УДК 004.75

СИСТЕМА ЗАПОБІГАННЯ ВТОРГНЕНЬ НА ОСНОВІ МАШИННОГО НАВЧАННЯ

Яцків В.В.[1)], Драчук М. Ф.[2)], Боднар В.М.[3)]
Тернопільський національний економічний університет
[1)]д.т.н., доцент; [2)-3)]магістрант

I. Постановка проблеми

Сучасна інформаційна безпека стикається з низкою труднощів, серед яких слід назвати величезні потоки подій, зниження експертизи і брак персоналу. При цьому число атак зростає, незважаючи на вжиті заходи захисту. В даний час середній період невиявлення загроз становить близько 200 днів, що стає результатом реактивності використовуваних захисних засобів. Тому сьогодні, як ніколи, важливо застосовувати нові методи боротьби зі шкідливою активністю, найперспективнішим з яких є машинне навчання.

Найбільш поширеним ризиком для безпеки мережі є вторгнення, відмова в обслуговуванні або навіть проникнення в мережу. Зі зміною структури поведінки мережі необхідно перейти до динамічного підходу для виявлення та запобігання таких вторгнень. Багато досліджень були присвячені цій галузі, і існує загальне визнання, що статичні набори даних не захоплюють необхідні композиції трафіку та втручання [1].

Маючи багато точок доступу в типовій бізнес-мережі, дуже важливо, щоб був спосіб контролювати за певними ознаками потенційні порушення, зловмисні випадки і неминучі загрози [2]. Сьогоднішні мережеві загрози стають все більш і більш складними і здатні проникнути навіть у найнадійніші рішення безпеки.

В останнє десятиліття відбулися значні досягнення в технології машинного навчання, що дозволяють автоматизувати і передбачати в масштабах, які важко було уявити раніше.

Система запобігання вторгненню (IntrusionPreventionSystem, IPS) це вид мережевої безпеки, яка працює для виявлення та запобігання виявлених загроз. Системи запобігання вторгнень постійно перевіряють мережу, шукають можливі зловмисні випадки та отримують інформацію про них. IPS повідомляє про ці події системним адміністраторам і здійснює попереджувальні дії, такі як закриття точок доступу та налаштування брандмауерів для запобігання майбутнім атакам. Рішення IPS можуть використовуватися для виявлення проблем з політикою корпоративної безпеки, стримування працівників і гостей мережі від порушення правил, що містяться в цій політиці [3, 4].

II. Мета роботи

Метою дослідження є розробка структури системи запобігання вторгнень з використанням глибокої моделі машинного навчання для виявлення аномалії в наборах даних.

III. Структура системи запобігання вторгнень

Число інцидентів, пов'язаних з інформаційною безпекою, за даними провідних аналітичних агентств, постійно зростає. Фахівці, відповідальні за захист інформації, відзначають зростаючу активність зовнішніх зловмисників, що використовують останні розробки в області нападу, що намагаються проникнути в корпоративні мережі.

Для мінімізації загроз інформаційній безпеці необхідне впровадження багаторівневої системи захисту інформації. Першим рівнем забезпечення інформаційної безпеки, яка блокує несанкціонований доступ хакерів в корпоративну мережу, є міжмережевий екран. Залежно від технології обробки інформації в організації міжмережеві екрани можуть поставлятися в різній комплектації і забезпечувати різний функціонал.

Основним функціоналом міжмережевих екранів є розмежування і контроль доступу, трансляція адрес, приховування топології обчислювальної мережі від зовнішнього світу і організація нейтральних зон. Однак необхідно відзначити, що міжмережеві екрани головним чином здійснюють фільтрацію і аналіз трафіку на третьому і четвертому рівнях моделі OSI, і лише обмежено — на більш високих рівнях.

Для моніторингу та боротьби з атаками і несанкціонованою мережевою активністю рекомендується використовувати спеціалізовані продукти мережевої системи виявлення та запобігання атак (IDS / IPS). Дані системи дозволяють відстежити і зареєструвати спроби несанкціонованої мережевої активності і опційно блокувати атаки в режимі реального часу. Схема роботи мережі з використанням системи запобігання вторгнень зображена на рисунку 1.

З метою покращення процесу виявлення вторгнень та зменшенням кількості помилкових спрацювань, не використовуючи статичний шаблон виявлення аномалій, доцільно реалізувати систему з використанням модуля машинного навчання.

На відміну від традиційних методів виявлення чогось поганого, що спираються на боротьбу з чимось знайомим, машинне навчання дозволяє нам розпізнати те, що ще невідомо. Щоб зробити це, на вхід моделі / алгоритму необхідно подати вхідні дані (багато даних), на яких модель буде навчатися. Після навчання моделі можна подавати на вхід нові дані, і вона почне виявляти в них шукане.

Машинне навчання базується на трьох ключових елементах: датасет, ознаки, Алгоритми / моделі.

Датасет. Щоб навчити модель розпізнавати щось погане, їй на вхід треба подати великі обсяги даних. Це може бути інтернет-трафік, мережеві потоки, логи, поштові повідомлення, активність користувача і багато іншого. Чим більше і різноманітніші навчальні дані, тим точніше буде результат передбачення. Щоб виявляти шкідливі домени, треба вивчати сотні мільярдів і трильйони DNS-запитів. Від якості датасета залежить ефективність машинного навчання.

Ознаки. Це те, що ми шукаємо в датасеті. Наприклад, доменне ім'я, IP-адрес, тривалість мережевий сесії, протокол, час дня і т.д. В залежності від розв'язуваної задачі можуть бути сотні різних ознак.

Алгоритми / моделі. Знайти за певними ознаками шукане в датасеті можна різними способами, вибір яких залежить від безлічі параметрів. Правильний вибір алгоритму або моделі - це завжди баланс між швидкістю роботи, акуратністю передбачення і складністю моделі.

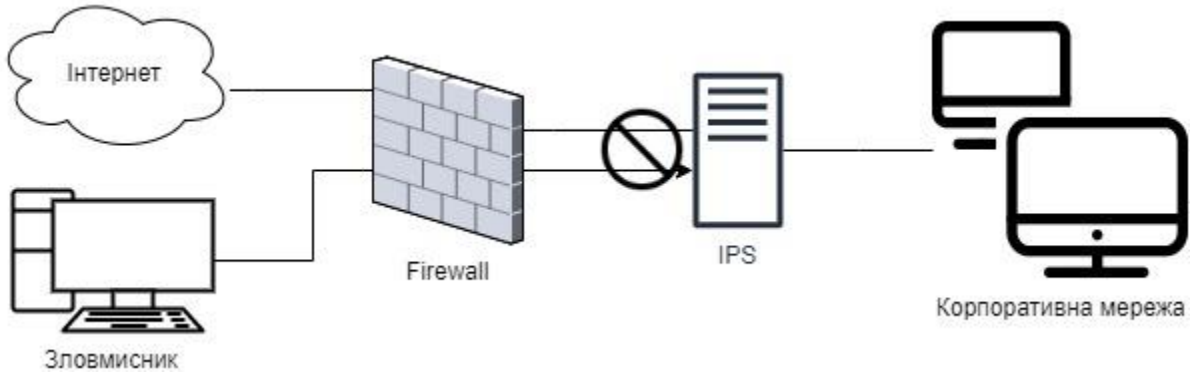

Рисунок 1 – Схема роботи мережі з використанням системи запобігання вторгнень

Щоб зробити роботу системи ефективною вона повинна автоматично опрацьовувати відомі види атак, а також вміти розпізнавати схожі до них атаки. Розроблену модель машинного навчання необхідно навчити розрізняти різницю між нормальними та атакуючими пакетами в мережі.

Внесення в модель машинного навчання визначеного набору даних з широким діапазоном видів вторгнень дозволить побудувати контрольовану модель та зменшити кількість помилкового виявлення. Для забезпечення високої ефективності виявлення вторгнень, набір даних повинен бути не тільки достовірний, але також таким, що модифікується, розширювальний та відтворний.

Висновок

У роботі розроблено структуру системи запобігання вторгнень з використанням моделі глибокого машинного навчання для виявлення аномалії в мережевому трафіку.

Список використаних джерел

1. Дубов Д.В. Кіберпростір як новий вимір геополітичного суперництва: монографія / Д.В. Дубов. – К.: НІСД, 2014. – 328 с.
2. Бурячок В.Л. Основи формування державної системи кібернетичної безпеки: Монографія. – К.: НАУ. – 2013. – 432 с.
3. Chio, C., Freeman, D. MachineLearningandSecurity: ProtectingSystemswithDataandAlgorithms." O'ReillyMedia, Inc.", 2018.
4. Грайворонський М.В., Новіков О.М. Безпека інформаційно-комунікаційних систем / М.В. Грайворонський – К.: Видавнича група BHV, 2009 – 608 с.
5. Завада А. Аналіз сучасних систем виявлення атак і запобігання вторгненням / А. Завада, О. Самчишин, В. Охрімчук. – Житомир: Збірник наукових праць ЖВІ НАУ, 2012. – 106 с.