

НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ
імені ІГОРЯ СІКОРСЬКОГО»

Факультет інформатики та обчислювальної техніки

Кафедра технічної кібернетики

До захисту допущено:

Завідувач кафедри

_____ Ігор ПАРХОМЕЙ

«__» _____ 2020 р.

Дипломний проєкт

на здобуття ступеня бакалавра

**за освітньо-професійною програмою «Програмне забезпечення
інтелектуальних та робототехнічних систем»
спеціальність 121 «Інженерія програмного забезпечення»
на тему: «Інтелектуальний додаток для ідентифікації аномалій
мережевого трафіку»**

Виконав (-ла):

студент (-ка) IV курсу, групи ІТ-62

Позднякова Олена Юріївна _____

Керівник:

доцент, к.т.н.

Батрак Євгеній Олександрович _____

Консультант з нормоконтролю:

доцент, к.т.н., доц.,

Пасько Віктор Петрович _____

Рецензент:

стар. викладач кафедри АУТС

Хмелюк Марина Сергіївна _____

Засвідчую, що у цьому дипломному
проєкті немає запозичень з праць інших
авторів без відповідних посилань.

Студент (-ка) _____

Київ – 2020 року

Національний технічний університет України
«Київський політехнічний інститут імені Ігоря Сікорського»
Факультет інформатики та обчислювальної техніки
Кафедра технічної кібернетики

Рівень вищої освіти – перший (бакалаврський)

Спеціальність 121 «Інженерія програмного забезпечення»

Освітньо-професійна програма «Програмне забезпечення інтелектуальних та робототехнічних систем»

ЗАТВЕРДЖУЮ

Завідувач кафедри

_____ Ігор ПАРХОМЕЙ

«___» _____ 2020 р.

ЗАВДАННЯ
на дипломний проєкт студенту
Поздняковій Олені Юріївні

1. Тема проєкту «Інтелектуальний додаток для ідентифікації аномалій мережевого трафіку», керівник проєкту Батрак Євгеній Олександрович, доцент, вчене звання, затверджені наказом по університету від «07» травня 2020р. № 1081-с

2. Термін подання студентом проєкту 29.05.2020

3. Вихідні дані до проєкту – інтелектуальний додаток, що ідентифікує аномалії мережевого трафіку

4. Зміст пояснювальної записки

Вступ

Аналіз предметної області

Аналіз вимог та визначення технічних засобів

Архітектура нейромережі

Реалізація та тестування

Висновки

5. Перелік графічного матеріалу (із зазначенням обов'язкових креслеників, плакатів, презентацій тощо): Повна архітектура нейронної мережі;

6. Консультанти розділів проєкту

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		завдання видав	завдання прийняв
Перевірка на співпадіння	доцент Лісовиченко О. І.		
Норм. контроль	доцент Пасько В. П.		

7. Дата видачі завдання « 01 » жовтня 2019р.

Календарний план

№ з/п	Назва етапів виконання дипломного проєкту	Термін виконання етапів проєкту	Примітка
1.	Огляд існуючих варіантів реалізації та постановка задачі	13.04.2020	
2.	Дослідження основних проблем для реалізації процесу	20.04.2020	
3.	Вибір та обґрунтування оптимальності технічних рішень та технічної реалізації	27.04.2020	
4.	Розробка функціональних вимог	04.05.2020	
5.	Архітектура/структура проєкту	11.05.2020	
6.	Оформлення документації дипломної роботи	18.05.2020	
7.	Норм. контроль	27.05.20 – 01.06.20	
8.	Перевірка на співпадіння	01.06.20 – 08.06.20	
9.	Попередній захист		
10.	Захист		

Студентка

Олена ПОЗДНЯКОВА

Керівник

Євгеній БАТРАК

АНОТАЦІЯ

У роботі розглянуто проблему ідентифікації мережевих аномалій, показані основні особливості існуючих рішень в цій сфері, їх переваги та недоліки. Було розглянуто кілька сучасних технологій для реалізації та обрано найвідповідніші до вимог швидкого, безпечного та зрозумілого додатку.

Розроблено інтелектуальний додаток, який визначає нормальний мережевий трафік та аномалії певних типів. Цей додаток для людей, які працюють із безпекою даних та мереж, та відстежують загальні порушення кібербезпеки. Додаток був розроблений у середовищі MATLAB, з використанням бібліотеки Neural Network Toolbox. Для навчання нейронної мережі був використаний протокол NetFlow, дані для навчання - вільно надані адміністрацією KDD Cup '99. Також, для реалізації додатку було вирішено використовувати MLP-архітектуру. Щоб користуватися додатком, користувачу необхідно мати ПК із тактовою частотою не менш 1100 МГц.

Ключові слова: нейронна мережа, кібербезпека, кіберпорушення, MATLAB, NetFlow.

Розмір пояснювальної записки – 55 аркушів, містить 16 ілюстрацій, 5 таблиць, 4 додатка.

ABSTRACT

The work examined the problem of identifying network anomalies, showing the main features of existing solutions in this field, their advantages and disadvantages. Several modern technologies for implementation were considered and the most appropriate for the requirements of a fast, safe and understandable application were selected.

An intelligent application has been developed that determines normal network traffic and certain types of anomalies. This application is for people who work with data and network security and monitor common cyber security breaches. The application was developed in MATLAB using the Neural Network Toolbox. The NetFlow protocol was used to train the neural network; training data was freely provided by the administration of the KDD Cup '99. Also, to implement the application, it was decided to use the MLP architecture. To use the application, the user needs to have a PC with a clock frequency of at least 1100 MHz.

Keywords: neural network, cybersecurity, cyber violation, MATLAB, NetFlow.

The explanatory note contains 55 pages, 16 illustrations, 5 tables, 4 applications.

**Пояснювальна записка
до дипломного проєкту
на тему: «Інтелектуальний додаток для
ідентифікації аномалій мережевого трафіку»**

ЗМІСТ

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ	8
ВСТУП	9
РОЗДІЛ 1 АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ	11
1.1 Нейронні мережі	11
1.2. Кіберпорушення і кібербезпека	13
1.3 Огляд схожих рішень	17
1.3.1. WireShark	18
1.3.2 Tcpdump	19
1.3.3 SolarWinds NetFlow Traffic Analyzer	20
1.4 Деякі відомі види кібер-злочинів	21
1.4.1 SNMP	21
1.4.2 Smurf	24
1.4.3 IPSweep	24
ВИСНОВКИ ДО РОЗДІЛУ 1	25
РОЗДІЛ 2 АНАЛІЗ ВИМОГ ТА ВИЗНАЧЕННЯ ТЕХНІЧНИХ ЗАСОБІВ	27
2.1 Вимоги	27
2.1.1 Вхідні дані	27
2.2 Визначення середовища розробки	29
2.2.1 MATLAB	29
2.2.2 Amygdala	31
2.2.3 NeuralBase	32
ВИСНОВКИ ДО РОЗДІЛУ 2	33
РОЗДІЛ 3 АРХІТЕКТУРА НЕЙРОМЕРЕЖІ	35
3.1 Про нейромережі	35
3.2 Архітектурні особливості нейромереж	36
3.3 Методика і типи навчання	38
3.4 FFNN і MLP	39
3.5 Процедура навчання MLP:	40
3.6 Особливості вибору і підготовки даних	42
ВИСНОВКИ ДО РОЗДІЛУ 3	44
РОЗДІЛ 4 РЕАЛІЗАЦІЯ ТА ТЕСТУВАННЯ	45
4.1 Формування концепції	45
4.2 Реалізація інтерфейсу	46
4.3 Навчання та тестування нейронної мережі	46
ВИСНОВКИ ДО РОЗДІЛУ 4	52
ВИСНОВКИ	53
ПЕРЕЛІК ПОСИЛАНЬ	55
ДОДАТКИ	56

					ІТ-62.27.1081.01 ПЗ			
Змн	Лист	Прізвище	Підпис	Дата	Інтелектуальний додаток для ідентифікації аномалій мережевого трафіку	Літ.	Лист	Листів
Розроб.		Позднякова О.Ю.						
Перевір.		Батрак Є.О.					7	55
Норм.		Пасько В.П.				КПІ ім. Ігоря Сікорського Каф. ТК Гр. ІТ-62		
Затв.		Пархомей І.Р.						

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ

PBC – розподілена обчислювальна система

TCP(Transmission Control Protocol) – протокол керування передачею

UDP(User Datagram Protocol) – протокол датаграм користувача

IDS(Intrusion Detection System) – система знаходження вторгнень

NTA(Network traffic analysis) – аналіз мережевого трафіку

MLP(Multiple layer perceptron) – багатошаровий перцептон

SOC(System-on-a-Chip) – однокришталева система

SNMP (Simple Network Management Protocol) – простий протокол
мережевого керування

NMS (Network Management System) – система керування локальною
мережею

FFNN (Feedforward neural network) – нейромережа з прямим зв'язком

					ІТ-62.27.1081.01 ПЗ	Арк.
						8
Змн.	Арк.	№ докум.	Підпис	Дат.		

ВСТУП

Безпека комп'ютерних систем та мереж – це проблема, яка існує навколо нас усіх на протязі значного періоду. Спочатку це стало проблемою лише урядів та військових, потім це зачепило численні фінансові установи (банки, страхові компанії, інвестиційні компанії і т.п.). Тепер, коли інтернет та мобільні технології активно використовуються в бізнесі та повсякденному житті, це стосується практично кожного. Кібербезпека важлива, оскільки урядові, військові, корпоративні, фінансові та медичні організації збирають, обробляють та зберігають небачений обсяг даних на комп'ютерах та інших пристроях. Значна частина цих даних може становити конфіденційну інформацію, чи то інтелектуальна власність, фінансові дані, особиста інформація, чи то інші типи даних, для яких несанкціонований доступ чи вплив може мати негативні наслідки. Організації передають конфіденційну інформацію по мережах та на інші пристрої в процесі ведення бізнесу. У міру зростання обсягів та витонченості кібератак компаніям та організаціям, особливо тим, кому належить забезпечити захист інформації, що стосується національної безпеки, охорони здоров'я чи фінансових записів, потрібно вжити заходів для захисту своєї чутливої інформації про бізнес та персонал.

Наприклад, людина покладається на питання безпеки навіть тоді, коли хоче щось придбати на “e-Bay”. Хоча на перший погляд це звичайний приклад, який здається досить простим для вирішення, фактично, він охоплює багато аспектів комп'ютерної безпеки, вимагає поглиблених знань математики, наполегливої роботи та своєчасного планування. Студенти коледжів та університетів, IT-фахівці, архітектори програмного забезпечення, розробники, адміністратори мереж та безпеки й багато інших зобов'язані бути добре ознайомлені з безпекою, конфіденційністю, захистом та подібними темами та проблемами. Це також важлива сфера для керівників компаній, оскільки безпека

					IT-62.27.1081.01 ПЗ	Арк.
						9
Змн.	Арк.	№ докум.	Підпис	Дат.		

в різних сферах, включаючи захист життєво важливої інфраструктури, є першочерговою для сучасних компаній та організацій.

Безпека інформаційної системи залежить від випадкового чи навмисного втручання, що чинять шкоду користувачам чи власникам інформації, в цілому залежить, від доступності (можливості за короткий часовий проміжок отримати потрібну інформаційну послугу); цілісності (актуальності й несуперечності інформації, її захищеності від пошкоджень і несанкціонованої деформації); конфіденціальності (захисту від несанкціонованого зчитування). Сьогодні інформаційна система являє собою складну систему, яка містить в собі безліч компонентів різного рівня автономності, які зв'язані один з одним і здійснюють обмін даними. Майже кожний компонент здатен піддаватися сторонньому втручанням або вийти з ладу.

Сучасна обчислювальна техніка та відкриття в області обробки даних надають змогу скористатися алгоритмами інтелектуального аналізу які направлені на ідентифікацію та класифікацію аномалій. Нейронні мережі є надзвичайно гнучкими системами, що здатні швидко перевчитись і аналізувати неоднорідні дані.

Нейронна мережа є принципово новим алгоритмом, що можна навчити. Результатом дипломного проєкту є програмний продукт, суть якого полягає в аналізі NetFlow-даних та класифікації аномалій, що, можливо, були знайдені.

					ІТ-62.27.1081.01 ПЗ	Арк.
						10
Змн.	Арк.	№ докум.	Підпис	Дат.		

РОЗДІЛ 1 АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ

1.1 Нейронні мережі

Нейронна мережа – це комп'ютерна модель, чия багаторівнева структура нагадує мережеву структуру нейронів у мозку людини з шарами зв'язаних вузлів. Вона може вчитись, тому її можна навчити розпізнавати шаблони, класифікувати дані й прогнозувати майбутні події.

Нейронна мережа розбиває введення на шари абстракції. На багатьох прикладах можна навчитись розпізнавати закономірності в мові або зображеннях, наприклад, як це робить людський мозок. Його поведінка визначається тим, як його окремі елементи пов'язані, і силою чи вагою цих зв'язків. Ці ваги автоматично коригуються під час навчання у відповідності з заданим правилом навчання, поки нейронна мережа не виконає бажану задачу правильно.

Нейромережі найбільш підходящі для розпізнавання образів, щоб ідентифікувати і класифікувати об'єкти або сигнали в мовних, зорових і керівних системах. Також, вони можуть бути використані для виконання прогнозування і моделювання часових рядів. Ті мережі, які працюють на двох або трьох шарах з'єднаних нейронних шарів, відомі як дрібні нейронні мережі. Мережі глибокого навчання можуть мати багато шарів, навіть сотні. Обидва є технікою машинного навчання, які навчаються безпосередньо з вхідних даних.

Глибинне навчання привертає багато уваги, і для цього є всі підстави. Це досягнення результатів, які раніше були неможливі. Найбільше, воно підходить для складних додатків ідентифікації, таких як розпізнавання облич, переклад тексту і розпізнавання голосу. Це також ключова технологія, яка використовується в сучасних системах і задачах допомоги водію, включно з класифікацією смуг руху й розпізнавання дорожних знаків.

					ІТ-62.27.1081.01 ПЗ	Арк.
						11
Змн.	Арк.	№ докум.	Підпис	Дат.		

Нейронна мережа об'єднує декілька шарів обробки, використовуючи прості елементи, які працюють паралельно і натхненні біологічною нервовою системою. Вона складається з вхідного шару, одного або декількох прихованих шарів і вихідного шару. Шари зв'язані між собою через вузли або нейрони, причому кожен шар використовує вихідні дані попереднього шару в якості вхідних даних.

Загальні методи машинного навчання для проєктування додатків нейронних мереж включають в себе навчання під наглядом і без нього, класифікацію, регресію, розпізнавання образів і кластеризацію.

Контрольовані нейронні мережі навчаються виробляти бажані результати у відповідь на вибірккові вхідні дані, що робить їх особливо відповідними для моделювання і керування динамічними системами, класифікації зашумлених даних і прогнозування майбутніх подій.

Штучним нейроном називається простий елемент, що спочатку обчислює зважену суму вхідних величин:

$$V = \sum_{i=1}^N W_i * x_i = \bar{W} * \bar{X}, \quad (1.1)$$

де V – зважена сума вхідних величин, N – розмірність простору вхідних сигналів, W_i у зваженій сумі зазвичай називають синаптичними коефіцієнтами або вагами, x_i – сигнал навчання, \bar{W} – кінцева вага, \bar{X} – кінцевий сигнал навчений нейромережею. Опісля, сума яку було отримано, порівнюється з пороговою величиною W_0 , а наступна, починає діяти нелінійна функція активації f . Саму ж зважену суму – W , назовемо потенціалом нейрона i . Вихідний сигнал тоді виглядатиме так $f(V)$. Величину порогового бар'єру можна розглядати додатковий ваговий коефіцієнт при постійному вхідному сигналі. В такому разі говориться про розширений вхідний простір: нейрон з N -мірним входом має $N+1$ ваговий коефіцієнт [1].

					ІТ-62.27.1081.01 ПЗ	Арк.
						12
Змн.	Арк.	№ докум.	Підпис	Дат.		

Якщо ввести в рівняння порогову величину W_0 , то воно буде виглядати так:

$$V = \sum_{i=1}^N W * x + W_0 \quad (1.2)$$

Залежно від способу перетворення сигналу і характеру активації виникають різні види нейронних структур. Існують детерміновані нейрони, коли активізуються, функція однозначно обчислює вихід по входу, і ймовірні нейрони, стан яких в момент t є випадкова функція потенціалу та стану в момент $t-1$ [1].

1.2. Кіберпорушення і кібербезпека

Віддаленна мережева атака – це інформаційний руйнівний вплив на розподілену обчислювальну систему (РВС), який втілюється по каналах зв'язку.

Через те, що здійснення віддаленої атаки досить важко виявити, а провести її відносно просто, такий різновид несанкціонованих дій посідає перше місце за ступенем небезпеки. За характеристикою впливу атаки бувають пасивні й активні. До перших відносяться такі, що не мають прямого впливу на роботу РВС, але здатні змінити її політику безпеки. Саме через нестачу прямого впливу на систему, таку атаку виявити складно. Активна дія на РВС – це така, яке безпосередньо впливає на саму роботу системи, порушує її працездатність, змінює конфігурацію і т. п. При активному типі атаки в системі виникають певні зміни, у той час коли при пасивному впливі не лишається видимих слідів.

При будь-якій атаці головна ціль, як правило, – це отримання стороннього доступу до інформації. Добування інформації буває двох видів: перехоплення і викривлення. При перехопленні отримують інформацію без можливості її змінити. Викривлення або підміна даних веде до порушення їх цілісності. Таким чином, за метою впливу мережеві атаки можна поділити на такі, що порушують коректну дієздатність системи, цілісність ресурсів або ж їх конфіденційність.

					ІТ-62.27.1081.01 ПЗ	Арк.
						13
Змн.	Арк.	№ докум.	Підпис	Дат.		

Мережеві та інформаційні технології розвиваються і змінюються настільки швидко, що інертні захисні механізми, такі як відділення доступу, системи ідентифікації не здатні, у багатьох випадках, забезпечити ефективний захист. Необхідні саме гнучкі методи, які дозволяють в короткий термін знаходити і попереджувати порушення безпеки. Одним з таких методів, що дозволяють відслідковувати злодіяння, які не ідентифікуються завдяки звичайним моделям контролю доступу, є процес знаходження втручань.

Знаходження втручань – це процес розпізнавання і реагування на сумнівний вплив, направлений на мережеві чи обчислювальні ресурси. Ефективність технології сильно залежить від того, які методи аналізу одержаної інформації застосовують. В теперішній час поряд зі статистичними методами використовується багато нових методик, таких як експертні системи і нейронні мережі.

Статистичний аналіз. Цей підхід має дві основні переваги:

- апарат математичної статистики;
- адаптація до поведінки суб'єкта.

З самого початку використання даного методу визначаються профілі для кожного суб'єкта досліджуваної системи. Будь-яке відхилення профілю від еталону розглядається як несанкціонована діяльність. Статистичні методи універсальні, так як не вимагають знань про можливі атаки і вразливості системи. Однак при їх використанні можуть виникати деякі труднощі, пов'язані, наприклад, з тим, що їх можна “навчити” сприймати несанкціоновані дії як звичайні. Тому поряд зі статистичним аналізом застосовуються додаткові методики:

– експертні системи. Цей метод для виявлення втручань є досить поширеним. За його використання, інформація про атаки формується у вигляді правил, які, найчастіше, записують як послідовність дій або в формі сигнатури. Якщо виконується будь-яке з тих правил, то тут же укладається рішення, щодо наявності несанкціонованої діяльності. Одне з головних переваг

					ІТ-62.27.1081.01 ПЗ	Арк.
						14
Змн.	Арк.	№ докум.	Підпис	Дат.		

цього методу – майже цілковита відсутність хибних тривог. Для того, аби експертні системи завжди залишалися актуальними, необхідно постійно оновлювати використовувані бази даних постійно. Недолік такого методу полягає в неможливості відображення невідомих втручань. Навіть, якщо атаку з бази даних трохи змінять, то це вже може бути серйозною перешкодою задля її знаходження;

– нейронні мережі. Через те, що хакерів і варіантів атак стає з кожним днем все більше, експертні системи, навіть в умовах постійного оновлення баз даних не можуть дати гарантії точної ідентифікації кожного можливого вторгнення. Як один із способів подолання даної проблеми використовуються нейронні мережі. Нейронна мережа аналізує інформацію і надає можливість дати оцінку, наскільки дані узгоджуються з розпізнаваними їй властивостями. Для цього її навчають точної ідентифікації, щодо підбраної вибірки прикладів з певної предметної області. Реагування нейронної мережі піддається аналізу, після чого систему налаштовують так, щоб досягти задовільних результатів. У темпі того, як нейромережа виконує аналіз даних, вона отримує додатковий досвід.

Одне з важливих переваг нейромереж – їх здатність враховувати властивості атак, ідентифікуючи елементи, які не є схожими на вивчені.

Через те, що перераховані методи пошуку атак мають свої недоліки, їх, як правило, використовують в сукупності для забезпечення більш надійного захисту.

Щоб забезпечити безпеку комп'ютера, потрібно знати, які мережеві атаки можуть йому загрожувати. Всі відомі загрози можна умовно розділити на три групи:

– сканування портів – дані загрози самі по собі атакою не є, але, як правило, їй передують, так як це один із способів отримати інформацію про віддалений комп'ютер. Суть даного способу полягає в скануванні портів, які використовуються мережевими сервісами на потрібному комп'ютері для

					ІТ-62.27.1081.01 ПЗ	Арк.
						15
Змн.	Арк.	№ докум.	Підпис	Дат.		

виявлення їх стану. Такий процес допомагає зрозуміти, які атаки на дану систему можуть бути вдалими, а які ні. Більш того, сканування дає зловмиснику необхідні відомості про операційну систему, що дозволяє підібрати ще більш відповідні типи атак;

– DoS-атаки – вони ще відомі, як “відмова в обслуговуванні”. Це такі атаки, в результаті дії яких система приходить в нестабільний стан або ж повністю перестає функціонувати. Їх наслідки можуть включати в себе пошкодження або руйнування інформаційних ресурсів і неможливість їх використання. DoS-атаки бувають двох типів. Перший тип – коли комп'ютеру-жертві відправляються спеціально сформовані пакети, які призводять до перезавантаження системи або її зупинки. Другий – комп'ютеру-жертві відправляється велика кількість пакетів в одиницю часу, він не справляється з їх обробкою. Наслідок – вичерпання ресурсів системи;

– атаки-вторгнення. Їх мета – “захоплення” системи. Такий тип атак найнебезпечніший, тому що при успішному їх виконанні зловмисник отримує максимально повну інформацію про систему. Атаки-вторгнення застосовуються в тих випадках, коли є необхідність в отриманні конфіденційних даних з віддаленого комп'ютера, такі як паролі та доступ до кредитних карток. Також метою таких атак може бути закріплення в системі для того, щоб згодом з метою зловмисника використовувати її обчислювальні ресурси. До цієї групи відноситься найбільша кількість атак.

Більш поширені види атак, які використовують мережеві сервіси операційної системи:

– атаки на переповнення буфера. Цей тип вразливостей в програмному забезпеченні виникає через відсутність або недостатній заходи контролю при роботі з масивами даних;

– атаки, засновані на помилках форматних рядків. Такий тип виникає через недостатній мірі контролю значень вхідних параметрів функцій форматного введення-виведення. У тому випадку, якщо така вразливість

					ІТ-62.27.1081.01 ПЗ	Арк.
						16
Змн.	Арк.	№ докум.	Підпис	Дат.		

знаходиться в програмному забезпеченні, то зловмисник може отримати абсолютний контроль над системою.

1.3 Огляд схожих рішень

Історично інспекцією трафіку займаються системи виявлення вторгнень (Intrusion Detection Systems, IDS). Однак вони функціонують тільки в режимі реального часу, тому з їх допомогою неможливо побачити атаки в минулому, просто довантажити нові сигнатури погроз і індикатори компрометації. Крім того, IDS-рішення зазвичай не застосовуються для виявлення небезпечної та підозрілої активності у внутрішньомережевих інформаційних потоках. У тих випадках, коли вони все ж використовуються для аналізу внутрішнього трафіку (переважно – за допомогою сигнатурних методів), в “сліпій зоні” залишаються приховані канали передачі даних, переміщення зловмисника по мережі і його закріплення в ній, несанкціонований доступ до мережесервісів (в тому числі – за допомогою легітимних інструментів) і всілякі нові техніки атакуючих. Крім того, просто виявити проблеми недостатньо. Необхідно оперативне розслідування інциденту і ефективне реагування. Зловмисники можуть діяти і закріплюватися в інфраструктурі жертви вельми швидко, тому важливо скоротити час між виявленням атаки і відгуком на неї.

Network Traffic Analysis (NTA). – нова категорія систем мережевої безпеки, які не тільки служать джерелами даних для центрів моніторингу (SOC), але і забезпечують потужні можливості ретроспективного пошуку, розслідування інцидентів, централізованого протидії шкідливої активності.

Системи аналізу мережевого трафіку (NTA) – призначені для перехоплення потоків даних і виявлення ознак складних, найчастіше цільових атак (APT). З їх допомогою можна проводити ретроспективне вивчення мережесервісів, виявляти і розслідувати операції зловмисників в інформаційній інфраструктурі підприємства, а також ефективно реагувати на відповідні події.

					ІТ-62.27.1081.01 ПЗ	Арк.
						17
Змн.	Арк.	№ докум.	Підпис	Дат.		

Рішення класу NTA відстежують трафік і комунікації всередині корпоративної мережі і застосовують різні техніки (поведінковий аналіз, машинне навчання) для того, щоб швидко виявити, проаналізувати і обробити загрози, які в іншому випадку залишалися б прихованими. NTA може застосовуватися в мережах будь-якого масштабу, так само як і будь-якої архітектури: локальної, хмарної, гібридної.

На сьогоднішній день на ринку представлений різноманітний вибір варіацій програмного забезпечення для аналізу мережевого трафіку. Цінова політика таких варіацій також відрізняється – від безкоштовних до рішень з досить дорогою корпоративною ліцензією.

1.3.1. Wireshark

Wireshark – це програма для аналізу мережевих протоколів з відкритим вихідним кодом, введена в експлуатацію Gerald Combs в 1998 році. Глобальна організація фахівців із мережевих технологій та розробників програмного забезпечення підтримує Wireshark і продовжує оновлювати нові мережеві технології і методи шифрування [2].

Цей додаток абсолютно безпечний та зрозумілий у використанні (рис.1.1). Урядові установи, корпорації, некомерційні та освітні установи використовують Wireshark для усунення неполадок і навчання.

Wireshark – інструмент для аналізу і аналізу пакетів. Він захоплює мережевий трафік в локальній мережі, а також зберігає ці дані для автономного аналізу. Wireshark захоплює мережевий трафік з Ethernet, Bluetooth, бездротових мереж (IEEE.802.11), Token Ring, Frame Relay і т. п.

Wireshark дозволяє фільтрувати журнал або до початку захоплення, або під час аналізу, так що надається можливість звузити і обнулити те, що необхідно в трасуванні мережі. Наприклад, можлива установка фільтру для перегляду трафіку TCP між двома IP-адресами. Також, можна встановити його тільки для відображення пакетів, відправлених з одного комп'ютера. Фільтри в Wireshark є

					IT-62.27.1081.01 ПЗ	Арк.
						18
Змн.	Арк.	№ докум.	Підпис	Дат.		

однією з основних причин, з якої він став стандартним інструментом для аналізу пакетів. Також, він відмінно працює з будь-яким із трьох найвідоміших сімейств операційних систем – Linux, Windows і Mac [2].

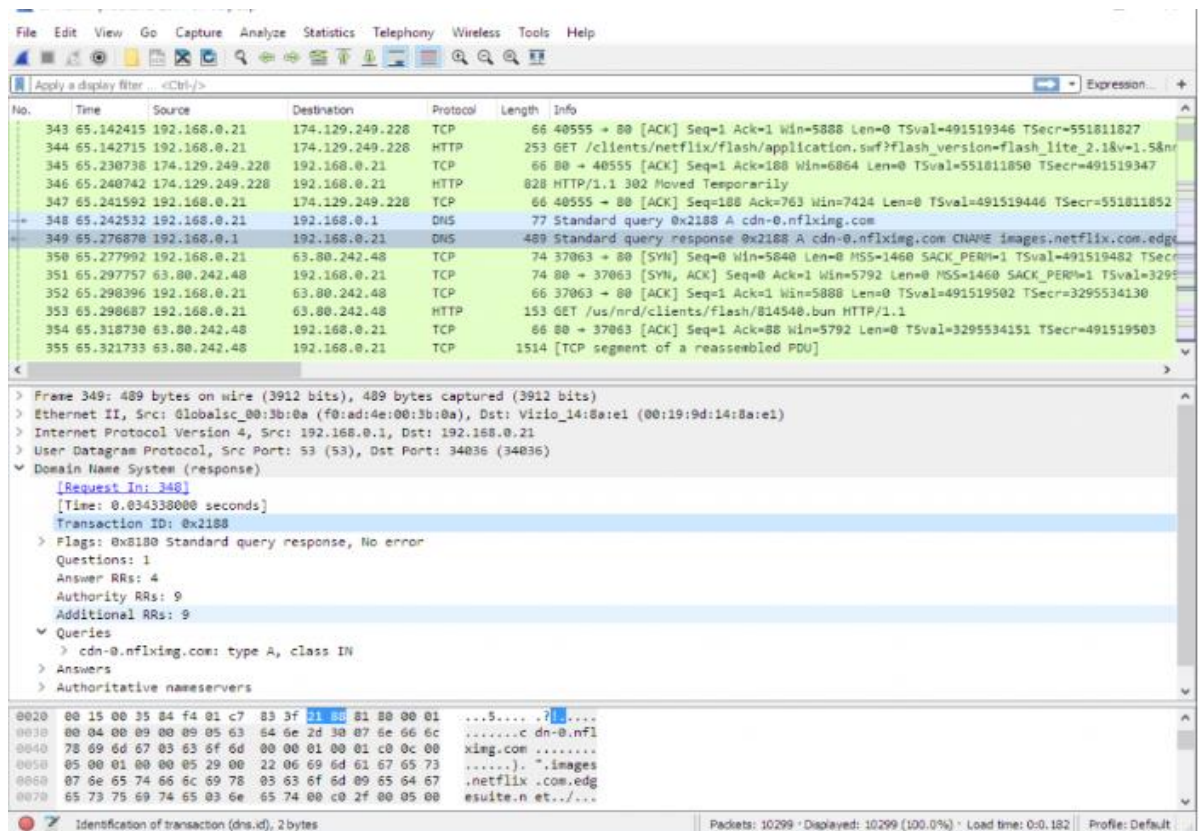


Рис. 1.1. WireShark

1.3.2 Tcpdump

Аналізатор трафіку Tcpdump виглядає як стародавній інструмент, і, насправді працює він так само (рис. 1.2). Не звертаючи увагу на те, що зі своєю роботою він справляється відмінно, при тому, що використовує для цього мінімальну кількість системних ресурсів, настільки, наскільки це взагалі можливо, багато кому з сучасних фахівців буде важко розібратися у великій кількості таблиць з даними.

Але інколи, в житті трапляються ситуації, в яких використання таких невибагливих і обрізаних відносно ресурсів рішень може бути корисним. У певних середовищах, або на ледве-ледве "живих" ПК, мінімалізм може несподівано виявитися єдиним допустимим варіантом [2].

```

suse1:~ # tcpdump
tcpdump: verbose output suppressed, use -v or -vv for full protocol decode
listening on eth0, link-type EN10MB (Ethernet), capture size 96 bytes
20:38:29.013544 IP text-lb.esams.wikimedia.org.http > 192.168.198.128.54543: FF
659455554:659455554(0) ack 3797162379 win 64240
20:38:29.013961 IP 192.168.198.128.54543 > text-lb.esams.wikimedia.org.http: F
1:1(0) ack 1 win 51120
20:38:29.014324 IP text-lb.esams.wikimedia.org.http > 192.168.198.128.54543: .
ack 2 win 64239
20:38:29.015350 IP 192.168.198.128.40566 > 192.168.198.2.domain: 16599+ PTR? 12
8.198.168.192.in-addr.arpa. (46)
20:38:29.043034 IP 192.168.198.2.domain > 192.168.198.128.40566: 16599 NXDomain
0/1/0 (95)
20:38:29.043533 IP 192.168.198.128.57315 > 192.168.198.2.domain: 34584+ PTR? 19
2.174.198.91.in-addr.arpa. (45)
20:38:29.101802 IP 192.168.198.2.domain > 192.168.198.128.57315: 34584 1/0/0 PT
R[[domain]
20:38:29.102332 IP 192.168.198.128.57083 > 192.168.198.2.domain: 41740+ PTR? 2.
198.168.192.in-addr.arpa. (44)
20:38:29.128455 IP 192.168.198.2.domain > 192.168.198.128.57083: 41740 NXDomain
0/1/0 (93)
20:38:29.177013 IP bits-lb.esams.wikimedia.org.http > 192.168.198.128.53928: FF
929274665:929274665(0) ack 3446487441 win 64240
20:38:29.177279 IP 192.168.198.128.53928 > bits-lb.esams.wikimedia.org.http: F
1:1(0) ack 1 win 58220
20:38:29.177618 IP bits-lb.esams.wikimedia.org.http > 192.168.198.128.53928: .
ack 2 win 64239
20:38:29.178046 IP 192.168.198.128.33843 > 192.168.198.2.domain: 3084+ PTR? 202

```

Рис. 1.2 Tcpdump

Спершу, програмне рішення tcpdump було розроблено для середовища *NIX, але в сьогодення, він також працює з декількома портами Windows. Він має всю базову функціональність, яку ви можете очікувати побачити в будь-якому аналізаторі трафіку – запис, захоплення і т.п., – але вимагати, щось набагато більшого від нього не варто [2].

1.3.3 SolarWinds NetFlow Traffic Analyzer

SolarWinds NetFlow Traffic Analyzer – це рішення для аналізу трафіку NetFlow та моніторингу смуги пропускання (рис. 1.3). Інструмент розроблений спеціально для аналізу даних трафіку NetFlow, а також записів потоків IPv4 і IPv6 і трафіку додатків. Користувачі також можуть візуально зіставляти невідповідності продуктивності і трафіку, відображаючи метрики поруч один з одним. SolarWinds NTA дозволяє збирати дані про мережеві потоках і перетворювати їх в легко читаються діаграми і таблиці. Візуальне відображення даних дозволяє швидко зрозуміти, як використовується корпоративна мережа, ким і які програми працюють [2].

					ІТ-62.27.1081.01 ПЗ	Арк.
						20
Змн.	Арк.	№ докум.	Підпис	Дат.		

SolarWinds NTA може бути об'єднаний з їх монітором продуктивності мережі для забезпечення комплексного вирішення для моніторингу та аналізу. Ці два продукти об'єднуються і забезпечують єдину консоль.

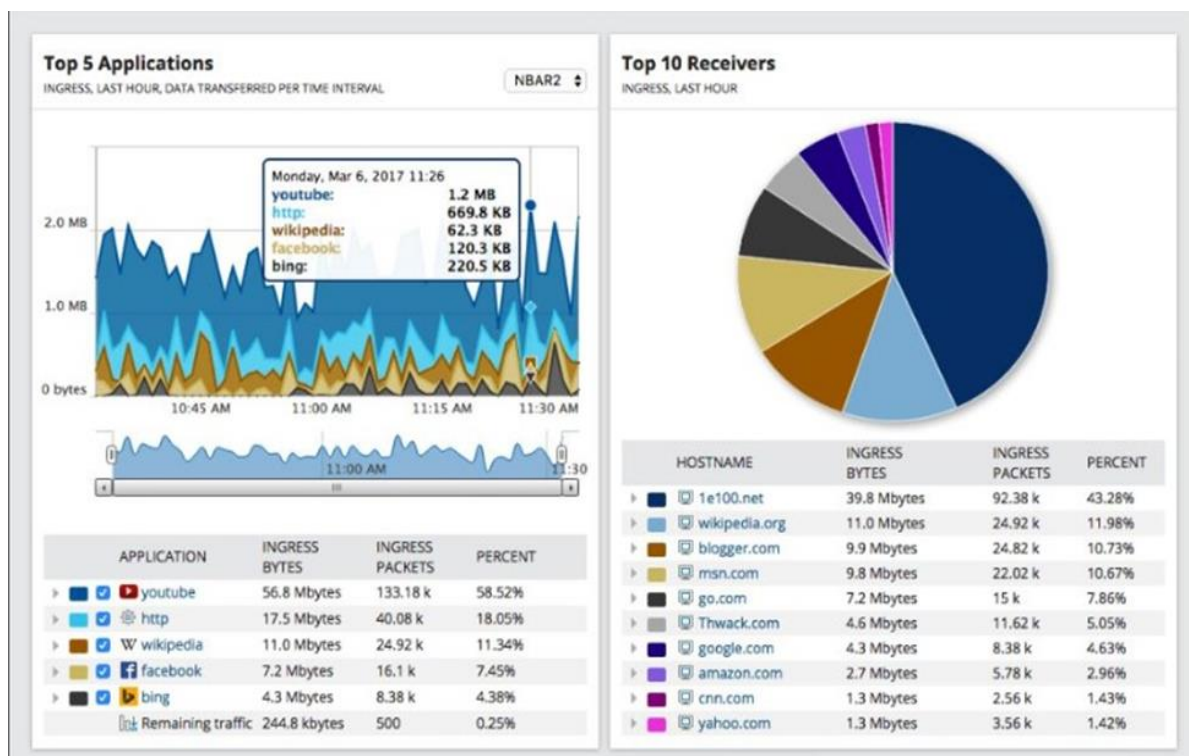


Рис. 1.3. SolarWinds NetFlow Traffic Analyzer

1.4 Деякі відомі види кібер-злочинів

1.4.1 SNMP

Ось уже більше десяти років мережеві адміністратори використовують протокол Simple Network Management Protocol для моніторингу мережевих пристроїв і управління ними. З моменту своєї появи в 1987 році SNMP став фактичним стандартом на мережеве управління; випущена вже третя версія цього протоколу [5].

Однак недавній звіт, підготовлений координаційним центром CERT Coordination Center, що спеціалізуються на питаннях комп'ютерної захисту, показав, що в реалізаціях SNMP є безліч недоліків, через які продукти більше ста виробників виявляються уразливими для атак.

При вмілому використанні цих недоліків хакер може отримати несанкціонований доступ з широкими привілеями, організувати DoS-атаку або зробити інші шкідливі дії.

Основний протокол зв'язку в SNMP – це UDP (User Datagram Protocol). Якщо агенти SNMP аналізують дані з порту UDP 161 при отриманні запитів від NMS, то NMS аналізує потік, який передається через порт UDP 162 для отримання асинхронних повідомлень. На рис. 1.4 показана спрощена архітектура SNMP-управління, де динамічний порт призначає операційна система. SNMP підтримує найпростішу аутентифікацію за допомогою імені спільноти, яка служить в якості пароля для отримання або зміни даних, суттєвих для задач управління [5].

Дослідники з фінського університету Оулу розробили тести, які дозволяють виявити безліч уразливих місць в реалізаціях SNMPv1. Тестові пакети, як правило, використовуються для аналізу протоколу і породжують повідомлення, які перевіряють виконання різних обмежень архітектури.

Дослідники виявили наступні вразливі місця SNMP:

- обробка повідомлень trap. Безліч недоліків було виявлено в тому, як різні станції NMS (network management station) декодують і обробляють trap-повідомлення;
- процедура розгляду заяв про. Тестування показало наявність похибок при декодуванні і обробці запитів SNMP різними агентами.

Ці вразливі місця виникли через незадовільну перевірку повідомлень SNMP в той момент, коли вони були отримані та оброблялися системою. Ці дефекти можуть привести до DOS-атак, уразливості формату рядка і переповнення буферів.

					ІТ-62.27.1081.01 ПЗ	Арк.
						22
Змн.	Арк.	№ докум.	Підпис	Дат.		

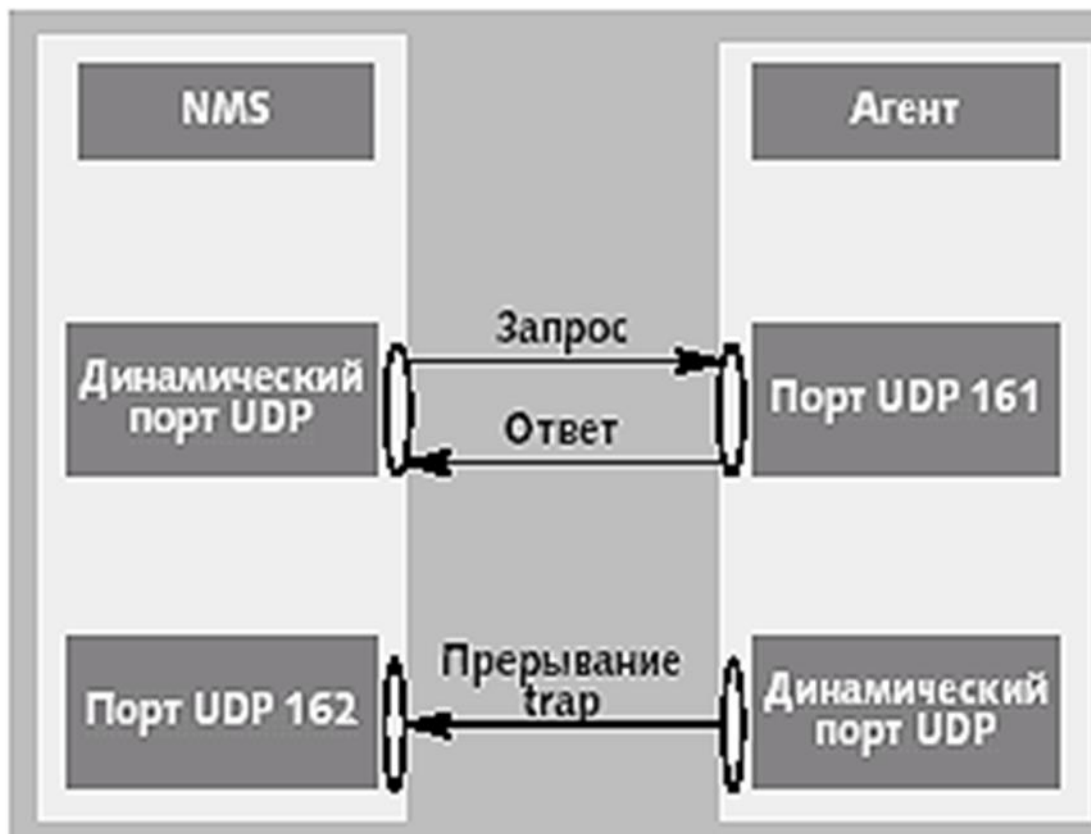


Рис. 1.4. Спрощена архітектура SNMP

Деякі з уразливих місць, виявлені Protos, не вимагають, щоб в повідомленнях SNMP було присутнє коректне ім'я спільноти, в силу чого використовувати такі вразливі місця дуже легко. Крім того, оскільки UDP – комунікаційний протокол, який не потребує встановлення з'єднання, агенти SNMP і NMS, що підтримують trap, приймають вхідні запити та повідомлення trap без будь-яких попередніх налаштувань сеансу. Більшість продуктів, орієнтованих на SNMP, випускаються з рядками спільноти за замовчуванням public (“загальнодоступний”) для доступу тільки на читання і private (“приватний”) для доступу з правами читання та запису. Рядок імені спільноти інтегрована в повідомлення SNMP і передається по мережі у вигляді звичайного тексту. Навіть якщо цей рядок вказаний коректно, вона вразлива для перехоплення пакетів [5].

Для блокування атак, що використовують ці дефекти захисту, контролю мережевого доступу також недостатньо, оскільки адреси відправників UDP

легко сфальсифікувати spoofing. Хакер може надіслати пакети, вказавши хибну адресу відправника, і вивести з ладу пристрій-одержувач. Більш того, деякі реалізації SNMP за замовчуванням дозволяють пересилати пакети SNMP широкомовно. Хакери можуть легко поширити широкомовні пакети, щоб вразити всю мережу цілком, навіть якщо їм не відома адреса конкретного пристрою й ім'я спільноти SNMP.

1.4.2 Smurf

Реалізується така атака шляхом відправлення ICMP ECHO запитів із підміненою адресою джерела (замість якого вказується адреса сервера-жертви) на широкомовну адресу маршрутизатора великої мережі. Задумка полягає в тому, що ICMP ECHO запит, який прийшов на широкомовну адресу, буде розісланий всім пристроям мережі, кожен з яких, у свою чергу, відповість на цей запит. Таким чином, утворюється потік ICMP ECHO відповідей, які направляються з сервера-жертви, який здатний частково або повністю вичерпати каналні й обчислювальні ресурси цільового сервера. В даний час атаки цього виду відбуваються інколи, тому що заводські налаштування більшості сучасних маршрутизаторів не передбачають обов'язкову розсилку кожному пристрою мережі пакетів, які прийшли на широкомовну адресу. Однак, в мережах все ще зустрічаються застарілі моделі мережевого обладнання, а також з неактуальною версією ПО, або некоректним настроюванням політик обробки пакетів, які можуть бути використані зловмисниками для реалізації таких атак [4].

Серед рекомендованих способів захисту можна назвати ретельне налаштування правил обробки пакетів, які прийшли на широкомовну адресу мережі, на маршрутизаторі.

1.4.3 IPSweep

Відбуваються тоді, коли хакер всередині мережі або за її межами видає себе за комп'ютер, якому можна довіряти. Хакер може робити це або

					ІТ-62.27.1081.01 ПЗ	Арк.
						24
Змн.	Арк.	№ докум.	Підпис	Дат.		

використовуючи IP-адресу, яка знаходиться в межах довірчих адрес для даної мережі, або авторизована зовнішня IP-адреса, с якої дозволений доступ до певних ресурсів мережі. IP spoofing часто є початковим етапом атак інших типів. Класичним прикладом є початок DoS-атаки з використанням підроблених адрес джерела з метою приховати реальні адреси хакера.

Зазвичай IP spoofing обмежується введенням небезпечних даних або команд в існуючий потік даних, який проходить між клієнтом і сервером, або при зв'язку рівноправних вузлів мережі. Щоб забезпечити двонаправлене з'єднання, хакер повинен змінити всі таблиці маршрутизації, щоб направити на підмінену IP-адресу. Іноді використовується інший хакерський підхід – просто не турбуватися з приводу отримання будь-якого відгуку від додатків. Якщо хакер намагається отримати критично важливий файл системи, відповіді додатків йому не важливі. Однак якщо хакер зможе змінити таблиці маршрутизації, щоб направити на підроблену IP-адресу, він зможе отримувати всі мережеві пакети, які направляються на підроблену адресу й відповідати саме так, як відповідав би довірений користувач.

ВИСНОВКИ ДО РОЗДІЛУ 1

Використання нейронних мереж заради аналізу трафіку є надзвичайно ефективним засобом для відстеження підозрілої активності. Правильно навчена мережа дозволяє виявити загрозу ще на етапі сканування портів, що дозволяє запобігти більш небезпечним несанкціонованим діям.

Шанси збільшити ефективність боротьби з атаками й несанкціонованим проникненням підвищується в десятки разів.

Звертаючи увагу на безліч існуючих на ринку продуктів, було виявлено, що вони працюють з даними мережі, підключеної в даний момент, або ж не мають комплексу для вибірки конкретних характеристик мережевого трафіку, що дозволило б визначити конкретні характеристики протоколу NetFlow, в яких один або група хакерів, залишили "сліди". Для кібер-фахівців було б зручно мати

					ІТ-62.27.1081.01 ПЗ	Арк.
						25
Змн.	Арк.	№ докум.	Підпис	Дат.		

в своєму розпорядженні інструмент, що дозволяє швидко і без зайвих дій перевірити конкретні характеристики NetFlow, в яких дії хакерів найчастіше виявляються. Доцільно створити програмний продукт, що працює не в режимі реального часу, адже більшість рішень або дуже дорогі (для корпоративного використання), або не надають зручного функціоналу для перевірки мережевого трафіку, вирваного з потоку реального часу, що може ускладнювати роботу з аналізу фахівцям по кібер-безпеки.

					ІТ-62.27.1081.01 ПЗ	Арк.
						26
Змн.	Арк.	№ докум.	Підпис	Дат.		

РОЗДІЛ 2 АНАЛІЗ ВИМОГ ТА ВИЗНАЧЕННЯ ТЕХНІЧНИХ ЗАСОБІВ

2.1 Вимоги

2.1.1 Вхідні дані

NetFlow – це протокол, розроблений Cisco Systems для запису всіх потоків IP-трафіку, що проходять через маршрутизатор або комутатор з підтримкою NetFlow. Він генерує статистику всередині цих пристроїв на рівні інтерфейсу і відправляє цю інформацію в записах потоків на основі UDP у зовнішній елемент, званий колектором – програмою, що працює на сервері, де статистика трафіку може зберігатися для аналізу балансування навантаження. NetFlow створює і відстежує інформацію про потоках або метадані для цих односпрямованих потоків IP-трафіку в кеш-пам'яті в пам'яті пристрою, що підтримує NetFlow. Ці метадані використовуються декількома мережевими адміністраторами, які використовують інструменти аналізу програмного забезпечення, щоб допомогти їм проаналізувати пропускну здатність мережі, втрату пакетів, перевантаження трафіку, виявити DDoS-атаки та інші випадки моніторингу мережі [7].

Було вирішено скористатись вільними даними, що були надані та поширені адміністрацією KDD Cup '99. Ці данні являють собою вибірку характеристик мережеских потоків, щоб збільшити продуктивність і швидкодію, а також усі символічні дані замінені числами. Для аналізу мережеских потоків було обрано 9 характеристик мережевого потоку:

- duration;
- protocol_type;
- service;
- src_bytes;
- dst_bytes;
- srv_count;
- dst_host_count;

					IT-62.27.1081.01 ПЗ	Арк.
						27
Змн.	Арк.	№ докум.	Підпис	Дат.		

- dst_host_srv_count;
- dst_host_same_srv_rate.

2.1.1 Вимоги

На основі проведеного аналізу нейронних мереж та кібербезпеки, огляду і аналізу існуючих рішень, та загальному огляді деяких способів атакування комп'ютерних мереж, були сформовані конкретні вимоги до майбутнього застосунка. На наступних трьох таблицях (табл. 2.1-2.3), вказані функціональні, нефункціональні та системні вимоги:

Таблиця 2.1. Функціональні вимоги

№	Вимога
F1	Визначення нормального трафіку
F2	Визначення аномалії типу snmpgetattack
F3	Визначення аномалії типу named
F4	Визначення аномалії типу xlock
F5	Визначення аномалії типу smurf
F6	Визначення аномалії типу ipsweep

Таблиця 2.2. Нефункціональні вимоги

№	Вимога
NF1	Мати графічний інтерфейс користувача
NF2	Неможливість отримати доступ до коефіцієнтів навченої нейромережі
NF3	Підтримуваними ОС мають бути: Windows 7, Windows 8, Windows 10
NF4	Бібліотеки, необхідні для стабільної роботи, мають бути встановленні разом із програмою одним інсталяційним пакетом

Таблиця 2.3. Системні вимоги

Процесор	Будь-який процесор із тактовою частотою не менш 1500 МГц
ОЗП	Ємність 2Гб та частота не менш 700 МГц
Зовнішні пристрої	Миша, клавіатура та монітор
Internet	Програма працює в режимі offline, мережевого підключення не вимагає

2.2 Визначення середовища розробки

2.2.1 MATLAB

MATLAB – це інтегрована технічна обчислювальна середа, яка поєднує в собі чисельні обчислення, просунуту графіку і візуалізацію, а також мову програмування високого рівня (рис.2.1). MATLAB використовується в різних областях, включаючи обробку сигналів і зображень, проєктування систем управління, фінанси, інженерні та медичні дослідження.

MATLAB і Neural Network Toolbox надають функції командного рядка і програми для створення, навчання і моделювання неглибоких нейронних мереж. Додатки полегшують розробку нейронних мереж для таких завдань, як класифікація, регресія (в тому числі регресія часових рядів) і кластеризація. Після створення мереж за допомогою цих інструментів можна автоматично генерувати код MATLAB.

Поліпшення здатності мережі узагальнювати допомагає запобігти переоснащенню, що є поширеною проблемою при проєктуванні нейронних мереж. Переоснащення відбувається, коли мережа запам'ятала навчальний набір, але не навчилася узагальнювати на нові входи. Переоснащення викликає відносно невелику помилку в навчальному наборі, але набагато більшу помилку, коли в мережу надходять нові дані.

					ІТ-62.27.1081.01 ПЗ	Арк.
						29
Змн.	Арк.	№ докум.	Підпис	Дат.		

Два рішення для поліпшення узагальнення містять у собі:

Регуляризація змінює функцію перфорації мережі (міра помилок, яку процес навчання зводить до мінімуму). Включаючи розміри ваг і зміщень, регуляризація створює мережу, яка добре працює з даними навчання і демонструє більш плавну поведінку при поданні нових даних.

При ранній зупинці використовуються два різних набори даних: навчальний набір для поновлення ваг і зміщень та набір перевірки для зупинки навчання, коли мережа починає бути відповідною до даних.

Більше мільйона інженерів та наукових співробітників з усього світу використовують MATLAB як мову технічних розрахунків. Якщо порівнювати MATLAB із традиційними мовами програмування (C / C ++, Java) то він, на відміну від них, дає змогу значно зменшити час вирішення звичних завдань, і значно спрощує розробку нових алгоритмів.

MATLAB є фундаментом усього сімейства MathWorks і є основним інструментом для вирішення різноманітних прикладних і наукових задач, у таких галузях як: розробка об'єктів управління та систем моделювання, проектування комунікативних систем, вимірювання сигналів і тестування, обробка сигналів і зображень, обчислювальної біології, фінансового моделювання та ін.

Ядро MATLAB дає змогу дуже легко працювати з матрицями комплексних, аналітичних та реальних типів даних, і структур даних, та таблиць пошуку.

MATLAB містить у собі функції лінійної алгебри, швидкі перетворення Фур'є, функції для роботи з поліномами, функції базової статистики та численних рішень диференціальних рівнянь; розширені математичні бібліотеки для Intel MKL.

					IT-62.27.1081.01 ПЗ	Арк.
						30
Змн.	Арк.	№ докум.	Підпис	Дат.		

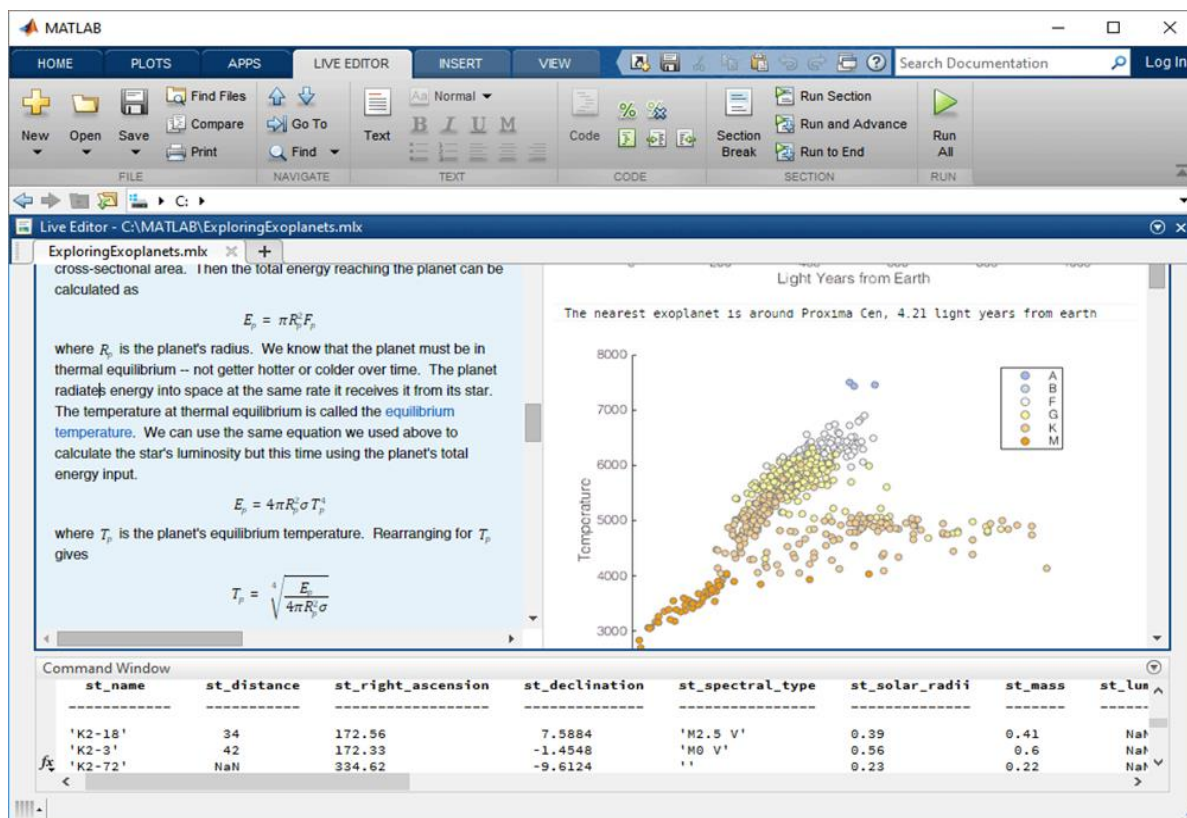


Рис. 2.1. Інтерфейс MATLAB

2.2.2 Amygdala

Amygdala – це програмне забезпечення з відкритим кодом для моделювання нейромереж. Це дуже специфічний вид "штучного мозку", який продемонстрував особливу роль в останніх наукових дослідженнях. SNN корисні в штучному інтелекті (AI) та розпізнаванні образів.

Мета проєкту Amygdala полягає в тому, щоб використовувати біологічно правильні моделі для створення мереж, здатних ефективно працювати в режимі реального часу; сильна увага приділяється підтримці багатоваріантної / розподіленої обробки, для збільшення масштабованості та потужності "мозку" Amygdala. Ще на своїх ранніх етапах Amygdala вже використовувалася для створення робочих нейронних мереж, а постійне вдосконалення спрямоване на підвищення ефективності та корисності Amygdala, а також на закладенні основ для ряду важливих додаткових функцій.

Amygdala в даний час реалізується як бібліотека C ++, придатна для використання програмістами. Майбутні вдосконалення розширить можливості використання Amygdala, зокрема, шляхом додавання графічного інтерфейсу користувача (GUI), придатного для не-програмістів. Також передбачено додаткові інтерфейси для мов програмування, відмінних від C ++, для розширення доступності та корисності Amygdala для більш ширшої бази розробників та додатків.

2.2.3 NeuralBase

NeuralBase – це бібліотека з відкритим вихідним кодом, що реалізує мережу Хопфілда і багат шарову нейронну мережу, які навчаються за алгоритмом зворотнього поширення. Бібліотека є інтелектуальною власністю компанії BaseGroup Labs – професійного постачальника програмних продуктів і рішень в області аналізу даних. Використання компонента є безкоштовним, але тільки для некомерційного використання. Компонент розроблений для Borland Delphi 6.0.

Бібліотека призначена переважно для інтеграції нейромереж в інформаційні системи, для збільшення системних аналітичних можливостей. Реалізація нейромереж як компонентів, і наявність відкритого коду, дозволяють дуже легко вбудовувати їх в інші програми. Об'єктно-орієнтований вигляд надає гнучкості, достатньо переписати декілька методів і ви можете отримати компонент, оптимізований під ваші завдання [6].

Існує три базових класи – TNeuron, TLayer, TNeuralNet. Всі інші – похідні від них. На рис. 2.2 представлена ієрархія класів, суцільними лініями показано успадкування (стрілкою вказаний нащадок), пунктиром позначено в яких класах вони використовуються.

TNeuron – базовий клас для нейронів, несе всю головну функціональність, має індексовані: властивість Weights, що являє собою вагові коефіцієнти,

					IT-62.27.1081.01 ПЗ	Арк.
						32
Змн.	Арк.	№ докум.	Підпис	Дат.		

властивість Output, яке є виходом нейрона і акумулятор, роль якого, виконує метод ComputeOut.

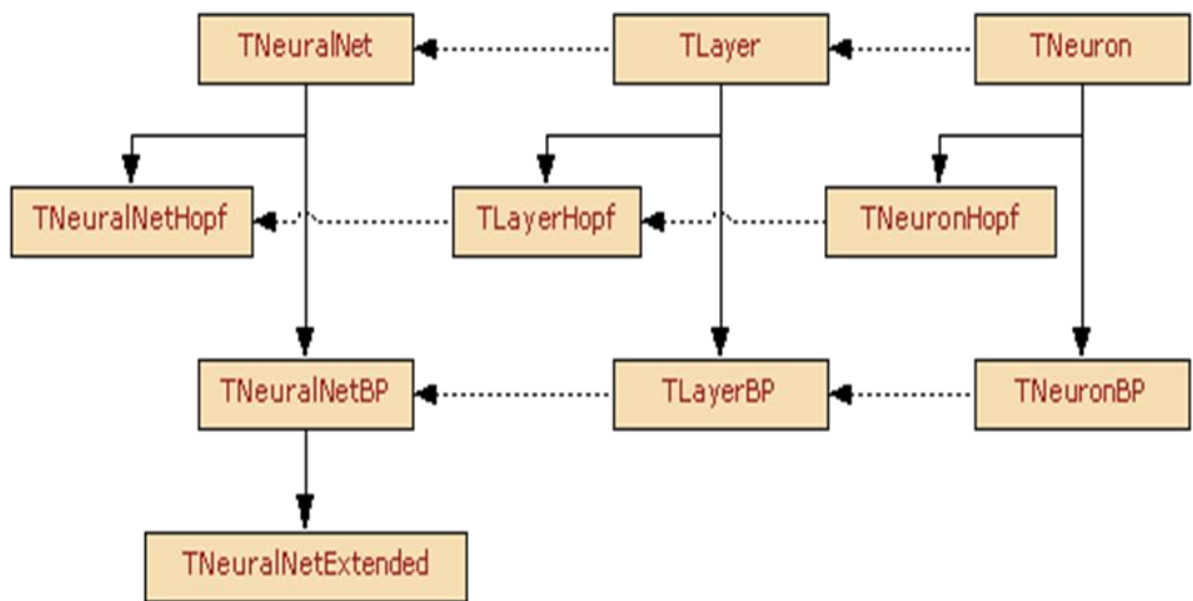


Рис. 2.2. Діаграма класів

TNeuronHopf – нащадок TNeuron, реалізовує нейрон використаний в нейромережі Хопфілда, єдиною відмінністю від базового класу, є використання функції активізації в перекритому методі ComputeOut [6].

ВИСНОВКИ ДО РОЗДІЛУ 2

Середовище Matlab в основі свого пакету інструментів має зручний інструмент для роботи з нейромережами під назвою "Neural Network toolbox", що і стало причиною вибору саме цього середовища.

Інструментарій створення графічного інтерфейсу і мова програмування не дуже зручні, проте не є дуже складними, головним пріоритетом є розробка надзвичайно продуктивної нейромережі.

Формат даних для навчання нейромережі нам необхідний у форматі протоколу NetFlow. Він має десятки характеристик мережевого потоку, завдяки чому нейромережа буде навчена декількома тестовими вибірками даних, щоб знайти найпродуктивніший варіант вхідних даних.

Данні для навчання вільно надані KDD Cup '99. Нейромережі у середовищі MATLAB приймають на вхід тільки числові данні, тому данні скоріш за все будуть потребувати попередньої обробки.

					IT-62.27.1081.01 ПЗ	Арк.
						34
Змн.	Арк.	№ докум.	Підпис	Дат.		

РОЗДІЛ 3 АРХІТЕКТУРА НЕЙРОМЕРЕЖІ

3.1 Про нейромережі

Нейронні мережі беруть натхнення з процесу навчання, що відбувається в мозку людини. Вони складаються зі штучної мережі функцій, які називаються параметрами, яка дозволяє комп'ютеру вивчати й налаштовувати себе, аналізуючи нові дані. Кожен параметр, іноді званий також нейроном, являє собою функцію, яка робить висновок після отримання одного або декількох входів.

Ці вихідні дані після обробки шаром нейронів передаються наступний прошарок нейронів, які використовуються в якості входних даних своєї власної функції та створюють додаткові вихідні дані. Далі процес повторюється, і так триває до тих пір, поки кожен шар нейронів не буде пройдений. Вихідні нейрони пізніше виводять остаточний результат для моделі, з якою нейронна мережа взаємодіє. На рис. 3.1 показано візуальне подання такої мережі.

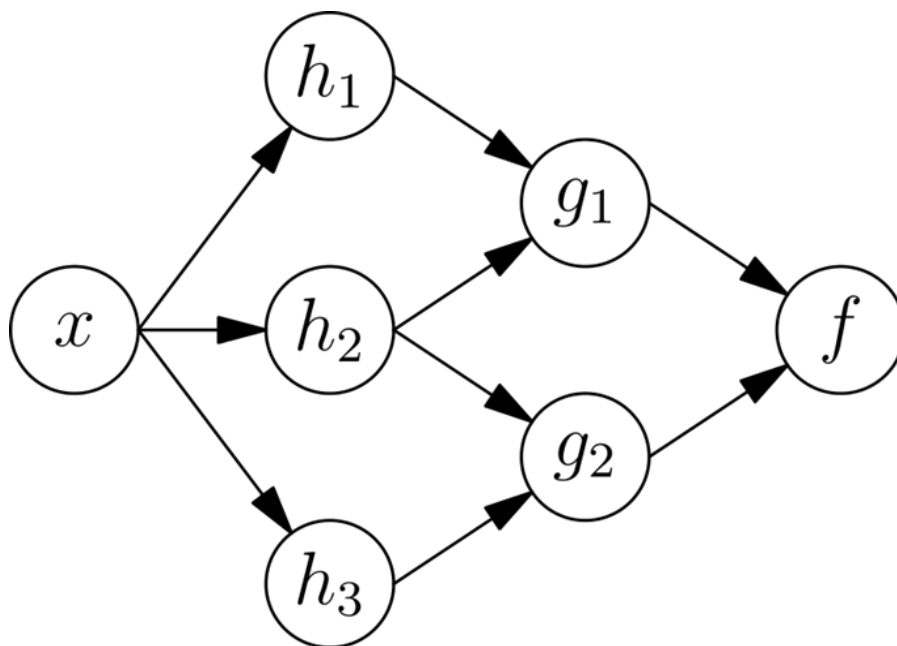


Рис. 3.1. Одношаровий персептон

Початковий сигнал – x , який потім передається першому шарові нейронів (h^*), де три функції розглядають вхід, який вони отримують, і генерують вихід.

Цей результат потім передається на другий шар (g^*). Там наступні вихідні дані розраховуються на основі вихідних даних з першого шару. Цей вторичний результат потім об'єднується для отримання остаточного результату моделі [1].

3.2 Архітектурні особливості нейромереж

Під нейронною мережею мають на увазі сукупність елементів, влаштованих подібно нейронам мозку, які пов'язані між собою та з зовнішнім середовищем. Зв'язки визначаються ваговими коефіцієнтами. Нейронні елементи існують трьох видів: вхідні, вихідні та приховані. Останні виконують основну роботу з перетворення вхідних даних у вихідні; і є основою нейронної мережі. Визначається вид нейрона з його розташування в топології мережі.

Архітектура нейромережі має такі складові:

- топологія зв'язків між нейронами;
- розділення певних підмножин нейронних елементів на вхідні, вихідні й оброблючі нейрони;
- способ навчання;
- наявність міжнейронної конкуренції;
- способи управління, синхронізації та передачі інформації між нейронами.

За топологією нейронні мережі класифікуються за такими типами:

- повнозв'язні;
- шаруваті (багатошарові);
- слабозв'язаних (з локальними зв'язками).

Кількість нейронів у кожному шарі може бути довільною і незалежною від інших шарів.

Серед архітектурного різноманіття різних нейромереж виділимо мережі прямого поширення. Вони описуються системою двонаправлених сигнальних зв'язків адаптивних елементів, які забезпечують перетворення сигналів при прямому і зворотньому поширенні.

					ІТ-62.27.1081.01 ПЗ	Арк.
						36
Змн.	Арк.	№ докум.	Підпис	Дат.		

Нейромережі можуть змінювати або один, або кілька нейронів за один момент часу. У цій відмінності нейронні мережі підрозділяються на синхронні і асинхронні. Тобто, асинхронна мережа за один момент часу може змінювати стан у групи нейронів, як правило, у всього шару одразу.

Під шаром мають на увазі один, або кілька нейронів, на входи яких подається однаковий загальний сигнал. Шарувату нейронну мережу визначає наявність окремих шарів таким чином, що обробка інформації відбувається по черзі між шарами.

У шаруватих мережах передача та обробка між шарами відбувається послідовно: нейрони початкового шару обробляють вступну інформацію, після чого передають її на наступний шар. Таким чином обробка буде проходити всі стадії аж до останнього шару, вихідні дані якого будуть зчитуватися інтерпретатором призначеного для користувача інтерфейсу.

В межах одного шару дані обробляються паралельно, а в межах нейронної мережі, – послідовно, від одного шару до іншого. Однак, це не правило, і сигнал не завжди подається на всі нейрони в шарі, як, наприклад, в когнітроні.

Когнітрон побудований з шарів нейронів, пов'язаних синапсами, але пресинаптичний нейрон в одному шарі пробуджує постсинаптичні нейрони в наступному шарі. Є два типи нейронів: збуджуючі клітини, які мають тенденцію пробуджувати постсинаптичні клітини, і гальмівні клітини, які зменшують цю тенденцію.

Активація нейрона залежить від зважених сум що його пробуджують, і гальмівних входів, однак фактичний механізм дещо складніше простого підсумовування. Кожен нейрон з'єднується тільки з нейронами в сусідній області, яка є його областю з'єднання. Обмеження на діапазон з'єднання схожий із анатомією зорової області кори головного мозку, де зв'язки між нейронами рідко встановлюються на відстані більше одного міліметра один від одного.

В моделі Фукусіми, нейрони розташовані в шарах, зі сполуками від одного шару до іншого. Знову ж таки, це схоже на шарувату структуру зорової кори, а також інших частин мозку [8].

3.3 Методика і типи навчання

Зворотнє поширення – це метод навчання нейронної мережі, що призводить до точного налаштування ваг нейронної мережі на основі частоти помилок, отриманих в попередню епоху (тобто ітерація). Правильна настройка ваг дозволяє знизити частоту появи помилок і підвищити надійність моделі за рахунок збільшення її узагальнення. Це стандартний метод навчання штучних нейронних мереж. Цей метод допомагає розрахувати градієнт функції втрат стосовно всіх ваг в мережі.

Метод навчання зворотнім поширенням здійснюється за таким алгоритмом:

1. Входи надходять через попередньо підключене введення.
2. Введення моделюється з використанням реальних ваг W , які підбираються випадковим чином.
3. Розрахунок виходу для кожного нейрона від шару введення до прихованих шарів, а потім до шарів виведення.
4. Розрахунок помилки рішення на вихідному шарі, отриманої з порівняння очікуваного й фактичного результату.
5. Перерахунок ваг прихованих шарів на основі помилки обчислення вихідного шару.

Навчання з вчителем – це варіант, який зустрічається найчастіше. В цьому випадку в наявності має бути присутня підбірка правильних відповідей. Або, якщо орієнтуватись й розроблювати нейромережу для біржі цінних паперів, точно відомий курс якихось акцій протягом тривалого часу.

Навчання без вчителя (кластеризація) полягає в поданні на вхід мережі деяких даних і передбачається, що в них можна виділити кілька класів цих даних.

					ІТ-62.27.1081.01 ПЗ	Арк.
						38
Змн.	Арк.	№ докум.	Підпис	Дат.		

Навчання без вчителя використовується в тих випадках, коли потрібно виявити внутрішні взаємозв'язки, залежності й закономірності, що існують між об'єктами, але заздалегідь невідомо, в чому ці закономірності проявляються.

3.4 FFNN і MLP

Нейронна мережа з прямим зв'язком (FFNN), – це тип нейронних мереж, в яких міжнейронні з'єднання не створюють нескінченних циклів. Нейронні мережі прямого поширення є найпростішою різновидністю класичних нейронних мереж. Інформація (вхідні дані), обробляється і передається тільки в одному напрямку (вперед) [3].

Багатошаровий персептрон (MLP) – це різновид штучної нейронної мережі з прямим зв'язком (ANN), показаний на рис.3.2. Термін MLP використовується неоднозначно, іноді для позначення будь-якого прямого зв'язку ANN, іноді строго для позначення мереж, що складаються з декількох шарів персептонів.

Складається ж багатошаровий персептрон з шарів трьох рівнів:

1. Вхідний.
2. Прихований.
3. Вихідний.

Кожен вузол є нейроном (якщо він не вхідний), що використовує для своєї активації нелінійну функцію. Від лінійного персептона відрізняється багатошаровістю і нелінійною активацією (дані, які не є лінійно ділимі, можуть бути розділені) [8].

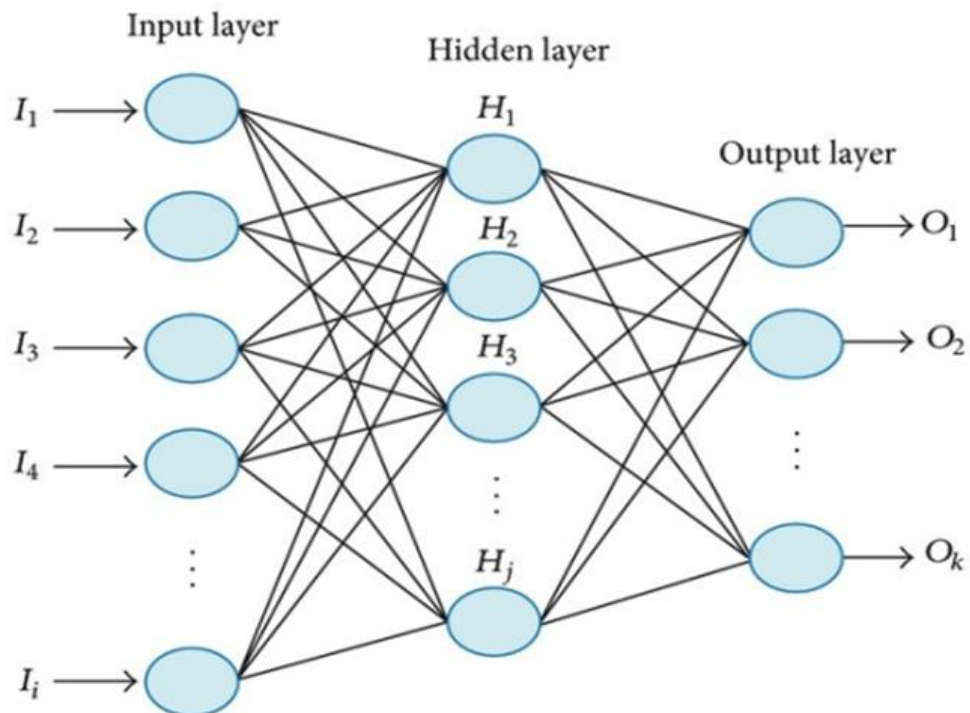


Рис. 3.2. Багатошаровий персептон

3.5 Процедура навчання MLP:

Процедура навчання MLP виглядає наступним чином:

1. Починаючи з вхідного шару, поширимо дані вперед на вихідний шар. Цей крок – пряме поширення.
2. Грунтуючись на вихідних дані, відбудеться розрахунок помилки (різницю між прогнозованим і відомим результатом). Помилка з кожної ітерації повинна бути зведена до мінімуму.
3. Помилка буде поширена на всі попередні шари, крім вхідного. Це означає перерахунок вагових коефіцієнтів кожного нейрона в залежності від його минулого значення і значення помилки.

Кроки будуть повторені N-разів, поки в залежності від показників помилки, статистики та градієнта поширення нейронна мережа не покаже задовільні результати [3].

На першому етапі необхідно обчислити одиницю активації прихованого шару за формулами (3.1) і (3.2):

$$z_1^{(h)} = a_0^{(in)} w_{0.1}^{(h)} + a_1^{(in)} w_{1.1}^{(h)} + \dots + a_m^{(in)} w_{m.1}^{(h)} \quad (3.1)$$

$$a_1^{(h)} = \phi(z_1^{(h)}), \quad (3.2)$$

де z – сигнал пропущений через нейромережу (кінцевий), $a_0^{(in)}$ – початковий сигнал, переданий у нейрон, w – певна вага нейрону, h – позначення належності до прихованого шару нейронної мережі. Одиниця активації є результатом застосування функції активації ϕ до значення z . Вона повинна бути диференційована, щоб мати можливість вивчати вагу з використанням градієнтного спуску [3].

Функція активації ϕ часто є сигмоїдальною (логістичною) функцією (рис. 3.3).

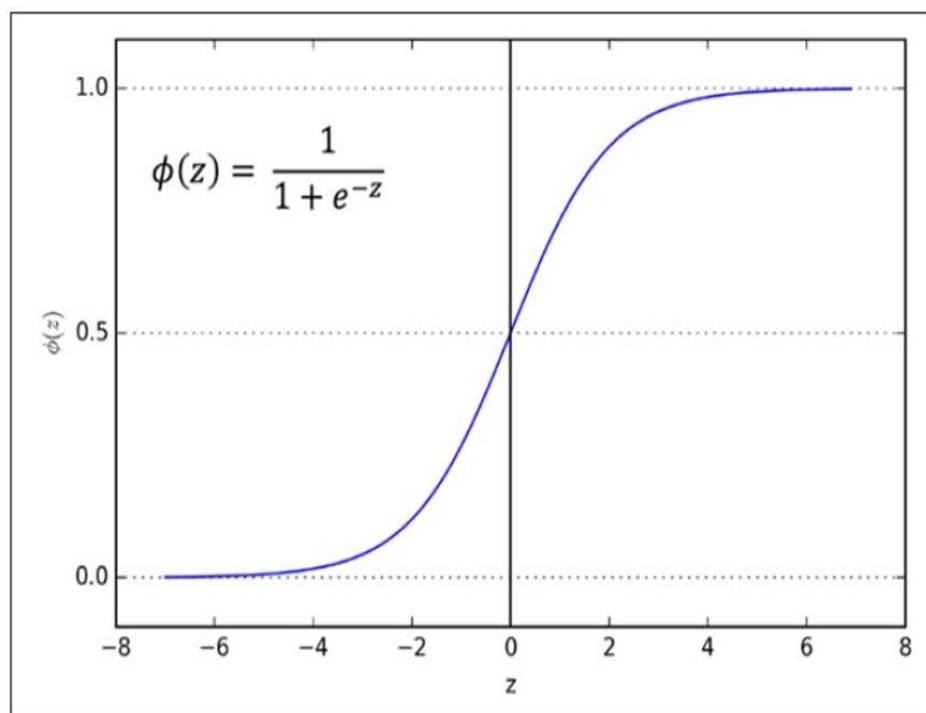


Рис. 3.3. Функція активації

Це дозволяє нелінійність, необхідну для вирішення складних завдань, таких як класифікація аномалій мережевого трафіку.

3.6 Особливості вибору і підготовки даних

Був обраний набір даних KDD Cup 1999, який проводився спільно з "П'ятою міжнародною конференцією з вивчення та інтелектуального аналізу даних".

Завдання конкурсу полягало в створенні детектора вторгнення в мережу, прогнозуючого моделі, який здатний розрізняти аномальні з'єднання, звані вторгненнями або атаками, та нормальні з'єднання. Ця база даних містить стандартний набір даних для аудита, який включає в себе широкий спектр вторгнень, що були змодельовані на базі локальної мережі військової бази ВПС США.

Lincoln Labs створила середовище для збору необроблених даних за дев'ять тижнів. Вони наповнювали її численими атаками, щоб змодельовати реальне навантаження атаками серверів ВПС США.

Вихідні дані навчання становили близько чотирьох гігабайт стислих даних мережевого трафіку. Це близько п'яти мільйонів рядків з даними про підключення. Ще за два тижні, тестових даних було отримано близько двох мільйонів записів.

Підключення – це послідовність пакетів, що починаються і закінчуються в деякі чітко визначені моменти часу, між якими дані передаються від і з вихідної IP-адреси, до цільової IP-адреси по деякому чітко визначеному протоколу. Кожне з'єднання позначається як нормальне або як атака з точно визначеним типом атаки. Кожен запис з'єднання складається з близько 100 байтів.

Атаки діляться на чотири основні категорії:

- DoS: відмова в обслуговуванні, наприклад "syn flood";
- R2L: несанкціонований доступ з віддаленого комп'ютера, наприклад, вгадати пароль;
- U2R: несанкціонований доступ до локальних привілеїв суперкористувача (root), наприклад, різні атаки “переповнення буфера”;

					IT-62.27.1081.01 ПЗ	Арк.
						42
Змн.	Арк.	№ докум.	Підпис	Дат.		

– зондування: спостереження і інше зондування, наприклад, сканування портів.

Важливо відзначити, що дані тесту не мають того ж розподілу ймовірностей, що і дані навчання, і включають конкретні типи атак, яких немає в системі адаптації. Це робить задачу більш реалістичною. Деякі експерти з вторгнень вважають, що більшість нових атак є варіантами відомих атак, і “сигнатури” відомих атак може бути достатньо для виявлення нових варіантів.

Набори даних містять в цілому 24 типи тренувальних атак, і додатково 14 типів, але тільки в тестових даних. Приклад характеристик з'єднання наведено в таблиці (табл. 3.1).

Таблиця 3.1 Характеристики з'єднання

Назва	Опис	Тип
duration	Тривалість підключення(сек)	Тривалий
protocol_type	Тип протоколу: tcp, udp, чи інший	Тривалий
service	Мережевий сервіс призначення	Дискретний
src_bytes	Кількість відправлених байтів	Тривалий
dst_bytes	Кількість прийнятих байтів	Тривалий
flag	Статус підключення (1 якщо помилковий)	Дискретний
land	1 якщо підключення з однакового порту	Дискретний
wrong_fragment	Кількість пошкоджених\хибних фрагментів	Тривалий
urgent	Кількість термінових пакетів	Тривалий

ВИСНОВКИ ДО РОЗДІЛУ 3

Нейронні мережі є складними моделями, які намагаються імітувати спосіб, яким людський мозок розробляє правила класифікації. Нейронна мережа складається з безлічі різних шарів нейронів, причому кожен шар отримує вхідні дані від попередніх і передає вихідні дані в інші шари. Те, як вихідні дані кожного шару стають вхідними для наступного шару, залежить від ваги, від розрахункової функції та від оптимізатора, певних для цієї конкретної мережі.

Нейронна мережа виконує ітерації протягом заздалегідь визначеного числа ітерацій, званих епохами. Після кожної епохи функція вартості аналізується, щоб побачити, де модель може бути поліпшена. Потім оптимізуюча функція змінює внутрішню механіку мережі, вагові коефіцієнти та зміщення, на основі інформації, наданої функцією ваги, але до тих пір, поки функція ваги не буде скасована.

В ході порівняльного аналізу основних класів архітектур нейронних мереж було виявлено, що навчання мережі спрямованої на класифікацію й аналіз таких обсягів даних, що були обрані, потребує десятки, а то й сотні обчислювальних одиниць.

Було вирішено використовувати MLP-архітектуру, так, як вона може бути прекрасно налаштованою на будь-які вхідні та вихідні дані із максимальним показником продуктивності класифікації. Однак, в цієї архітектури є і свої недоліки: неспроможність перевчитись та високий час навчання.

РОЗДІЛ 4 РЕАЛІЗАЦІЯ ТА ТЕСТУВАННЯ

4.1 Формування концепції

Необхідно розробити інтерфейс, здатний вмістити в собі інформативність і лаконічність, не втрачаючи при цьому функціональності. Сучасні інтерфейси бувають жовх видів:

- одновіконні;
- багатовіконні.

Подавляюча більшість сучасних інтерфейсів – багатовіконні. Це може бути виправдано багатofункціональністю та перевантаженністю додатковими модулями.

В даному проекті буде використаний одновіконний інтерфейс, адже цільове завдання одне, – класифікувати різні типи відхилень мережевого трафіку від норми, що ергономічно було б зробити в одному вікні.

Введення даних підключення буде організоване за допомогою випадających списків і текстових полів, що надасть можливість продемонструвати користувачеві конкретні характеристики, по яким буде проводитись аналіз, та результат, нічого надлишкового. Принцип швидкої зміни поля введення буде спиратися на принцип зміни фокусу елемента, що надасть можливість переключатися використовуючи клавішу Tab у Windows-системах.

Рішення щодо кольору повинно відображати тільки важливі деталі, не кидатися в очі, не відволікати від основної задачі. Інтерфейс повинен бути виконаний у гаммі нейтральних світлих відтінків, з використанням виділення елементів вводу, в яких будуть аномальні значення, а також сам вердикт нейронної мережі про підключення, прихованої за графічним вікном користувача.

					ІТ-62.27.1081.01 ПЗ	Арк.
						45
Змн.	Арк.	№ докум.	Підпис	Дат.		

4.2 Реалізація інтерфейсу

MATLAB надає всі інструменти не тільки для розробки, тренування та перевірки нейронної мережі, а ще й для написання сучасних гнучких інтерфейсів користувацького вводу.

В ході тестування та налагодження мережі виникла проблема надлишку характеристик трафіку для навчання нейронної мережі, що призвело до надто довгого навчання. Були відібрані найбільш вагомні характеристики, всього яких десять.

Був створений прототип с десятима полями для введення характеристик і однією кнопкою, котра запускає процес тестування. У тестовій експлуатації інтерфейс виявився зручним і приємним на вигляд (рис.4.1).

Netflow Analyze Tool

NetFlow Analyzer

Flow datas

Duration (sec): 10 Protocol: UDP Service: private Source bytes: 1.032e+04

Destination bytes: 7.332e+04 Server count: 2 Count: 4 Dest. host count: 10

dst_host_srv_count: 31

dst_host_same_srv_rate: 1.8

Analyze

Smurf

Рис. 4.1. Інтерфейс розроблюваного додатку

4.3 Навчання та тестування нейронної мережі

Навчання нейромережі – найважчий етап розробки. Не через технологічну складність, а через калібровку та перенавчання. Навчання нейромережі — нешвидкий процес, що зайняв найбільший час.

Продемонстрована нижче нейромережа (рис.4.2) не задовільняє мінімальним вимогам якості, вона детектувала лиш 37% аномалій.

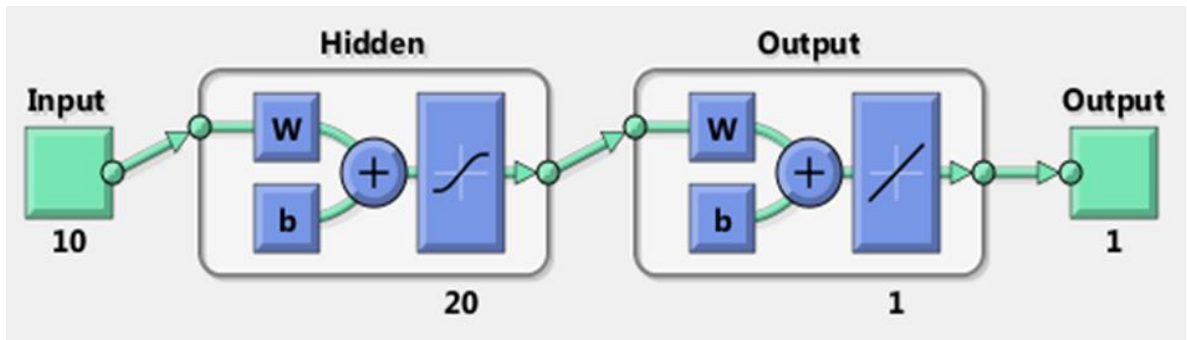


Рис.4.2. Початкова архітектура нейромережі

На графіку (рис.4.3) наведений градієнт функції втрат протягом різних етапів створення нейронної мережі. Нижнє вимірювання – це номер епохи. Як можна побачити, градієнт функції втрат став досить високим. Градієнт розподілу дорівнював 0.0403, станом на 26 епоху.

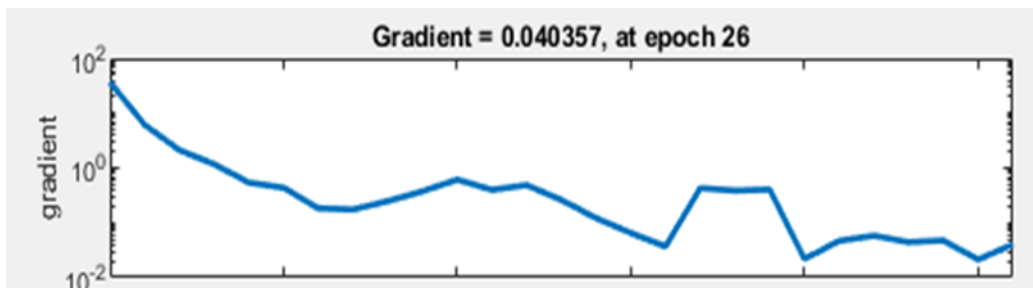


Рис.4.3. Градієнт функції втрат

У процесі створення та тренування нейромережі виникла проблема з навчанням. Даних було так багато, що створювана нейромережа спочатку навчалася три тижні, без перерви. Показники в неї були не надто високі, як для такого часу, виділеного на навчання. Вона видавала усього лише 37% правильних відповідей. Це був одношаровий персептрон із двадцяти нейронів у прихованому шарі.

У даних KDD Cup '99 всього 41 характеристика трафіку. Задача відбору даних полягає у знаходженні найвагоміших характеристик трафіку, що дозволять дуже чисто та правильно детектувати обрані аномалії.

Самих підключень виявилось надто багато, було окремо відібрано три вибірки: для навчання, для тестування та для експлуатаційного тестування нейронної мережі.

Було вирішено відібрати найдинамічніші й найважливіші параметри підключення, які мають найбільшу вагу у відмінностях між різними аномальними станами мережі.

В кінцевому результаті були знайдені найпродуктивніші параметри трафіка (табл.4.1), далі будуть показані результати та графіки навчання.

Таблиця 4.1. Ітогові характеристики для навчання нейрмережі

Опис	Тип
Тривалість з'єднання	Тривалий
Тип протоколу	Дискретний
Сервіс підключення	Дискретний
Відправлено байт	Тривалий
Прийнято байт	Тривалий
Кількість підключених сервісів	Тривалий
Кількість інших підключень	Тривалий
Вихідні підключення до віддаленого сервера	Тривалий
Відношення прийнятих/отриманих пакетів від віддаленого серверу	Тривалий

Архитектура нейронної мережі мультишарового персептона в MATLAB виглядає таким чином (рис.4.4):

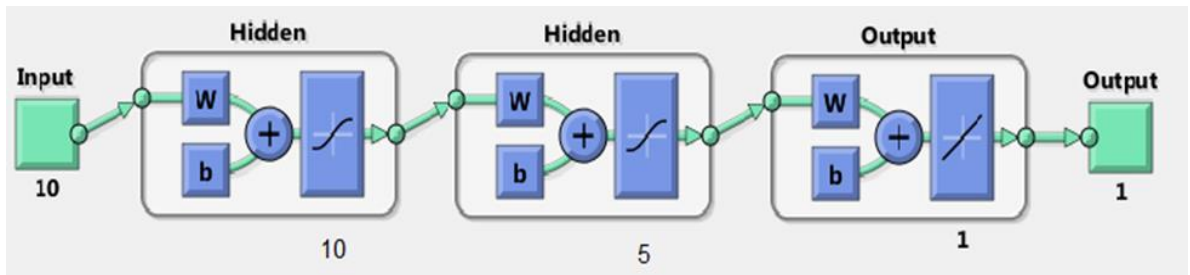


Рис. 4.4. Архітектура мережі мультишарового персептона

Десять нейронів прихованого шару необхідно для того, щоб виконати першим етапом аналіз відхилень характеристик від норми, а другий шар вже здійснював порівняння результатів, на основі чого приймає рішення, щодо кожної конкретної аномалії. До того ж ця мережа краще справляється з поставленою задачею, ніж 20-ти нейронна одношарова архітектура, що була розглянута раніше.

Нейромережа вчиться катастрофічно довго (рис. 4.5), має малий градієнт функції втрат, що є добре для розпізнавання конкретних категорій атак та аномалій, та має надзвичайно малу похибку.



Рис. 4.5. Звіт навчання

На рис. 4.6, можна робачити, що піку навчання нейромережа досягає вже десь на дадцятому циклі навчання.

Згідно графікам на рис. 4.7, нейромережа навчилася класифікувати аномалії без помилки, тобто для неї ця задача не є важкою і якщо робити систему безпеки мережевого трафіку знову, то можна зробити систему з декількох

нейромереж, та навчити їх працювати між собою, заради більшої продуктивності та безпеки, чи обрати більш багатошарову архітектуру.

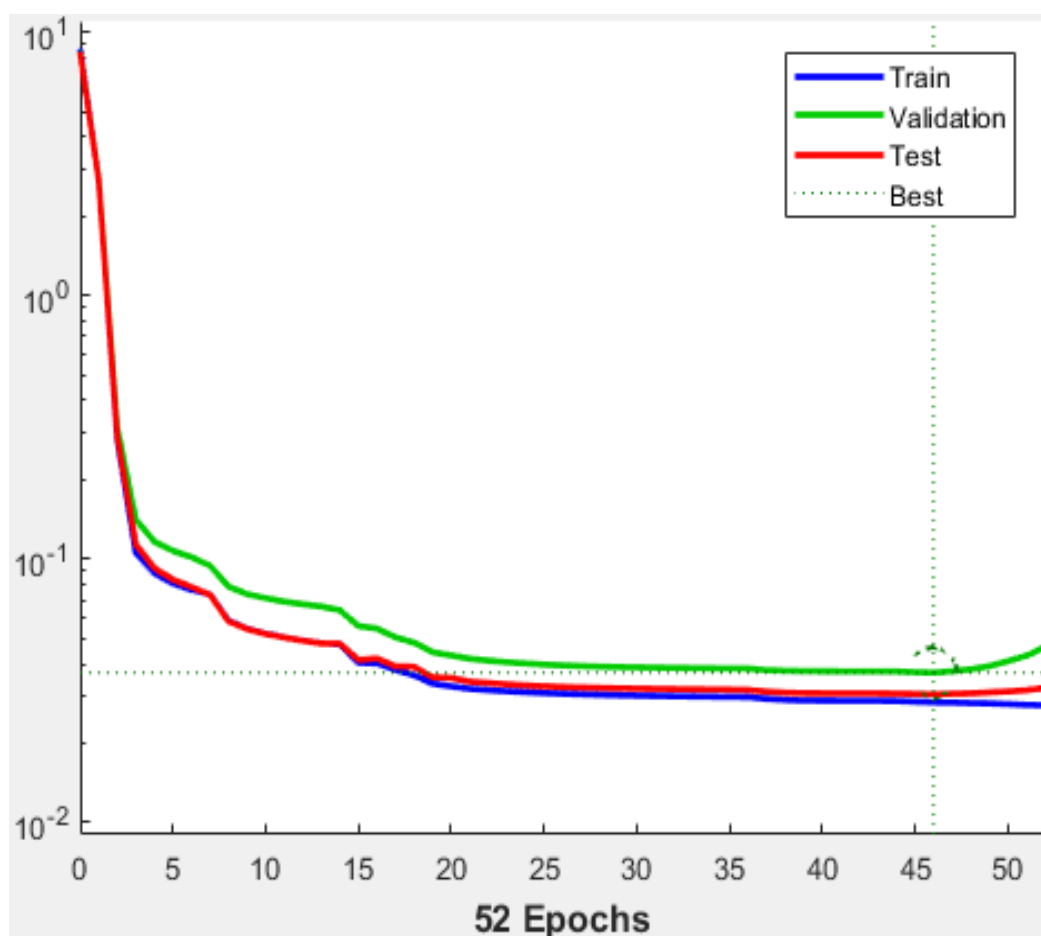


Рис. 4.6. Графік середньоквадратичної похибки

Архітектура, що була побудована, відрізняється надвисокою швидкістю роботи та дуже високими показниками продуктивності. Відсоток вірних відповідей у тестувальній експлуатації склав аж 100%.

В дипломному проєкті нейромережа навчилася ідентифікувати трафік практично без помилок.

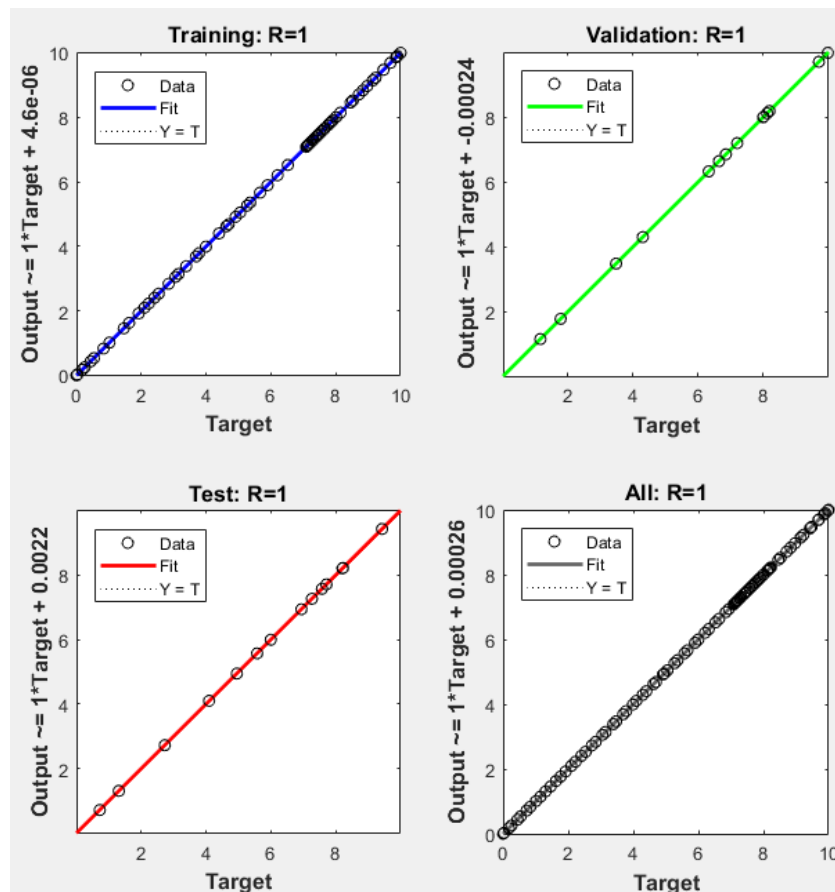


Рис. 4.7. Результат навчання

При необхідності робити систему безпеки мережевого трафіку знову, доцільно буде модифікувати цю систему додатковими нейромережами та навчити їх працювати між собою заради більшої продуктивності та безпеки.

ВИСНОВКИ ДО РОЗДІЛУ 4

В реальності архітектура персептона виявилась визначальною для нейромережі. Інші архітектури або занадто вибагливі до навчання, способу представлення даних, обчислювальним потужностям, а персептон виділяється своєю простотою.

Було перебрано декілька основних архітектур персептона для спроби навчити систему успішно визначати основні класи загроз і аномалій. Виникли великі труднощі в пошуку й обробці даних для найбільш ефективного процесу навчання з максимально високим відсотком вірно розпізнаних загроз. Нейронні мережі здатні навчатися тижнями, місяцями, тому спосіб подачі даних має досить велике значення.

Данні, що були взяті з KDD Cup 1999, вимагали додаткової обробки і селекції. Довелося перебрати всі дані, точно і детально проаналізувати, крім того, передивитись динаміку зросту від часу й загальну амплітудність характеристик серед класів атак, щоб знайти найвагоміші характеристики для розроблюваної нейронної мережі. Врешті решт було обрано десять характеристик протоколу NetFlow.

В дипломному проєкті нейромережа навчилася ідентифікувати трафік практично без помилок. При необхідності робити систему безпеки мережевого трафіку знову, доцільно буде модифікувати цю систему додатковими нейромережами та навчити їх працювати між собою заради більшої продуктивності та безпеки.

Інтерфейс додатку був зібраний і виконаний стандартними засобами MATLAB 2019 року. Ця середа має лаконічний і зрозумілий код скриптів, що дозволяє швидко в ній розібратися та зібрати користувацький графічний інтерфейс під будь-яку задачу. В данному проєкті був зібраний досить лаконічний продукт.

					ІТ-62.27.1081.01 ПЗ	Арк.
						52
Змн.	Арк.	№ докум.	Підпис	Дат.		

ВИСНОВКИ

Програмний продукт, описаний в даній роботі, відповідає вимогам і був розроблений в зазначені терміни. В процесі розробки мали місце деякі ускладнення, що були вирішені. Наприклад, існувала необхідність підібрати загальну модель нейронної мережі та розширити під конкретну задачу, що була поставлена.

Були розглянуті основні класи нейронних мереж, включаючи моделювання специфіки цих архітектур на поставлене завдання. Фінальним рішенням був багатошаровий перцептон (MLP).

В процесі створення і тренування нейронної мережі виникла проблема з навчанням. Даних було так багато, що створювана нейронна мережа спочатку навчалася 3 тижні без перерви. Самих же підключень виявилось занадто багато.

Показники у неї були не надто високі для такого часу навчання. Вона видавала всього-на-всього 37% правильних відповідей. Це був одношаровий перцептон з 20 нейронів в прихованому шарі. Тому було відібрано окремо 3 вибірки: для навчання, для тестування нейронної мережі безпосередньо, а також для тестування експлуатацією завершеного продукту.

Нейронні мережі здатні вчитися тижнями, місяцями, тому спосіб подачі даних має досить велике значення. Підсумкова мережа нараховує 10 вхідних нейронів, 10 в першому прихованому шарі, 5 у другому прихованому шарі, і 1 нейрон на вихідному шарі.

Саме така топологія була побудована, спираючись на попередні аналіз і обробку початкових даних.

Також було вирішено відібрати найдинамічніші і важливіші параметри підключень, що мають великий діапазон припустимих значень.

Інтерфейс програми був зібраний і втілений засобами Matlab 2019 року, що дало перевагу в зв'язуванні інтерфейсу користувача з попередньо навченої і откалиброваної нейронною мережею.

					ІТ-62.27.1081.01 ПЗ	Арк.
						53
Змн.	Арк.	№ докум.	Підпис	Дат.		

Серед Matlab має лаконічну і зрозумілу скриптову мову, що дозволяє швидко в ній розібратися і зібрати для користувача графічний інтерфейс під будь-яке завдання. У даній роботі був зібраний призначений для користувача інтерфейс продукту, що відповідає всім як функціональним, так і нефункціональним вимогам.

Нейронна мережа з легкістю стала відображати заявлені аномальні стани мережевого підключення, яких було 5, з дуже малою часткою помилок.

Були отримані теоретичні знання побудови нейронних мереж різних складнощів, що дає можливість в подальшому доопрацювати програму ретроспективним аналізом, або вдосконалити архітектуру.

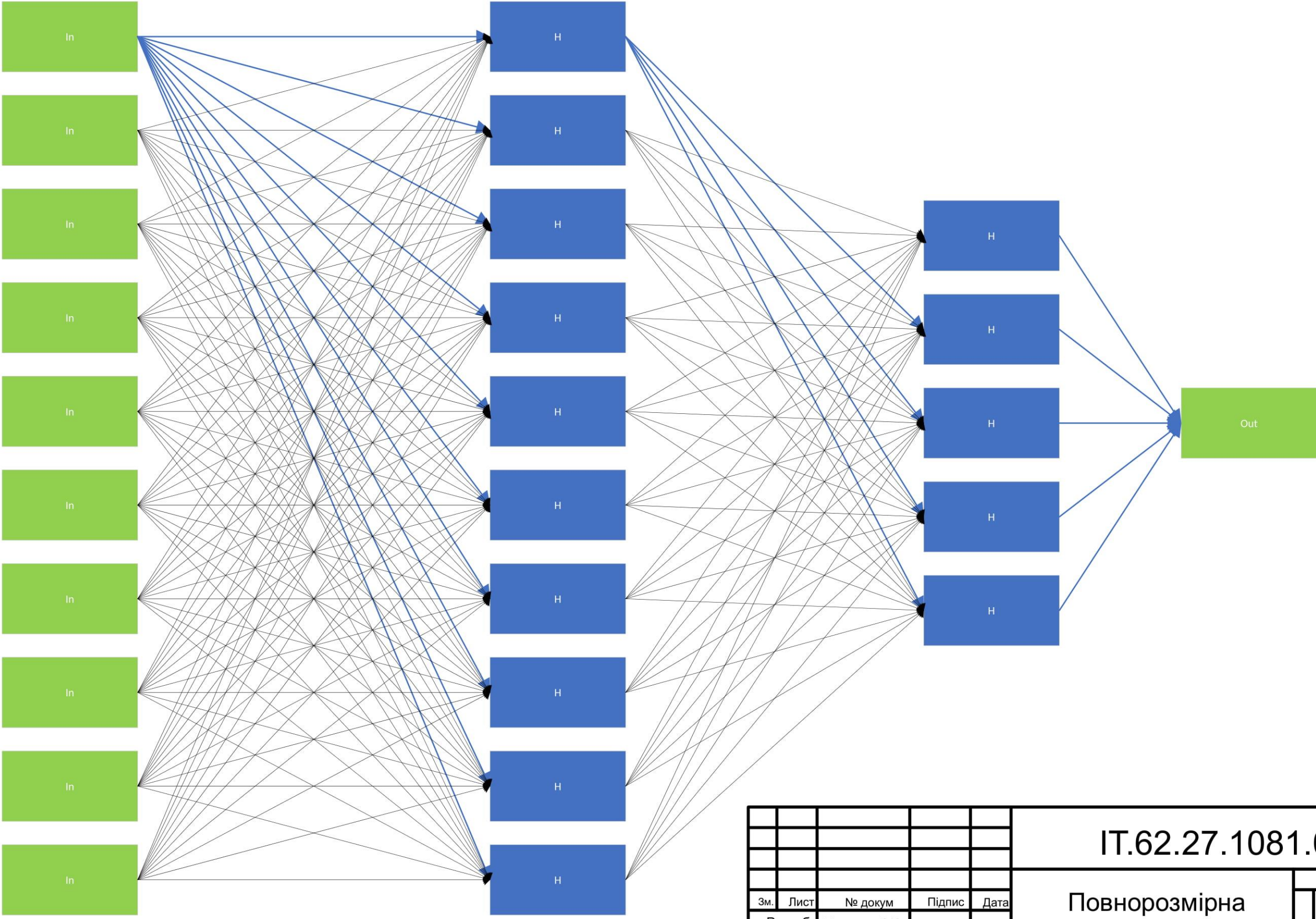
					ІТ-62.27.1081.01 ПЗ	Арк.
						54
Змн.	Арк.	№ докум.	Підпис	Дат.		

ПЕРЕЛІК ПОСИЛАНЬ

1. Григорівський Є.С. Адміністрування в інформаційних системах. [Електронний ресурс]. – 2009. – Режим доступу до ресурсу: <http://uadoc.zavantag.com/text/15386/index-1.html/>
2. Игорь Панов. 8 лучших программ для анализа сетевого трафика. [Електронний ресурс]. – 2015-2020. – Режим доступу до ресурсу: <https://networkguru.ru/8-luchshikh-programm-dlia-analiza-setevogo-trafika/>
3. Feedforward neural network [Електронний ресурс]. – Режим доступу до ресурсу: https://en.wikipedia.org/wiki/Feedforward_neural_network/
4. Атака ширококешательными ICMP ECHO пакетами (Smurf Attack) [Электронний ресурс]. – Режим доступу до ресурсу: <https://ddos-guard.net/ru/terminology/ataka-shirokoveshatelnyimi-icmp-echo-paketami-smurf-attack/>
5. SNMP протокол – принципы, безопасность, применение [Електронний ресурс]. – Режим доступу до ресурсу: <http://www.iso27000.ru/chitalnyi-zai/setevaya-bezopasnost/snmp-protokol-principy-bezopasnost-primenenie/>
6. NeuralBase – нейросеть за 5 минут [Електронний ресурс]. – Режим доступу до ресурсу: <https://basegroup.ru/community/articles/fastneuralnet/>
7. І. Івановіч та С. Гайін. (2016). Recommendations for Network Traffic Analysis Using the NetFlow Protocol. 40-42.
8. Eugenio Culurciello. Архитектуры нейросетей [Електронний ресурс]. – 2018. – Режим доступу до ресурсу: <https://habr.com/ru/company/nix/blog/430524/>

					IT-62.27.1081.01 ПЗ	Арк.
						55
Змн.	Арк.	№ докум.	Підпис	Дат.		

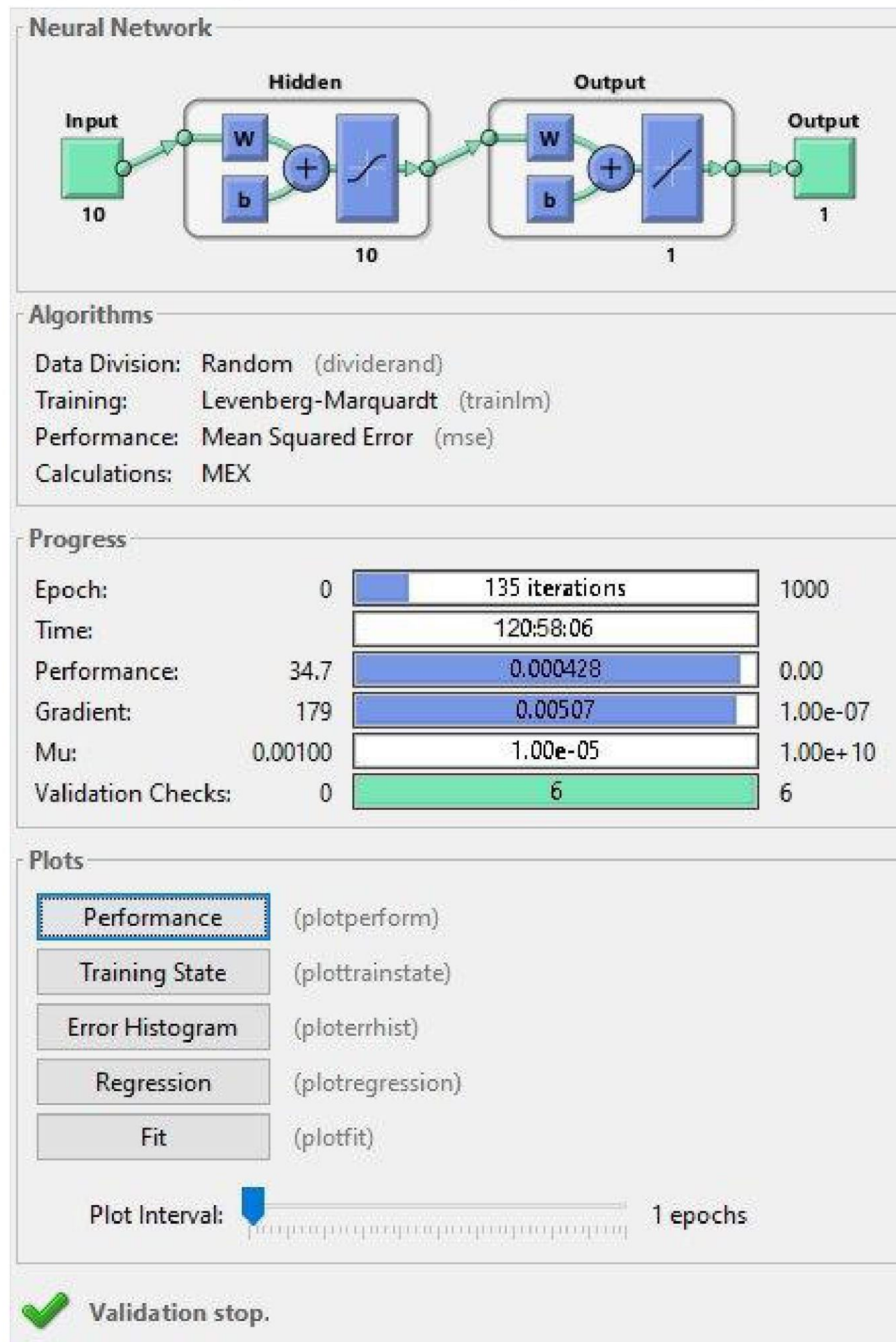
IT.62.27.1081.01



Інв. № ориг.	Підпис і дата	Взам. інв. №	Інв. № дубл.	Підпис і дата

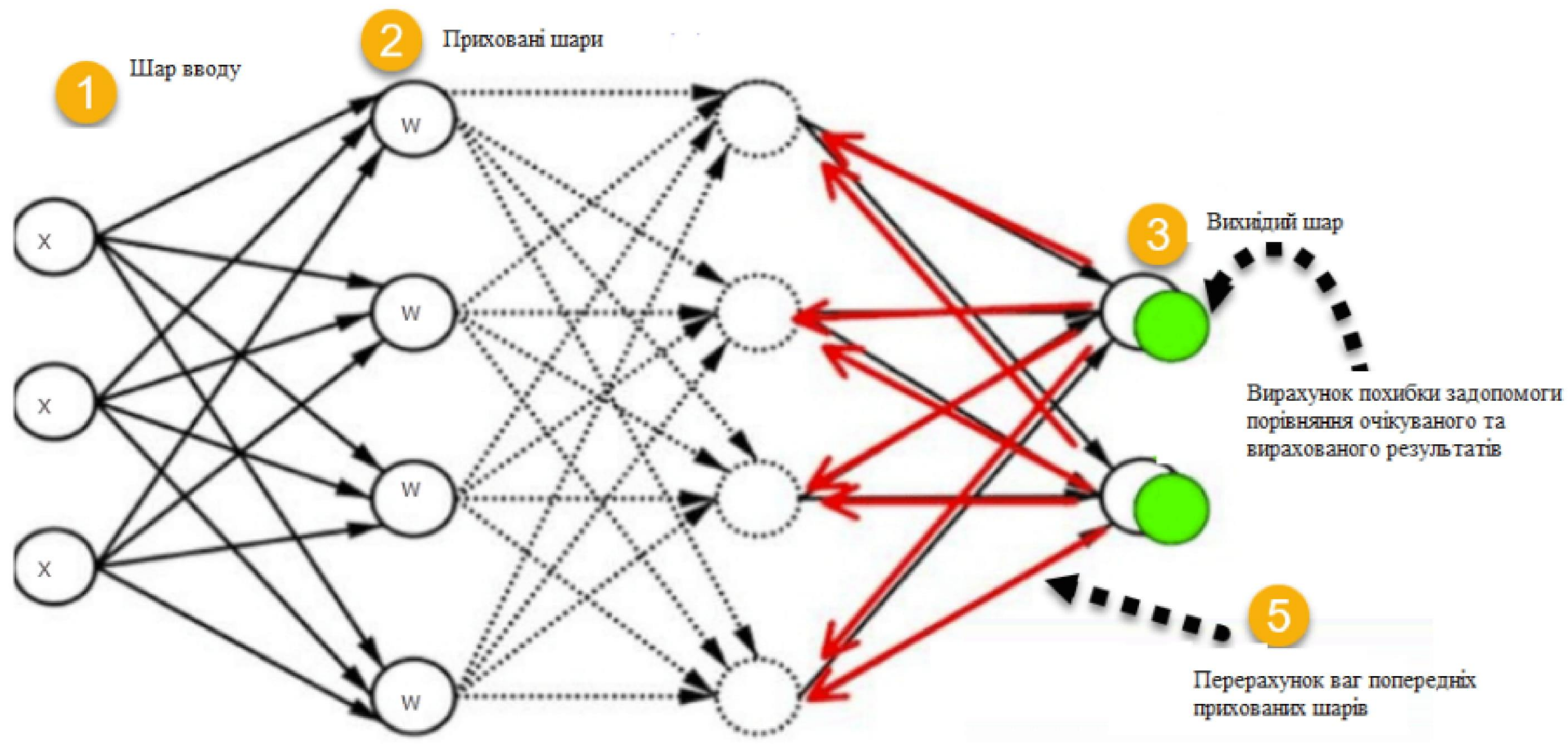
					ІТ.62.27.1081.01 ВЗ					
					Повнорозмірна архітектура мережі	Літ.		Маса	Мірило	
Зм.	Лист	№ докум	Підпис	Дата		Лист		Листів		
Розроб.		Позднякова О.Ю.								
Перев.		Батрак Є.О.								
Н.-Контр.		Пасько В.П.			Кафедра Технічної кібернетики	КПІ ім. Ігоря Сікорського				
Затв.		Пархомей І.Р.								

Одношаровий персептон, який довелося навчати 120 годин



Демонстраційний плакат №1
до дипломного проєкту на тему
„Інтелектуальний додаток для ідентифікації
аномалій мережевого трафіку”

Розробила: студентка групи IT-62 Позднякова О.Ю.
Прийняв: доц. к.т.н кафедри ТК Батрак Є.О.



Підпис і дата	Інв. № дубл.	Взам. інв. №	Підпис і дата	Інв. № ориг.

					IT.62.27.1081.01 ВЗ							
					Алгоритм навчання нейромережі	Літ.			Маса	Мірило		
Зм.	Лист	№ докум	Підпис	Дата								
Розроб.		Позднякова О.Ю.										
Перев.		Батрак Є.О.										
						Лист			Листів			
Н.-Контр.		Пасько В.П.			Кафедра Технічної кібернетики	КПІ ім. Ігоря Сікорського						
Затв.		Пархомей І.Р.										

Власник документу:
Лісовиченко Олег Іванович

ID перевірки:
1003791067

Дата перевірки:
04.06.2020 22:52:54 EEST

Тип перевірки:
Doc vs Internet + Library

Дата звіту:
04.06.2020 23:03:51 EEST

ID користувача:
76913

Назва документу: IT-62_ПоздняковаОленаЮрївна

ID файлу: 1003804949 Кількість сторінок: 41 Кількість слів: 8523 Кількість символів: 63254 Розмір файлу: 120.56 KB

5.61% Схожість

Найбільша схожість: 1.01% з джерело бібліотеки. ID файлу: 5922694

1.61% Схожість з Інтернет джерелами

66

Page 43

4.96% Текстові збіги по Бібліотеці акаунту

193

Page 43

0% Цитат

Не знайдено жодних цитат

0% Вилучень

Вилучений текст відсутній

Підміна символів

Заміна символів

1