortScan
Detect	0,6526	0,7412	0,5482	0,8560	0,8984
FDetect	0,3474	0,2588	0,4518	0,144	0,1016

Виходячи з отриманих результатів відзначимо, що найвищий рівень виявлення атак Patator та PortScan. Дані атаки спрямовані на злам сервера за допомогою пароля та сканування портів системи. Найменший показник виявлення показали мережеві атаки, такі як XSS та SQL-ін'єкції.

Для підтримки різноманітності клітин у системі та підвищення якості детектування, механізми роботи були додані операції клонування та мутації. Вони є основою АКС. Ефективність використання такої методики була перевірена при тестуванні системи 100000 клітин на 50000 тестових запитів. Після кожного тестування до системи застосовувалась клональна селекція. Експериментальні результати наведені у табл. 3.4.

Таблиця 3.4 – Вплив клональної селекції на показники виявлення

Число кроків	0	10	50	100
Рівень виявлення	0,8452	0,8567	0,8894	0,8926
Рівень хибних спрацьовувань	0,1548	0,1433	0,1106	0,1047

У результаті експериментів було встановлено, що у початковому етапі застосування алгоритму дає невелике збільшення рівня виявлення, після чого відзначається стабілізація значень показників. Стабілізація може бути викликана тим, що деякі вектори зміщуються в області нормального трафіку в область аномального. При постійному зміщенні такі запити роблять внесок у загальний показник помилкових спрацьовувань (рис. 3.5).

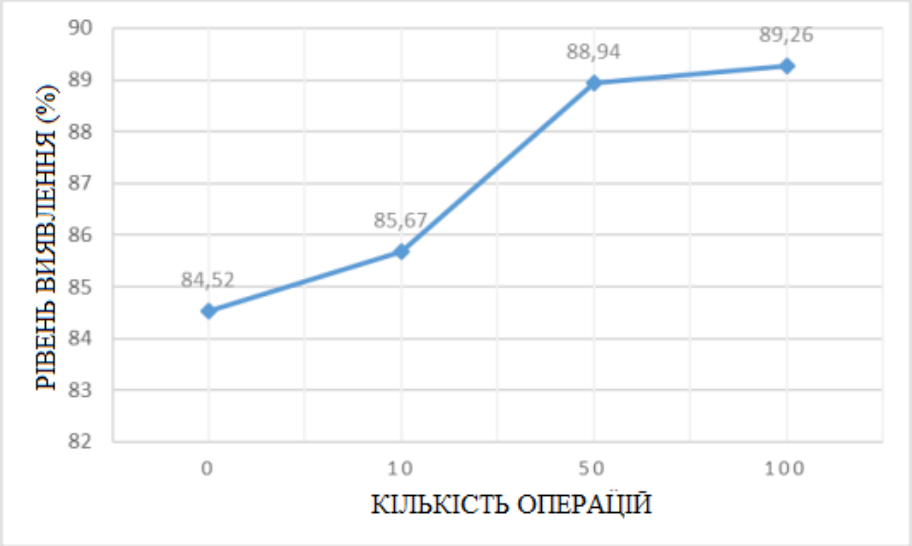

Рисунок 3.5 – Залежність показника виявлення від кількості операцій

3.3 Аналіз ефективності імунних систем розпізнавання

Для оцінки ефективності даного підходу було проведено такі експерименти:

– дослідження залежності результату від порогового значення афінності;
– аналіз залежності результату від розміру навчальної вибірки;
– дослідження перехресної перевірки;

– оцінка ефективності застосування АКС.

Проведені дослідження були спрямовані на оцінку ефективності застосування ICP для аналізу трафіку та виявлення АЗ. У ході дослідження було проведено експерименти, подібні до досліджень ФІС.

Особливістю застосування цього підходу є залежність результату від заданого граничного значення афінності. Ухвалення рішення про аномальність запиту виникає у разі знаходження клітини в системі, афінна відстань до якої перевищує поріг афінності. Отже, неправильний вибір порога афінності може призвести до збільшення хибних спрацьовувань або зниження якості детектування атак.

При проведенні експерименту ICP була навчена на 100 000 АЗ, а в якості тестової вибірки було обрано 50 000 запитів різних класів. У ході експерименту було виявлено, що за даних умов рівень виявлення перевищує 70% починаючи з показника порога афінності рівного 0,85 (рис. 3.6).

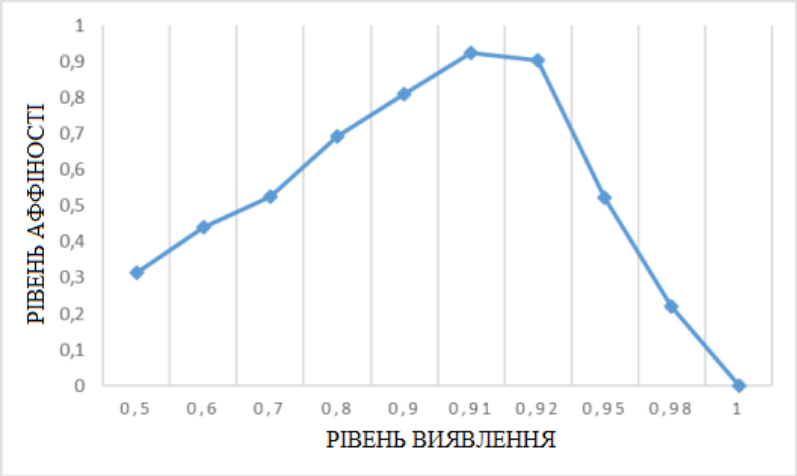

Рисунок 3.6 – Залежність рівня виявлення від порога афінності

Пошук оптимального значення здійснювався з інтервалу від 09 до 093. Результати тестування системи з вибраними показниками наведено в табл. 3.5.

Таблиця 3.5 – Результати тестування за різних порогів афінності

Порогове значення афінності:	0,90	0,91	0,92	0,93
Кількість правильних спрацьовувань (R)	40530	46280	45205	41165
Кількість помилков. спрацьовувань (L)	9470	3770	4795	8835
Помилка 1 роду (Err1)	0,11552	0,03492	0,06836	0,13428
Помилка 2 роду (Err2)	0,07388	0,04048	0,02752	0,0424
Рівень виявлення (Detect)	0,8106	0,9246	0,9041	0,8233
Рівень хибних спрацьовувань (FDetect)	0,1894	0,0754	0,0959	0,1767
Точність (Precision)	0,8081	0,9063	0,9286	0,8733
Повнота (Recall)	0,7292	0,9181	0,8397	0,6853

Таким чином, найкращий результат показала система, значення порога афінності якої дорівнювало 0,91. Також було зазначено, що до цього порога переважало значення помилки першого роду, і з перевищенням цієї позначки стало переважати кількість хибно нормальних запитів. Можливою причиною такого результату є зменшення ймовірності збігу більшої кількості атрибутів АЗ при підвищенні порога афінності.

При дослідженні залежності системи від розміру навчальної вибірки було встановлено, що зі збільшенням кількості клітин у системі зростає рівень виявлення. Для тестування бралася вибірка, що складається із 20000 запитів різних класів. Кращий показник рівня виявлення 94,55% отримано при 300786 клітинах. Проміжні результати тестування системи представлені у табл. 3.6.

Таблиця 3.6 – Залежність рівня виявлення кількості клітин у системі

Параметри	Експ. №1	Експ. №2	Експ. №3
Кількість клітин у системі (N)	20132	100411	300786
Кількість правильних спрацьовувань (R)	17940	18448	18910
Кількість помилкових спрацьовувань (L)	2060	1552	1090
Рівень виявлення (Detect)	0,897	0,9224	0,9455
Рівень хибних спрацьовувань (FDetect)	0,103	0,0776	0,0545
Точність (Precision)	0,8836	0,9235	0,9624
Повнота (Recall)	0,9281	0,9297	0,9366

У результаті експериментів було встановлено, що точність системи збільшується пропорційно розміру популяції клітин. Повнота системи також зростає, але збільшення показника не перевищує 1%. Динаміку зміни даних параметрів показано на рис. 3.7.

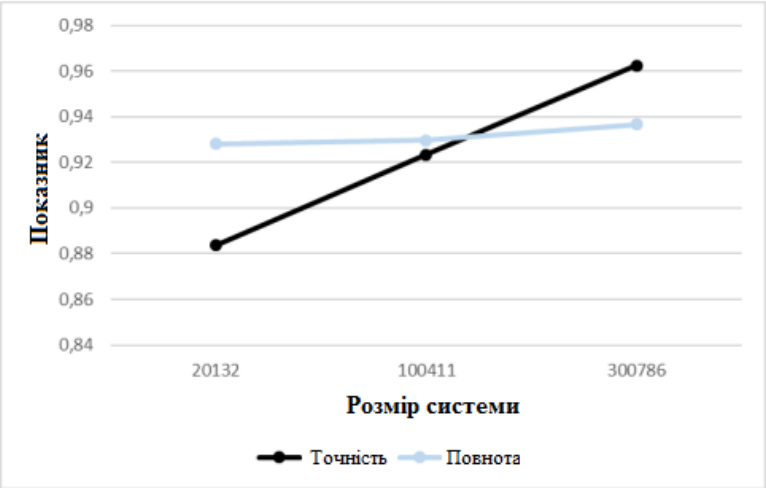

Рисунок 3.7 – Зміна показників повноти та точності системи

Як перевірку працездатності системи при виникненні атаки, яка не була раніше зареєстрована, використовувався метод перехресної перевірки. Як навчальна вибірка були сформовані файли, які не містять одного з типів атак. Навчені в такий спосіб системи тестувалися на невідомих їм атаках. Результати цього дослідження показані у табл. 3.7.

Таблиця 3.7 – Результати перехресної перевірки

Перевірка атаки	DDoS	Dos	Web attack	Patator	PortScan
Detect	0,7786	0,7968	0,8569	0,6914	0,8838
FDetect	0,2214	0,2032	0,1431	0,3086	0,1016

На відміну від методів ФІС, ІСР, що використовуються як підсистема аналізу, показують найкращий результат на атаках типу PortScan і Web (XSS, SQL-ін'єкції), а найгірший результат на атаках типу Patator (рис. 3.8).

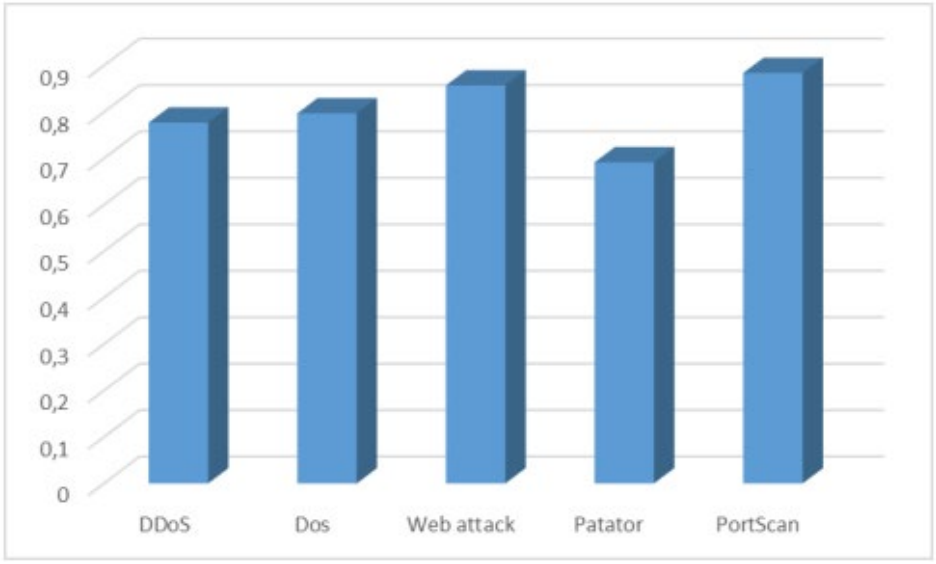

Рисунок 3.8 – Показники рівня виявлення різних типів атак

Для динамічного розмаїття популяції системи, у цьому підході також застосовувався АКС. Визначення клітин, які додаватимуться у систему, відбувалося з урахуванням ідей еволюційних алгоритмів. Ефективність застосування даного алгоритму була перевірена на системі з 100000 клітин та 50000 тестових запитах. Після кожного тестування до системи застосовувалась клональна селекція. Результати експерименту представлені у табл. 3.8.

Таблиця 3.8 – Вплив клональної селекції на показники виявлення

Число кроків	0	10	50	100
Рівень виявлення	0,9031	0,9254	0,9338	0,9389
Рівень хибних спрацьовувань	0,0969	0,0746	0,0662	0,0611

В результаті дослідження було встановлено, що при перших 10 тестуваннях відзначено найвищий показник збільшення рівня виявлення. При подальших експериментах рівень виявлення збільшується, але показник, який збільшується результат, падає (рис. 3.9).

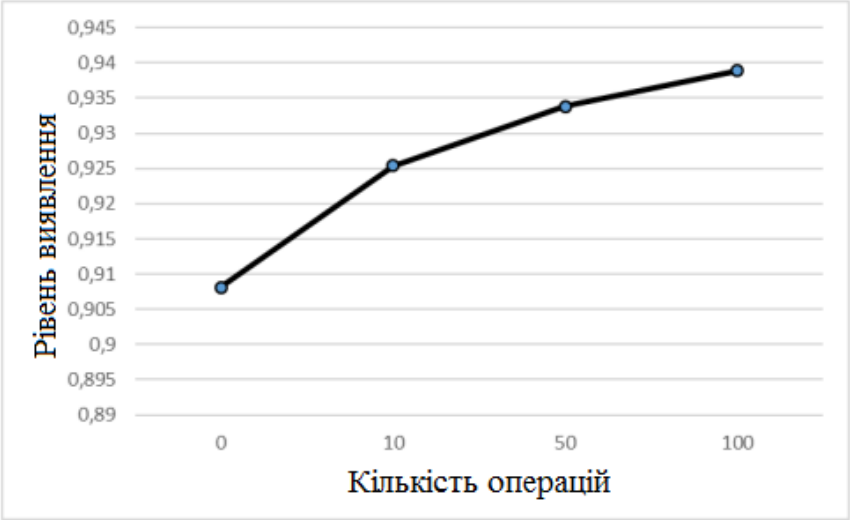

Рисунок 3.9 – Ефективність застосування АКС

3.4 Аналіз ефективності комбінації методів

Кожна з досліджених систем має свої переваги і недоліки, тому було запропоновано об'єднати роботу двох систем. Система, побудована на основі комбінації результатів двох підходів, була протестована для оцінки ефективності її роботи. Показники роботи комбінованої системи при тестуванні на вибірці з 50000 запитів наведені в табл. 3.9.

Таблиця 3.9 – Результати тестування комбінованого підходу

Параметри	Експ. №1	Експ. №2	Експ. №3
Кількість клітин у системі (N)	40155	200918	601405
Кількість тестових запитів (M)	50000	50000	50000
Кількість правильних спрацьовувань (R)	45680	47390	48235
Кількість помилкових спрацьовувань (L)	4320	2610	1765
Істинно нормальні запити (RN)	21625	22251	22865
Аномальні Запити (RA)	24037	25139	25371
Помилково нормальні запити (LN)	2463	1361	1129
Невірно аномальні запити (LA)	1857	1249	635
Помилка 1 роду (Err1)	0,03714	0,02498	0,0073
Помилка 2 роду (Err2)	0,04926	0,02722	0,02258
Рівень виявлення (Detect)	0,9136	0,9478	0,9647

Продовження таблиці 3.9

Рівень хибних спрацьовувань (FDetect)	0,0864	0,0522	0,0353
Точність (Precision)	0,9276	0,9526	0,9755
Повнота (Recall)	0,9079	0,9486	0,9573

Застосування комбінації методів показує результати виявлення вище, ніж застосування методу ФІС та ІСР. Найкращий показник (96,47%) досягається при використанні комбінації систем, сумарний обсяг яких перевищує 600 000 клітин. Порівняльна характеристика рівня виявлення різних підходів наведена на рис. 3.10.

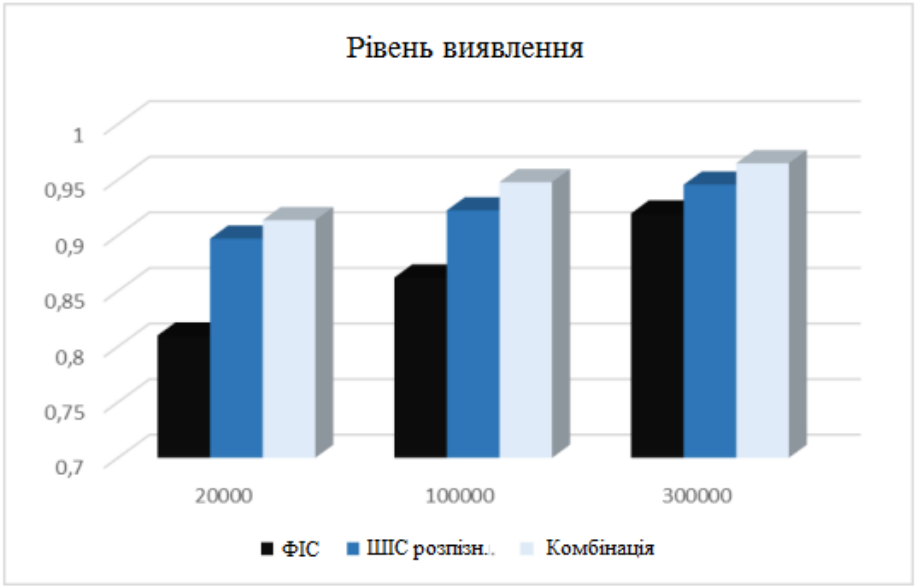

Рисунок 3.10 – Порівняння показників рівня виявлення

У ході досліджень було зазначено, що при застосуванні ФІС кількість запитів, хибно класифікованих нормальними більша, ніж кількість запитів, хибно визнаних аномальними. При дослідженні ІСР було зазначено, що результати протилежні до застосування ФІС. Таким чином, при комбінації методів було зазначено, що показники помилок зменшилися, і стали розподілятися більш рівномірно. Графічне подання результатів показано на рис. 3.11, 3.12.

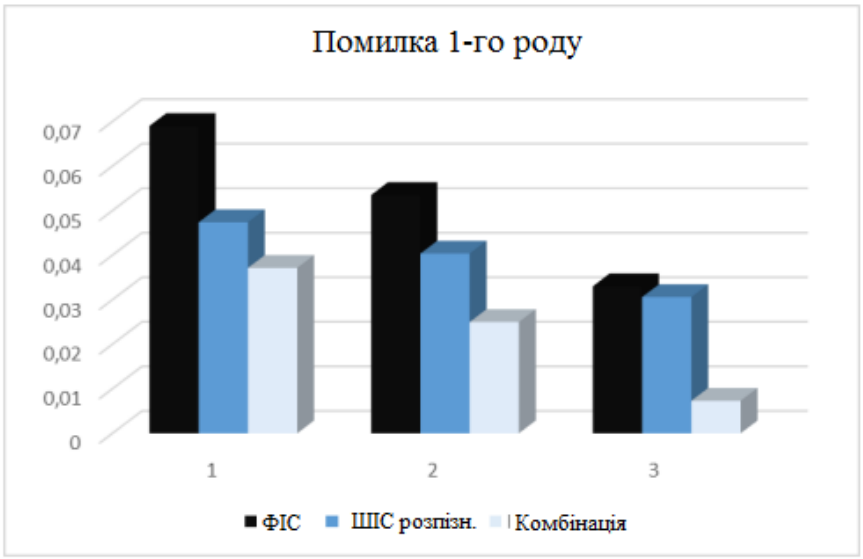

Рисунок 3.11 – Порівняння показників помилки 1 роду

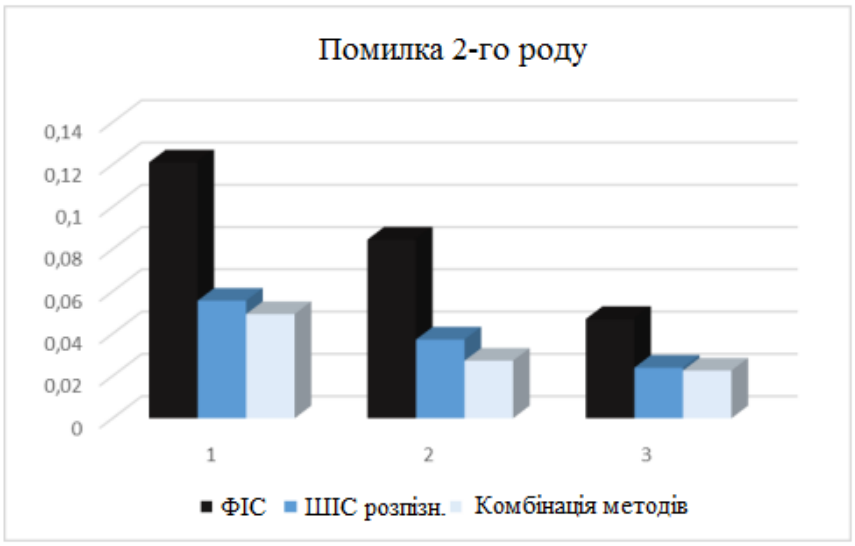

Рисунок 3.12 – Порівняння показників помилки 2 роду

У ході досліджень було зазначено, що застосування комбінації методів ФІС та ІСР дає найкращі результати. При застосуванні такого підходу підвищується ефективність виявлення та знижуються показники хибних класифікацій. Також для отримання хороших показників необхідно навчати системи на добре підібраній вибірці, що містить різні типи запитів, і оптимально підбирати параметри систем.

Результати перехресної перевірки у межах тестування комбінації методів наведено у табл. 3.10.

Таблиця 3.10 – Перехресна перевірка комбінації методів

Параметри	DDoS	Dos	Web attack	Patator	PortScan
Кількість правильних спрацьовувань (R)	37160	36770	34455	41610	45055
Кількість помилкових спрацьовувань (L)	12840	13230	15545	8390	4945
Аномальні запити (RA)	37160	36770	34455	41610	45055
Невірно аномальні запити (LA)	12840	13230	15545	8390	4945
Рівень виявлення (Detect)	0,7432	0,7354	0,6891	0,8322	0,9011
Рівень хибних спрацьовувань (FDetect)	0,2568	0,2646	0,3109	0,1678	0,0989

Виходячи з отриманих результатів та порівняння їх з результатами дослідження кожного методу окремо зауважимо, що застосування комбінації методів у більшості випадків дає результат вище значення ФІС, але нижче за показник ІСР. Трохи гіршим виявився результат під час перевірки DoS -атак, а збільшення показника було досягнуто під час перевірки атак типу PortScan (рис. 3.13).

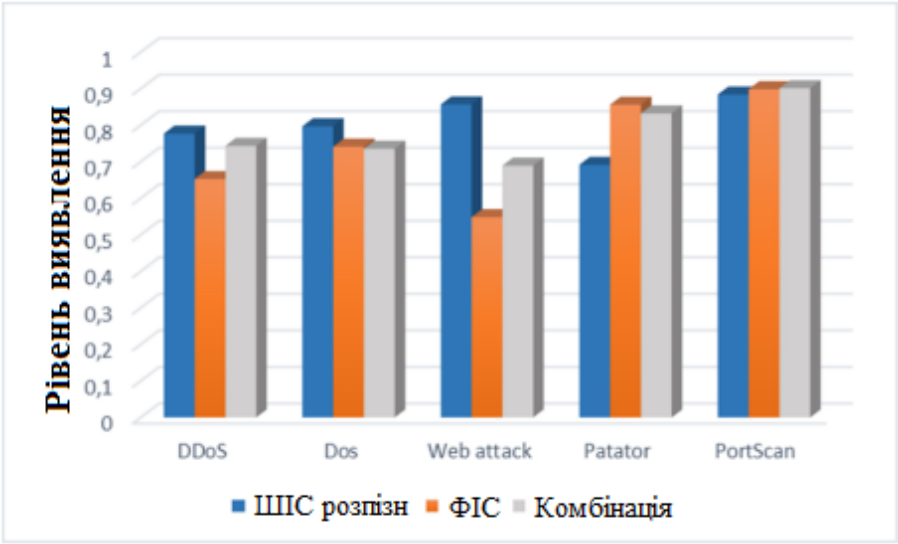

Рисунок 3.13 – Порівняння результатів перехресної перевірки

3.5 Висновки до третього розділу

У цьому розділі було проведено аналіз ефективності функціонування СВВ. Описано загальні параметри оцінки системи. Кожна методика аналізу запитів тестувалася шляхом проведення експериментів, скерованих на оцінку різних характеристик функціонування системи.

Результати експериментів розробленої СВВ дозволяють стверджувати, що ШІС можна ефективно застосовувати для детектування різного роду мережевих вторгнень. Проте для покращення роботи СВВ потрібно подальше удосконалення.

4 ОХОРОНА ПРАЦІ ТА БЕЗПЕКА В НАДЗВИЧАЙНИХ СИТУАЦІЯХ

4.1. Охорона праці

Метою кваліфікаційної роботи магістра є дослідження дослідження та створення СВВ з урахуванням архітектури ШІС. Оскільки, проведення робіт з розробки та використання системи передбачає використання комп’ютерної техніки, зокрема ПК та периферійних пристроїв, то обов’язковим є дотримання вимог з охорони праці і техніки безпеки.

Для ефективної і безпечної роботи колективу працівників з створення СВВ з урахуванням архітектури ШІС, в тому числі і фахівців зі створення СВВ, необхідно організувати безпечні умови праці. При цьому керівник організації несе безпосередню відповідальність за порушення нормативно-правових актів з охорони праці [34]. Окрім цього, на робочих місцях працівників необхідно забезпечити дотримання вимог, затверджених Наказом Мінсоцполітики від 14.02.2018 за № 207 «Про затвердження Вимог щодо безпеки та захисту здоров’я працівників під час роботи з екранними пристроями». Згідно Вимог приміщення, де розміщені робочі місця операторів, крім приміщень, у яких розміщені робочі місця операторів великих ЕОМ загального призначення (сервер), мають бути оснащені системою автоматичної пожежної сигналізації відповідно до цих вимог;

– переліку однотипних за призначенням об'єктів, які підлягають обладнанню автоматичними установками пожежогасіння та пожежної сигналізації, затвердженого наказом Міністерства України з питань надзвичайних ситуацій та у справах захисту населення від наслідків Чорнобильської катастрофи від 22.08.2005 N 161, зареєстрованого в Міністерстві юстиції України 05.09.2005 за N 990/11270 (НАПБ Б.06.004-2005);

– Державних будівельних норм "Інженерне обладнання будинків і споруд. Пожежна автоматика будинків і споруд", затверджених наказом Держбуду України від 28.10.98 N 247 (далі - ДБН В.2.5-56:2014, з димовими пожежними сповіщувачами та переносними вуглекислотними вогнегасниками.

В інших приміщеннях допускається встановлювати теплові пожежні сповіщувачі. Приміщення, де розміщені робочі місця операторів, мають бути оснащені вогнегасниками, кількість яких визначається згідно з вимогами ДСТУ 4297:2004 «Пожежна техніка. Технічне обслуговування вогнегасників». Загальні технічні вимоги і з урахуванням граничнодопустимих концентрацій вогнегасної рідини відповідно до вимог НАПБ А.01.001-2014. Приміщення, в яких розміщуються робочі місця операторів сервера загального призначення, обладнуються системою автоматичної пожежної сигналізації та засобами пожежогасіння відповідно до вимог ДБН В.2.5-56:2014, ДБН В.2.5-56:2010, НАПБ А.01.001-2014 і вимог нормативно-технічної та експлуатаційної документації виробника. Проходи до засобів пожежогасіння мають бути вільними.

Лінія електромережі для живлення комп'ютера та периферійних пристроїв повинні бути виконаними як окрема групова трипровідна мережа шляхом прокладання фазового, нульового робочого та нульового захисного провідників. Нульовий захисний провідник використовується для заземлення (занулення) електроприймачів. Не допускається використовувати нульовий робочий провідник як нульовий захисний провідник. Нульовий захисний провідник прокладається від стійки групового розподільного щита, розподільного пункту до розеток електроживлення. Не допускається підключати на щиті до одного контактного затискача нульовий робочий та нульовий захисний провідники.

Площа перерізу нульового робочого та нульового захисного провідника в груповій трипровідній мережі має бути не менше площі перерізу фазового провідника. Усі провідники мають відповідати номінальним параметрам мережі та навантаження, умовам навколишнього середовища, умовам розподілу провідників, температурному режиму та типам апаратури захисту, вимогам НПАОП 40.1-1.01-97.

У приміщенні, де одночасно експлуатуються понад п'ять комп'ютерів, на помітному, доступному місці встановлюється аварійний резервний вимикач, який може повністю вимкнути електричне живлення приміщення, крім освітлення. Комп'ютери повинні підключатися до електромережі тільки за допомогою справних штепсельних з'єднань і електророзеток заводського

виготовлення.

У штепсельних з'єднаннях та електророзетках, крім контактів фазового та нульового робочого провідників, мають бути спеціальні контакти для підключення нульового захисного провідника. Їхня конструкція має бути такою, щоб приєднання нульового захисного провідника відбувалося раніше, ніж приєднання фазового та нульового робочого провідників. Порядок роз'єднання при відключенні має бути зворотним. Не допускається підключати комп'ютери до звичайної двопровідної електромережі, в тому числі – з використанням перехідних пристроїв. Електромережі штепсельних з'єднань та електророзеток для живлення комп'ютерної техніки повинні бути виконаними за магістральною схемою, по 3-6 з'єднань або електророзеток в одному колі. Штепсельні з'єднання та електророзетки для напруги 12 В та 42 В за своєю конструкцією мають відрізнятися від штепсельних з'єднань для напруги 127 В та 220 В. Штепсельні з'єднання та електророзетки, розраховані на напругу 12 В та 42 В, мають візуально (за кольором) відрізнятися від кольору штепсельних з'єднань, розрахованих на напругу 127 В та 220 В.

При підвищенні ефективності контролю доступу в приміщення, де для забезпечення безпеки мешканців, співробітників і збереження майна використовуються ДС, важливим, з точки зору охорони праці, є забезпечення достатньої величини природного та штучного освітлення, які визначені у НПАОП 0.00-7.15-18. Організація робочого місця фахівця із дослідження методів та програмно-апаратних засобів оптимізаційних процесів на основі ГА повинна забезпечувати відповідність усіх елементів робочого місця та їх розташування ергономічним вимогам ДСТУ 8604:2015 «Дизайн і ергономіка. Робоче місце для виконання робіт у положенні сидячи. Загальні ергономічні вимоги». Відстань від екрана до ока фахівців, які працюють за комп'ютером визначається згідно з вимогами ДСанПіН 3.3.2.007-98.

Розміщення принтера або іншого пристрою введення-виведення інформації на робочому місці має забезпечувати добру видимість екрана комп'ютера, зручність ручного керування пристроєм введення-виведення інформації в зоні досяжності моторного поля згідно з вимогами ДСанПіН

3.3.2.007-98.

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 фахівців з дослідження та розробки СВВ з урахуванням архітектури ШІС.

4.2. Функціонування державної системи спостереження, збирання, оброблення та аналізу інформації про стан довкілля під час надзвичайних ситуацій мирного та воєнного часу

Моніторинг довкілля – це система спостереження, збирання та аналізу інформації про ситуацію, що може скластись під час надзвичайних ситуацій мирного та воєнного часу. Також це система спостереження за визначеними об'єктами, явищами та процесами з метою оперативного оцінювання їх стану, виявлення результатів впливу на них зовнішніх чинників та прийняття відповідних управлінських рішень (ДСТУ 3891:2013) (див. ДСТУ 7295:2013).

Моніторинг потенційно небезпечних об'єктів це спостереження, контролювання за зміною параметрів технологічних режимів з метою збирання, збереження, передавання та аналізування інформації щодо поточного стану потенційно небезпечних об'єктів, наявності та кількості порушень вимог безпеки, відпрацювання рекомендацій щодо проведення 98 робіт із запобігання та ліквідування техногенних надзвичайних ситуацій та їх наслідків (ДСТУ 7295:2013).

Моніторинг джерел надзвичайних ситуацій це система спостереження за об'єктами, які можуть бути джерелами надзвичайних ситуацій, що має на меті виявлення небезпеки, збирання, узагальнення та аналізування оперативної інформації стосовно стану об'єктів моніторингу та розроблення науково-обґрунтованих рекомендацій щодо проведення заходів із запобігання та ліквідування надзвичайних ситуацій (ДСТУ 7295:2013).

Моніторинг довкілля – це систематичні спостереження і контролювання,

які проводять регулярно, за єдиною програмою для оцінювання стану довкілля, аналізування процесів, які відбуваються в ньому і своєчасне виявлення тенденцій його змінювання (ДСТУ 7295:2013).

Моніторинг надзвичайних ситуацій (НС) – система спостереження за об'єктами, які можуть бути джерелами надзвичайних ситуацій, що має на меті виявлення небезпеки, збирання, узагальнення та аналізування оперативної інформації щодо об'єктів моніторингу та розроблення науково обґрунтованих рекомендацій щодо проведення заходів із запобігання та ліквідування НС [35].

Моніторинг небезпечних явищ та процесів це система спостереження та контролювання за розвитком небезпечних та стихійних природних явищ і процесів, чинниками, які спричинюють їх формування та розвиток, аналізування, збереження та передавання інформації щодо виявлення тенденцій їх змінювання, розроблення комплексу заходів щодо запобігання природним надзвичайним ситуаціям та ліквідування їх наслідків. Небезпечні природні явища і процеси підрозділяють на геофізичні, геологічні, гідрологічні, метеорологічні, медико-біологічні та пожежі в природних екосистемах (ДСТУ 7295:2013).

Моніторинг пожеж в екосистемах це спостереження, контролювання, збирання, аналізування, збереження та передавання інформації щодо 99 пожежної небезпеки в природних екосистемах (умов погоди, стану горючих матеріалів, інших пожежонебезпечних чинників), з метою своєчасного планування та здійснення заходів щодо запобігання виникненню і ліквідування пожеж та їх наслідків (ДСТУ 7295:2013).

Моніторинг радіаційної безпеки це спостереження і контролювання рівня радіоактивного забруднення місцевості, повітря, води, продовольства, об'єктів господарювання, дозових навантажень на населення з метою прийняття оперативних рішень щодо запобігання виникненню та ліквідування надзвичайних ситуацій та їх наслідків (ДСТУ 7295:2013).

Моніторинг хімічної небезпеки це спостереження, контролювання, збирання, аналізування, збереження та передавання інформації щодо визначення ступеня і характеру хімічного забруднення довкілля, санітарногігієнічний нагляд за дотриманням установлених нормативів з метою виявлення джерела

надходження небезпечних хімічних речовин, запобігання виникненню та ліквідування надзвичайних ситуацій та їх наслідків (ДСТУ 7295:2013).

Збір та аналіз інформації про стан довкілля під час мирного та воєнного стану дає можливість приймати оперативні рішення для адекватного реагування на ситуацію [35].

4.3. Висновки до розділу

В цьому розділі проаналізовано важливі питання охорони праці та безпеки в надзвичайних ситуаціях, висвітлено питання функціонування державної системи спостереження, збирання, оброблення та аналізу інформації про стан довкілля під час надзвичайних ситуацій мирного та воєнного час.

ВИСНОВКИ

У результаті виконаної роботи було реалізовано дослідницький прототип СВВ, в основі інтелектуального компонента якої використовувалися методи ШІС. Вибір ШІС обумовлений тим, що вони включають переваги генетичних алгоритмів і нейронних мереж.

До позитивних властивостей ШІС можна віднести самоорганізацію, адаптивність, динамічні донавчання та гнучку масштабованість. У ході роботи були відзначені такі недоліки: складність навчання; велика кількість констант, які потрібно задати на початку роботи алгоритму (рівень клонування, рівень мутації, граничне значення афінності тощо); нетривіальність операцій мутації та клонування.

Як методики виявлення АЗ були обрані ФІС та ІСР. Запропоновані підходи після реалізації були інтегровані в компонент аналізу прототипу СВВ. При проведенні досліджень було відзначено, що показники виявлення ФІС нижчі, ніж показники ІСР. Після аналізу отриманих результатів було запропоновано поєднати роботу двох підходів. Це дозволило підвищити якість детектування атак до 96,47%. Для отримання високого показника виявлення необхідно підібрати для кожного підходу навчальну вибірку, яка містить велику кількість запитів різного типу. Також на результати впливає налаштування параметрів системи. Створення імунної мережі з великої кількості клітин є складним завданням та потребує великої кількості обчислювальних ресурсів.

Прототип СВВ було створено як клієнт серверного додатку. Клієнтська програма містить користувальницький інтерфейс, за допомогою якого клієнт може керувати процесом аналізу трафіку. У серверній частині знаходяться основні компоненти збору та аналізу трафіку, а також інтерфейс для роботи та налаштування імунних систем. Навчання системи та генерація трафіку здійснювалася на основі НД CICIDS 2017.

В ході проведення експериментального дослідження отриманого прототипу СВВ було встановлено, що архітектура ШІС може успішно застосовуватися для виявлення мережевих вторгнень. Однак необхідне подальше

вдосконалення одержаного продукту, для оптимізації роботи та збільшення продуктивності. Також залишається відкритим питання про розширення архітектури СВВ для роботи в інших напрямках та на різних машинах.

ПЕРЕЛІК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. DIGITAL 2023:GLOBAL OVERVIEW REPORT [Електронний ресурс] - Режим доступу: https://datareportal.com/reports/digital-2023-global-overview-report (дата звернення: 14.11.2023).

2. Литвиненко В. І. Методи та засоби гібридних штучних імунних систем в задачах інтелектуального аналізу даних. – Дис… докт.техн.н. – Львів, 2010. 36 с.

3. Корабльов М. М. Гібридні методи і моделі обробки нечіткої інформації на основі штучних імунних систем: автореф. дис. ... д-ра техн. наук: 05.13.23 Харків, 2012. 38 с.

4. Черник О.А. Системи виявлення вторгнень // Інформаційні моделі, системи та технології: Праці XI наук.-техн. конф. (Тернопіль, ТНТУ ім. І. Пулюя, 13-14 грудня 2023 р.) – Тернопіль, 2023. – С. 126.

5. Top 10 Common Types of Network Security Attacks Explained [Електронний ресурс] - Режим доступу: https://cisomag.com/top-10-common-types-of-network-security-attacks-explained/ (дата звернення: 14.11.2023).

6. IDS – що це таке? Система виявлення вторгнень (IDS) як працює? [Електронний ресурс] - Режим доступу: https://poradumo.com.ua/49510-ids-sho-ce-take-sistema-viiavlennia-vtorgnen-ids-iak-pracuye/ (дата звернення: 16.11.2023).

7. The History of Intrusion Detection Systems [Електронний ресурс] - Режим доступу: https://www.threatstack.com/blog/thehistory-of-intrusion-detection-systems-ids-part-1 (дата звернення: 14.11.2023).

8. Classification of intrusion detection systems [[Електронний ресурс] - Режим доступу: https://www.academia.edu/11395235/CLASSIFICATION_OF_INTRUSION_DETECTION_SYSTEMS (дата звернення: 14.11.2023).

9. Коробейнікова Т., Цар О. Аналіз сучасних відкритих систем виявлення та запобігання вторгнень. – Grail of Science, (27), 2023. с. 317–325.

10. Аналіз сучасних систем виявлення та запобігання вторгнень в інформаційно-телекомунікаційних системах [Електронний ресурс] - Режим доступу: https://ela.kpi.ua/bitstream/123456789/17609/1/meshkov.pdf (дата

звернення: 14.11.2023).

11. Ямпольський, Л. С. Нейротехнології та нейрокомп'ютерні системи / Л. С. Ямпольський, О. І. Лісовиченко, В. В. Олійник; НТУУ «КПІ». – Київ : Дорадо-друк, 2016. – 631 с.

12. Методи аналізу та моделювання безпеки розподілених інформаційних систем: навч. посіб. / В.В. Литвинов, В.В. Казимир, І.В. Стеценко та ін. – Чернігів: Чернігівський національний технологічний університет, 2016. – 254 с.

13. Зоріна Т.І. Системи виявлення і запобігання атак в комп'ютерних мережах / Т.І. Зоріна // Вісник східноукраїнського національного університету імені Володимира Даля. – 2013. – № 15 (204) ч.1. – С. 48 – 54.

14. Сучасна імунологія (курс лекцій) / І.А.Іонов, Т.Є.Комісова, О.М. Сукач, О.О. Катеринич. – ПП Петров В.В. , 2017 .– 107 с.

15. Dasgupta D. and F. A. Gonzalez. An Immunogenetic Approach to Intrusion Detection, CS Technical Report (No. CS-01-001), The University of Memphis. May, 2001.

16. Dasgupta. Immunity-Based Intrusion Detection Systems: A General Framework In the proceedings of the 22nd National Information Systems Security Conference (NISSC), October 18-21, 1999.

17. Dasgupta D. An Overview of Artificial Immune Systems and Their Applications. Chapter 1 in the book entitled Artificial Immune Systems and Their Applications, Publisher: Springer-Verlag, Inc., pp 3-23, January 1999.

18. Clever Algorithms: Nature-Inspired Programming Recipes [Електронний ресурс] - Режим доступу: http://www.cleveralgorithms.com/nature-inspired/ index.html (дата звертання 02.12.2023).

19. Литвиненко В.І. Побудова штучних імунних систем // Наукові праці. Комп'ютерні технології . – 2010. – Вип. 121. – Т.134. – С. 166 – 178.

20. Basic Immune Inspired Algorithms [Електронний ресурс] - Режим доступу: http://www.artificial-immune-systems.org/algorithms.shtml (дата звертання 02.12.2023).

21. Castro De, L.N. & Timmis, J.I. Artificial Immune Systems: A New Computational Intelligence Approach, London: Springer-Verlag 2000), September, –

357 p.

22. Intruction Detection Evaluation Dataset (CICIDS 2017) [Електронний ресурс] - Режим доступу: https://www.unb.ca/cic/datasets/ids-2017.html (дата звертання: 02.12.23).

23. I. Sharafaldin, A. H. Lashkari ,A. A. Ghorbani “Towarding Generating a New Instruction Detection Dataset and Intrusion Traffic Characterization” // 4-th International Conference on Information Security and Privacy (ICISSP), Portugal. – 2018. – С.108-116.

24. ROC-крива [Електронний ресурс] - Режим доступу: https://uk.wikipedia.org/wiki/ROC-крива (дата звертання: 03.12.23).

25. Помилки першого і другого роду [Електронний ресурс] - Режим доступу: https://uk.wikipedia.org/wiki/ Помилки_першого_і_другого_роду (дата звертання: 03.12.23).

26. Read M., Andrews P. S., Timmis J. An Introduction to Artificial Immune Systems, Handbook of Natural Computing, G. Rozenberg, T. Bäck, и J. N. Kok, Ред. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, pp. 1575– 1597.

27. Chelly Z. and Elouedi Z. A survey of the dendritic cell algorithm, Knowl Inf Syst, vol. 48, issue. 3, pp. 505–535, sept. 2016.

28. Malim M. R., Halim F. A. IMMUNOLOGY AND ARTIFICIAL IMMUNE SYSTEMS, Int. J. Artif. Intell. Tools, t. 21, ed. 06, p. 1250031, dec. 2012.

29. Cohen I. R. Real and artificial immune systems: computing the state of the body, Nat Rev Immunol, vol. 7, issue. 7, pp. 569–574, jul. 2007.

30. McEwan C., Hart E. Representation in the (artificial) immune system, J. Math. Model. Algorithms, vol. 8(2), pp. 125-149, 2009.

31. Mendao M., Timmis J., Andrews P.S., M. Davies The immune system in pieces: Computational lessons from degeneracy in the immune system, in Proc. Foundations of Computational Intelligence (FOCI 2007), 2007, pp. 394-400.

32. Katsikis, Peter D., Stephen P. Schoenberger, and Bali Pulendran, eds. Crossroads between Innate and Adaptive Immunity. Boston, MA: Springer US, 2007.

33. Korablev N. M., Ivaschenko G. S. Parallel immune algorithm of short-term

forecasting based on model of clonal selection, Radio Electronics, Computer Science, Control, vol. 0, issue. 2, nov. 2014.

34. Зеркалов Д.В. Безпека життєдіяльності та основи охорони праці. Навч. посібник. К.: «Основа». 2016. 267 с.

35. Сакевич В.Ф., Поліщук О.В. Цивільна оборона. Теоретичні основи. Навч. посібник. Вінниця : ВНТУ, 2009. 136 с.

ДОДАТКИ

ДОДАТОК А
Тези конференції

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ

ТЕРНОПІЛЬСЬКИЙ НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ ІМЕНІ ІВАНА ПУЛЮЯ

М А Т Е Р І А Л И

XI НАУКОВО-ТЕХНІЧНОЇ КОНФЕРЕНЦІЇ

«ІНФОРМАЦІЙНІ МОДЕЛІ, СИСТЕМИ ТА ТЕХНОЛОГІЇ»

13-14 грудня 2023 року

ТЕРНОПІЛЬ
2023

УДК 004.056
О.А. Черник
Тернопільський національний технічний університет імені Івана Пулюя, Україна

СИСТЕМИ ВИЯВЛЕННЯ ВТОРГНЕНЬ

O.A.Chernyk

INTRUSION DETECTION SYSTEMS

Система виявлення вторгнень СВВ (IDS) відстежує мережевий трафік на предмет аномального впливу і повідомляє користувачів про виявлення аномального трафіку. З кінця 1980-х років адміністратори корпоративних мереж стали включати СВВ до інструментів, що забезпечують безпеку мережі, і застосовували їх у роботі. Однак при функціонуванні були виявлені недоліки використання систем, такі як неможливість виявити атаки нульового дня, оскільки трафік порівнювався тільки зі списком відомих сигнатур, та використання великої кількості ресурсів для постійного сканування. Вирішення проблеми виявлення нових атак прийшло з впровадженням нового методу – виявлення аномалій (ВА). Він був заснований на виявленні нетипових поведінок у мережі та забезпечував виявлення аномальних ситуацій [1].

В даний час найбільший інтерес становлять СВВ, що працюють на хмарних обчисленнях. На підставі даних компанії Thread Stack, IDS є одними з технологій, що найбільше продаються в галузі забезпечення безпеки.

СВВ відрізняються одна від одної, в основному, за місцем збору інформації, механізмом виявлення і за швидкістю реагування. На основі даних критеріїв наведемо класифікацію СВВ [1].

За місцем збору інформації IDS поділяються на хостові (HIDS) - представник OSEEC, та мережеві (NIDS) - Cisco Secure IDS, Dragon Enterasys. Перші контролюють безпеку комп'ютера чи системи від внутрішніх та зовнішніх атак. Внутрішні – аномальні види поведінки програм, спроба доступу до заборонених ресурсів. Зовнішні – аномальна активність або вміст пакетів, виявлені під час аналізу мережевих взаємодій. До мінусів таких систем можна віднести збір даних на рівні ядра та пасивність системи. Другі відстежують мережевий трафік, і при ВА, які повідомляють про це адміністраторів.

За швидкістю реагування IDS: динамічні системи - ІКС UTM+, Рубікон (працюють у реальному часі) та статичні системи - Hummingbird (аналізують мережу, перевіряють зміни мережевих сервісів). Дані IDS як реального часу відстежують трафік на наявність аномальних ознак. Статичні здійснюють аналіз мережі, перевіряють зміни мережевих сервісів. Надають дані про атаку та допомагають усунути заподіяну шкоду.

За механізмом ВА IDS: сигнатурні – Suricat, та ВА - Snort, Bro-IDS. У першому типі систем кожен пакет мережного трафіку порівнюється з шаблонами атак, які у базі даних атак. До переваг даного методу відносяться простота використання та низький показник хибних спрацьовувань при хорошому підборі сигнатури атак. Основними недоліками є необхідність збирання якісної бази сигнатур атак та складність детектування раніше невідомих або не внесених до бази атак. Другий тип IDS, засновані на ВА, аналізує дії, що відбуваються в мережі. Стандартна поведінка в мережі визначається адміністратором або за допомогою навчального набору даних при розробці системи. Дії, які не вписуються у рамки стандартної поведінки, вважаються аномальними. Аналіз трафіку щодо аномалій дає можливість виявляти атаки, невідомі системі раніше. До недоліків такого типу систем належать завдання правил визначення аномалій та сильна залежність ефективності роботи від них.

Література

1. Коробейнікова Т., Цар О. Аналіз сучасних відкритих систем виявлення та запобігання вторгнень. – Grail of Science, (27), 2023. с. 317–325.

ДОДАТОК Б
Фрагмент програмного коду

Клієнтський додаток:

```
//Клас зберігання ідентифікаційних даних запиту
class Request {
            public string datetime;
            public string sourseip;
            public string sourceport;
            public string destination;
            public string protocol;
            public string result;
         public Request(string date, string ip, string port,
                              string dest, string protoc, string res)
            {
                datetime = date;
                sourceport = port;
                sourseip = ip;
                destination = dest;
                protocol = protoc;
                result = res;
            }
        }

//Функція надсилання повідомлення на сервер
static Request SendMessageFromSocket(string host, int port, string
            message, DataGridView dataGrid)
{
      // Буфер для вхідних даних
      byte [] bytes = New byte [1024];

      // З'єднуємося з віддаленим пристроєм
      // Встановлюємо віддалену точку для сокету
      IPAddress ipAddr;
      if (!IPAddress.TryParse(host, out ipAddr))
          ipAddr = Dns.GetHostEntry(host).AddressList[0];
      IPEndPoint ipEndPoint = новий IPEndPoint(ipAddr, port);
      Socket sender = новий Socket(ipAddr.AddressFamily,
                        SocketType.Stream, ProtocolType.Tcp);

      // З'єднуємо сокет із віддаленою точкою
      sender.Connect(ipEndPoint);
      Console.WriteLine("Сокет з'єднується з {0}",
```

```
                    sender.RemoteEndPoint.ToString()); byte[]
 msg = Encoding.UTF8.GetBytes(message);

// Надсилаємо дані через сокет
int bytesSent = sender.Send(msg)

 // Отримуємо відповідь від сервера
 int bytesRec = sender.Receive(bytes);
 string answer = Encoding.UTF8.GetString
 (bytes, 0, bytesRec);
 Console.WriteLine("\nВідповідь від сервера:
 {0}\n\n", Encoding.UTF8.GetString(bytes, 0,
 bytesRec));
 string[] mess = message.Split(',');
 Request res = new Request(DateTime.Now.ToString(),
 mess[1], mess[2], mess[3], mess[5], answer);
 InsertUpdateDB(res);

 // Звільняємо сокет
 sender.Shutdown(SocketShutdown.Both);
 sender.Close(); return res;
 запуску аналізу трафіку
 private async void button1_Click(object sender,
 EventArgs e)

 //Повідомлення про хід виконання завдання
 Progress<string> progess = new
 Progress<string>(text => this.label2.Text = text);

 //Токен скасування
 _tokenSource = new CancellationTokenSource();
 CancellationToken cancelToken = _tokenSource.Token;

 //Запускаємо завдання
 try {
 label2.Text = "Починаємо..."; this.label2.Text =
 await Task.Run(() => DoSomething(cancelToken,
 progess), cancelToken);
 }
 catch (OperationCanceledException)
 {
 this.label2.Text = "Завдання скасовано.";
 }
```

```
catch (Exception ex)
{
this.label2.Text = $"У задачі сталася помилка:
{ ex. Message } ";
}
}

//Функція формування повідомлення до сервера
private string DoSomething(CancellationToken
cancelToken, IProgress<string> progess)
{
bool work = true; int i = 0; while(work)
try
{

//повідомляємо про прогрес
progess.Report($"Request: {i}");

//Затримка між етапами с.
cancelToken.WaitHandle.WaitOne(TimeSpan.FromSeconds
(Random.Next (1, 6)));
string req = GenerateTraffic.GenerateRequest();
Request answ =
SendMessageFromSocket("192.168.131.1",
11000, req, dataGridView1);
i++;

//Виняток у разі натискання на кнопку скасування
cancelToken.ThrowIfCancellationRequested();
}
catch (OperationCanceledException)
{
work = false;
}
}
return "Готово!";
}
. . . . . .
```