ДОДАТОК А

Графічний матеріал кваліфікаційної роботи

ХАРКІВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ РАДІОЕЛЕКТРОНІКИ

КВАЛІФІКАЦІЙНА РОБОТА

НА ТЕМУ: Метод виявлення вторгнень у розподілених інформаційних системах

ВИКОНАВ:
Студент гр. СПм-21-1 Кізлевич Я.О.

КЕРІВНИК:
к.т.н. доц. Мартовицький В. О.

ХАРКІВ
2022р.

Мета роботи

Об'єктом дослідження є процес моніторингу стану інформаційного та комунікаційного середовища розподілених інформаційних систем.

Предметом дослідження є методи і алгоритми моніторингу в розподілених інформаційних системах із застосуванням технологій інтелектуального аналізу даних.

Метою роботи є покращення показників виявлення аномалій функціонування РІС в умовах кібернетичних впливів зовнішнього та внутрішнього середовища шляхом побудови моделей і методів на основі технологій інтелектуального аналізу даних.

Завдання

Для досягнення поставленої мети вирішуються наступні задачі:

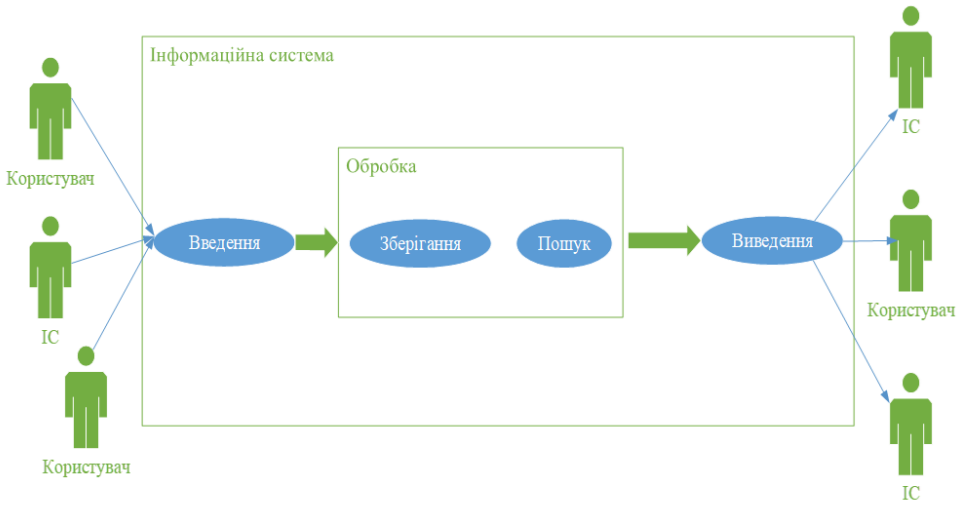
– проаналізувати підходи щодо забезпечення безпеки розподілених інформаційних систем;

– розглянути особливості архітектури розподілених інформаційних систем і виділити базові компоненти систем;

– розробити метод виявлення аномалій системи з використанням методів інтелектуального аналізу для класифікації стану функціонування комп'ютерних систем;

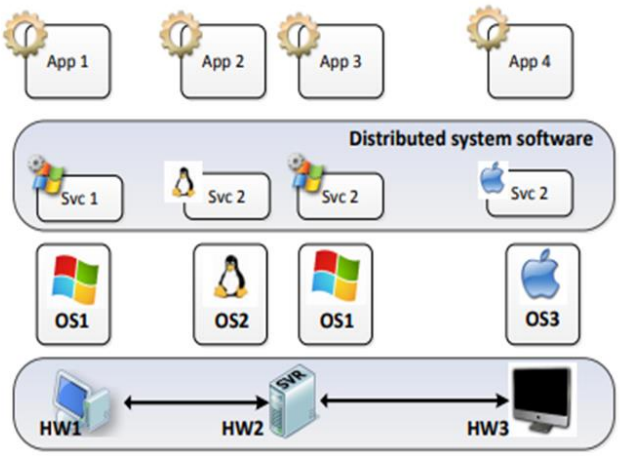
3

Загальні функції інформаційної системи



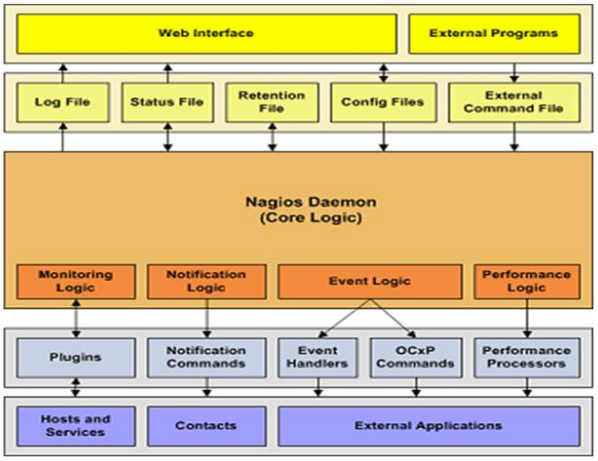
4

Модель розподіленої комп’ютерної системи



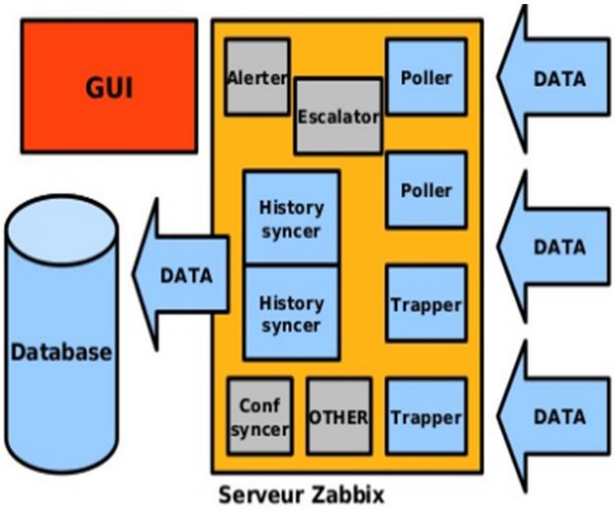
5

Архітектура системи Nagios



6

Архітектура системи Zabbix



7

Таблиця 1 – TCP параметри.

Параметр	Опис
Duration	Тривалість з'єднання(секунди)
protocol _ type	Протокол транспортного рівня
Service	Сервіс прикладного рівня
number of data bytes source - destination	Потік, що входить, байт
number of data bytes destination - source	Вихідний потік, байт
Flagstatus	Прапори, встановлені в заголовку TCP- пакету.
Land	Адреси співпадають, 0 інакше
wrong _ fragment	Число неправильних фрагментів
urgentpackets	Наявність термінових даних в пакеті(прапор URG).

Таблиця 2 – Характеристики сеансу.

Параметр	Опис
Hot	Число «hot» індикаторів
num_failed_logins	Число невдалих спроб входу
logged_ n	Успішний вхід
num_compromised	Доступ з адміністративними повноваженнями
root_shell	Число спроб доступу з правами адміністратора
num_root	Число операцій з файлами контролю доступу
num_file_creations	Кількість операцій створення файлу
num_access_files	Кількість операцій на файлах управління доступом
num_compromised	Кількість скомпрометованих статусів
is_hot_login	Приналежність користувача до «hot» списку
is_guest_login	Чи є користувач гостем

8

Таблиця 3 – Статистика з'єднання за 2 секунди.

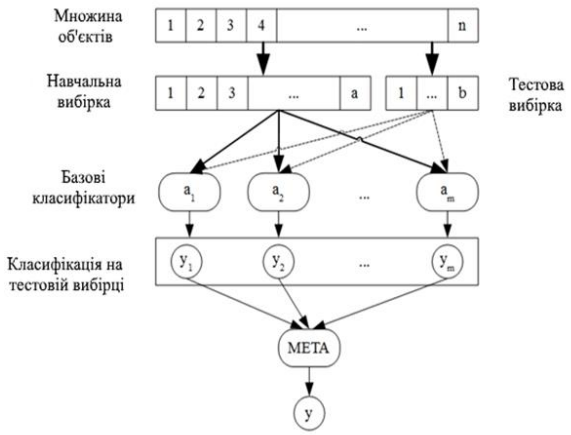
Параметр	Опис
Count	Число з'єднань із співпадаючим хостом
serror_rate	% з'єднання з помилкою "SYN"
Service	% з'єднань з помилкою "REJ" / % з'єднань з однаковим початковим портом
rerror_rate	% з'єднань з однаковим сервісом
same_srv_rate	% з'єднань з різним сервісом
diff_srv_rate	Число з'єднань із співпадаючим сервісом
srv_count	Число з'єднань із співпадаючим сервісом
srv_serror_rate	% з'єднань з помилкою "REJ"
srv_rerror_rate	% з'єднань з помилкою
srv_diff_host_rate	% з'єднань з хостами, що розрізняються

Таблиця 4 – Характеристики файлової системи кластера(Lustre).

Параметр	Опис
num _ exports	Кількість експорту на MDT - це клієнти, у тому числі інші сервери Lustre
Stats	Перераховані клієнтські з'єднання по NID.
lock _ count	Кількість блокувань
pool.granted	Luster розподілений менеджер блокування(ldlm) надав блокування
grant _ rate	ldlm заблокований рівень відміни, що має назву 'GR'
cancel _ rate	Ldlm заблокований рівень відміни, що має назву 'CR'

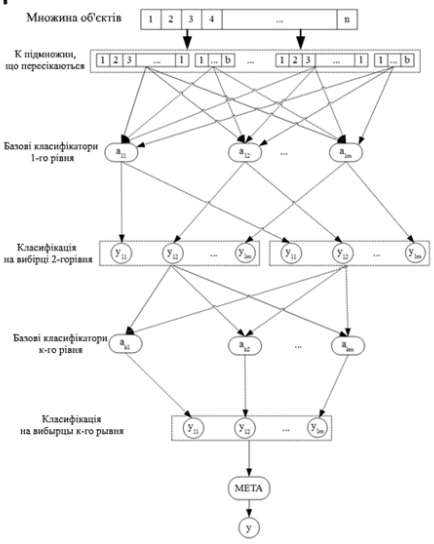
9

Схема класичного стекинга.



10

Схема модифікованого стекінга



11

Таблиця 5 – Класи атак в навчальній і тестовій вибірці.

Навчальна вибірка	Тестова вибірка
back, buffer_overflow, ftp_write, guess_passwd, imap, ipsweep, land, loadmodule, multihop, neptune, nmap, phf, pod, portsweep, rootkit, satan, smurf, spy, teardrop, warezclient, warezmaster, normal	guess_passwd, imap, ipsweep, land, loadmodule, multihop, neptune, nmap, phf, pod, portsweep, rootkit, satan, smurf, spy, teardrop, warezclient, warezmaster apache2, httptunnel, mailbomb, mscan, named, perl, processtable, ps, saint, sendmail, snmpgetattack, snmpguess, sqlattack, udpstorm, worm, xlock, xsnoop, xtermbuffer_overflow, Neptune, warezmaster, smurf, normal

Таблиця 6 – Класи атак в навчальній і тестовій вибірці

Навчальна вибірка	Тестова вибірка
back, neptune, pod, spy, normal	back, neptune, pod, smurf, spy, teardrop, normal

12

Кількість випробувань	Кількість вірних рішень		Кількість невірних рішень			
			Помилки I роду		Помилки II роду	
	Навч-на вибірка	Тестова вибірка	Навч-на вибірка	Тестова вибірка	Навч-на вибірка	Тестова вибірка
Наївний класифікатор Байеса	80,59	64,90	15,68	23,40	3,73	11,70
Класифікатор kNN	85,40	70,40	10,60	19,97	4,00	9,63
SVM	84,30	66,47	10,60	19,93	5,10	13,60
Дерева класифікації	85,70	72,80	8,10	15,60	6,20	11,60
Стандартний стекинг	86,64	71,38	3,09	15,82	10,27	12,8
Модифікований стекинг	92,01	84,19	4,7	10,49	3,29	5,32

Висновок

Основні результати виконаної роботи полягають в наступному:

1. Проведено аналіз існуючих методів і засобів моніторингу розподілених інформаційних систем, який виявив відсутність надання цілісної оцінки стану функціонування РКС.
2. Розглянуто особливості архітектури розподілених комп'ютерних систем, що дозволило виявити базові компоненти розподілених комп'ютерних систем, які потребують постійного моніторингу їх стану.
3. Визначено множину параметрів для оцінки стану кожного елемента системи, що дає можливість для подальшої оцінки функціонування системи в цілому та скоротити час навчання ансамблю класифікаторів на 30,79%.
4. Розроблено метод виявлення аномалій в мережі на основі стекінгу методів інтелектуального аналізу даних, який на відміну від існуючих алгоритмів виявлення аномалій дозволяють підвищити точність виявлення до 92,7% і знизити кількість помилкових другого роду до 4,05% за рахунок використання різнорідних методів інтелектуального аналізу даних і до навчання системи на даних підчас роботи системи моніторингу.