КІБЕРБЕЗПЕКА: освіта, наука, техніка
CYBERSECURITY: EDUCATION, SCIENCE, TECHNIQUE
№ 4 (12), 2021
ISSN 2663 - 4023
DOI 10.28925/2663-4023.2021.12.132142
УДК 004.8:[004.056:007]

Іваніченко Євген Вікторович
кандидат технічних наук, доцент кафедри ком'ютерних наук і математики
Київський університет імені Бориса Грінченка, м. Київ, Україна
ORCID ID: 0000-0002-6408-443X
y.ivanichenko@kubg.edu.ua

Сабліна Милана Андріївна
викладач кафедри ком'ютерних наук і математики
Київський університет імені Бориса Грінченка, м. Київ, Україна
ORCID:0000-0001-9452-1867
m.sablina@kubg.edu.ua

Кравчук Катерина Володимирівна
студентка факультету інформаціонних технологій та управління
Київський університет імені Бориса Грінченка, м. Київ, Україна
ORCID ID: 0000-0002-3589-8784
kvkravchuk.fitu16@kubg.edu.ua

ВИКОРИСТАННЯ МАШИННОГО НАВЧАННЯ В КІБЕРБЕЗПЕЦІ

Анотація. Актуальність теми - інтеграція технологій машинного навчання в системи кібербезпеки. Ознайомившись з технічною літературою було сформульовано основні технології машинного навчання які реалізуються в організації кібербезпеки. Ознайомлено з основним типом штучної нейронної мережі, яка використовуються під час попередження і виявлення кіберзагрози та встановлено, що основною для розгляду загального застосування технологій машинного навчання є штучні нейронні мережі, засновані на багатошаровому персептроні із зворотним поширенням помилок. Запропоновано використовувати індикатори компромісних кібератак як початкової інформації для систем автоматичного машинного навчання . Акцентовано увагу на основні типи даних, які можуть бути використані підсистемами спостерігання засобів захисту інформації та організації кібербезпеки для виконання завдань і попередження, класифікації та прогнозування подій кібербезпеки. За результатами аналізу визначено основні проблемні напрямки щодо їх реалізації в системах інформаційної безпеки. Проблему використання машинного навчаня (ML) в кібербезпеці складно вирішити, оскільки досягнення в цій області відкривають багато можливостей, з яких складно обрати дієві засоби реалізації та прийняття рішень. Окрім цього, ця технологія також може використотуватись хакерами для створення кібератаки. Метою дослідження є реалізація машинного навчання в технології інформаційної безпеки та кібербезпеки, та зобразити модель на основі самонавчання.

Ключові слова: машинне навчання; кібербезпека; нейронні мережі; кібератака; кіберзахист з використанням машинного навчання.

ВСТУП

Кібербезпека завойовує дедалі більше уваги, кожна кібербезпека набуває дедалі більше уваги щороку. Кількість кібератак значно зросла з 2009 року завдяки оцифруванню всього в сучасному світі. Відповідно до циклу Гартнера Хайпа [1], машинне навчання (ML) представляє великий інтерес у світі технологій. ML полягає у розумній поведінці в системі, включаючи сприйняття, міркування, навчання, спілкування та працювати в складному інформаційному середовищі [2]. Такий широкий інтерес до ML пов'язаний з двома критичними факторами : По-перше, він може автоматизувати процеси, які раніше вимагали участі людини.

Наприклад, контроль роботизованих механізмів у виробництві (тобто ML бере на себе людські обов'язки).

По-друге, він може швидко обробляти та аналізувати величезні обсяги інформації та розраховувати варіанти використовуючи безліч змінних. У цих сферах ML забезпечує якісно кращі результати порівняно з людськими можливостями.

ML надає великий функціонал з організації кібербезпеки. Поточні реалізації широко використовуються в системах IDS, системи «пісочниці» та інших сфер кібербезпеки - розвідки загроз для передової автоматизованої цифрової криміналістики. Насправді 71% американського бізнесу планує використовувати ML у своїх інструментах кібербезпеки у 2019 р. [3]

Становили понад третину (36%) [3] організації зазнали шкідливих кібератак у 2018 році. Більшість (83%) [3] це визнають кіберзлочинці, та використовують ML для створення кібератаки на організації. Проблема використання ML у кібербезпеці важко вирішити, оскільки досягнення в цій галузі дають стільки можливостей, що складно знайти хороші та вигідні приклади використання для впровадження та прийняття рішень. Більше того, важко визначити, наскільки безпечна система безпеки, яка використовується у виробництві, і як захистити організацію від кібератак, що проводяться через ML. Головною метою, є робота з дослідження та використання ML в кібербезпеці та дослідити випадки використання, пов'язані з використанням супротивником ML у кібератаках.

ОСНОВНІ ВИЗНАЧЕННЯ

ML - це процес, за допомогою якого машини вчаться на наданій інформації, будуючи логіку та прогнозуючи вихід для даного входу [4]. ML має три підкатегорії: навчання під наглядом, без нагляду навчання та підкріплення навчання [5]. Під контролем навчання використовується набір даних, позначаються символом правильні відповіді для вивчення. Такі мітки визначають характеристики кожного набору данних. Після того, як модель виконає самонавчання, вона може почати прогнозувати або приймати рішення щодо нових даних чи ситуацій, які їй передаються.. При навчанні без нагляду немає необхідності в такому позначеному наборі даних. Як тільки модель отримає набір даних, вона автоматично знаходить шаблони та взаємозв'язки, створюючи в ній кластери. Однак такий тип навчання не може нічого передбачити. Коли додаються нові дані, модель призначає їх одному з існуючих кластерів або створює новий. Підкріплення навчання - це здатність системи взаємодіяти з навколишнім середовищем та визначати найкращі результати.

Система «винагороджується» або «карається» балом за правильну або неправильну відповідь, а на основі отриманих позитивних балів винагороди модель формується автоматично Подібним чином, пройшовши навчання, він готується прогнозувати нові дані, представлені йому.

Глибоке навчання (DL) - це клас алгоритмів ML [6], який використовує кілька шарів для поступового вилучення функцій вищого рівня з вихідного входу. Основні відмінності між ML та DL полягають у наступному: алгоритми ML майже завжди вимагають структурованих даних, тоді як мережі DL покладаються на рівень штучних нейронних мереж (ANN).

Часто в ML, втручання людини необхідне для отримання подальших результатів з більшим набором даних, тоді як в DL це не потрібно. Однією з основної концепцій DL є ANN.

ISSN 2663 - 4023

ANN - це модель, яка побудована на принципі організації та функціонування мозку людини (тобто мереж нервових клітин у живому організмі). Іншими словами, алгоритм нейронної мережі намагається створити функцію, яка відображатиме вхідні дані до бажаних результатів. Нейронні мережі (NN), як правило, організовані шарами (рис. 1). Шари складаються з безлічі взаємопов'язаних «вузлів», що містять «функцію активації». Шаблони подаються в мережу через "вхідний рівень", який передає дані одному або декільком "прихованим шарам", де фактична обробка здійснюється через систему зважених "з'єднань". Потім приховані шари посилають на "вихідний рівень", де відповідь - вихідний.

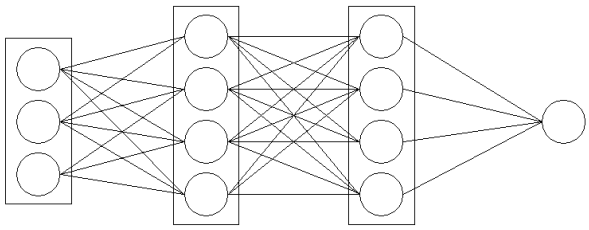

Рис. 1. Приклад нейронної мережі

Наприклад, при обробці зображень нижчі шари можуть ідентифікувати краї, тоді як вищі шари можуть ідентифікувати такі поняття, що стосуються людини, такі як цифри, літери чи грані. Якщо NN мають більше двох прихованих шарів, вони називаються глибокими нейронними мережами (DNN) [7]. DNN використовується для розпізнавання зображень, розпізнавання мови та інших програм. Окрім цього, створені та існуючі технології для створення нових фотографій, які виглядають принаймні поверхнево автентичними для спостерігачів - людей завдяки багатьом реалістичним характеристикам. Наприклад, відома спроба синтезувати фотографії собак, яка вводила експерта в оману, вважаючи, що вони справжні [8].

Це приклад технології, званої генеративною загальною мережею (GAN), алгоритмом ML некерованого навчання, побудованого на поєднанні двох NN: одна мережа G (генератор) генерує {{1} } нові приклади, а одна мережа D (дискримінатор) намагається класифікувати приклади як реальні, так і хибні [9].

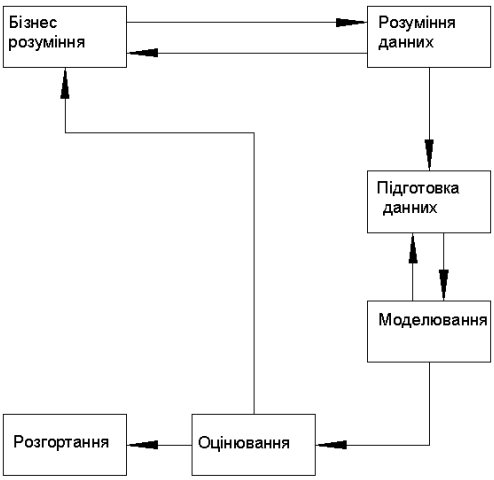

Рис. 2. CRISP-DM процес видобутку даних

Одним із процесів, який нерозривно пов'язаний з ML та DL, є видобуток даних. Використовуючи видобуток даних у великих наборах даних, можна виявити нові закономірності, використовуючи методи статистики та систем баз даних [10].

Міжгалузевий стандартний процес видобутку даних (CRISP-DM) описує міжгалузевий процес видобування даних [11]. CRISP-DM розбиває процес на шість основних фаз: розуміння бізнесу, розуміння даних, підготовка, моделювання, оцінка та розгортання даних (рис. 2). Перші дві фази з'єднані між собою. Їх головна мета - визначити цілі проекту, поставити завдання для ML та зібрати дані. Ці цілі можуть бути скориговані на основі даних. Наступний етап стосується процесу роботи з даними: очищення даних, комбінування даних, якщо потрібно, та форматування даних.

На етапі моделювання до даних застосовуються різні методи моделювання. Моделі будуються, і їх параметри доводяться до оптимальних значень. Через особливі вимоги до даних у різних моделях ми можемо повернутися до фази підготовки даних. На етапі оцінки модель вже побудована, і отримані кількісні оцінки її якості. Перш ніж впроваджувати цю модель, нам слід переконатися, що ми досягли всіх бізнес-цілей.

Залежно від вимог фаза розгортання може бути простою (наприклад, підготовка остаточного звіту) або складною (наприклад, автоматизація процесу аналізу даних для вирішення бізнес-проблем).

ВИКОРИСТАННЯ ML ДЛЯ ЗАХИСТУ

Сфера використання ML у кібербезпеці величезна, починаючи з виявлення аномалій та підозрілої або незвичної поведінки і закінчуючи виявленням уразливостей нульового дня та виправленням відомих. Ділект [12] представив найбільш вичерпний огляд застосувань методів ML.

Reathi та Malathi [3] представили набір алгоритмів ML, навчених набору даних виявлення вторгнень NSL-KDD для виявлення зловживання. Тим часом Бучак та інші [6] зосереджені на виявленні вторгнення в мережу за допомогою ML.

Меліхер та співавтори запропонував використовувати NN для перевірки стійкості до підбору паролів. Вони стиснули модель до сотень кілобайт та розробили інструмент JavaScript на стороні клієнта. Подібний експеримент був проведений Ciaramella та інші. Для активної перевірки надійності паролів вони використовують NN, такі як багатошаровий персептрон (MLP) та одношарові персептрони (SLP). Примітно, що MLP забезпечують кращі результати, ніж SLP, при тестуванні наборів даних.

Більше того, кількість шарів дорівнює 10, і таким чином отримується кращий результат. Аналітика поведінки користувачів та сутності (UEBA) використовує можливості ML для аналізу журналів поведінки та мережевого трафіку в режимі реального часу та належного реагування у разі атаки . Цей процес здійснюється шляхом змушення користувача ввійти знову, блокуючи атаку або оцінюючи рівень ризику та попереджаючи службовців інформаційної безпеки компанії, щоб вони могли вжити необхідних заходів.

Більшість методів ML та DL, такі як навчання ансамблю, кластеризація та дерево рішень [18], використовуються для виявлення зловживань, аномалій та гібридного кіберпроникнення.

Як згадується в Офіційному блозі Євгена Касперського, Касперський виявляє 99% кіберзагроз за допомогою технології ML. Інтервал часу між розкриттям підозрілої поведінки на захищеному пристрої та випуском відповідної нової "таблетки" триває в середньому 10 хвилин.

DARPA співпрацювала з BAE Systems, щоб розробити систему, яка дозволяє нам налаштовувати датчики та застосовувати захисні заходи "на швидкості машини". Ця ініціатива називається програмою CHASE,що розшифровується як Кіберполювання в масштабі, прагне розробити автоматизовані засоби виявлення та характеристики нових векторів атак, збору правильних контекстних даних та розповсюдження захисних заходів як всередині підприємств, так і між ними [2].

Кібератаки, здійснені хакерами, стосуються загальної думки про гучні новини. Інформація, зібрана із соціальних медіа, може допомогти передбачити подібні випадки, використовуючи методи NLP та ML [11].

Більше того, ми можемо використовувати ML для ідентифікації автора програми. Рейчел Грінштадт та Ейлін Каліскан розробили систему, яка може "деанонімізувати" програмістів , аналізуючи вихідний код або скомпільовані двійкові файли. Визначити розробника шкідливого програмного забезпечення набагато простіше.

Інший спосіб контролю систем та мереж на предмет зловмисних дій чи порушень політики - це система виявлення вторгнень (IDS). Система запобігання проникненню (IPS) - це система, пов’язана з IDS; ці системи виконують виявлення вторгнень і зупиняють виявлені випадки.

Обидві системи використовують контрольовані та неконтрольовані методи ML для виявлення точкових аномалій, контекстних аномалій та колективних аномалій [12].

Основне завдання брандмауерів - забезпечити мережеву систему безпеки, яка контролює та контролює вхідний та вихідний мережевий трафік. Брандмауери дозволяють або блокують трафік, порівнюючи його характеристики із заздалегідь визначеними шаблонами (тобто правилами брандмауера). У своїй роботі Ucar та Ozhan представили результат автоматичного виявлення аномалій у сховищі правил брандмауера на основі ML та високопродуктивних обчислювальних методів, таких як Naive Bayes, kNN, таблиця рішень та HyperPipes. Всі шість правил брандмауера з даних 93 правил були виявлені системою та перевірені експертами як аномалія. Брандмауери фільтрують вміст між серверами, а також є рішення, спеціально призначене для вмісту веб-додатків. Брандмауер веб-додатків (WAF) розгортається перед веб-програмами; він аналізує двонаправлений веб-інтерфейс (HTTP) трафіку та виявляє та блокує будь-що зловмисне. WAF запобігає використанню вразливостей у веб-програмах із зовнішніх загроз. Для реалізації такої функціональності у WAF розробники використовують регулярні вирази, токени, поведінковий аналіз, аналіз репутації та технології ML.

Серед методів ML спеціальні методи прогнозування також можуть бути використані для запобігання втраті / витоку даних (DLP) для зменшення ризику порушень або витоків. Програмні рішення DLP дозволяють нам встановлювати ділові правила, які класифікують конфіденційну та не конфіденційну інформацію, щоб їх не можна було розкрити зловмисно або випадково несанкціонованими кінцевими користувачами. Цей процес можна зробити, використовуючи контрольовані алгоритми навчання та два типи прикладів: позитивні приклади (тобто вміст, який потрібно захищати) та контрприклади (тобто документи, подібні до позитивного набору, але не повинні захищатися).

ВИКОРИСТАННЯ ML ПІД ЧАС КІБЕРАТАК

У цьому розділі описано, як кібератака може досягати успіху за допомогою ML. Автоматизоване сканування вразливості - одне з найбільш очевидних і поширених завдань в кібератаці. Наприклад, CSRF міститься лише у 5% додатків, як повідомляється у Топ-10 OWASP 2017 року, оскільки більшість фреймворків включають захист CSRF.

Відповідно, Calzavara et al. представив Mitch, перший на основі ML інструмент для виявлення чорної скриньки CSRF, який дозволяє ідентифікувати 35 нових вразливостей CSRF на 20 веб-сайтах з 10 000 веб-сайтів Alexa і три раніше не виявлені вразливості CSRF на виробничому програмному забезпеченні, які вже були проаналізовані за допомогою ультрасучасного інструменту. Мітч - це двійковий класифікатор, що позначає чутливі або нечутливі запити за допомогою випадкового лісового алгоритму на 49-вимірному просторі об'єктів.

Порівняно з евристичними класифікаторами BEAP та CsFire, Мітч демонструє найкращі показники F1 і точність (таблиця 1). Маркетологи використовують методи ML для профілювання. Trustwave випустила інструмент розвідки з відкритим кодом, який використовує розпізнавання обличчя для автоматичного відстеження тем у соціальних мережах. Розпізнавання обличчя допомагає цьому процесу, видаляючи хибні спрацьовування в результатах пошуку, прискорюючи перегляд даних для оператора-людини.

Таблиця 1

Заходи дійсності випробуваних класифікаторів (BEAP, CsFire, Mitch)

Класифікатор	Точність	Час відповіді	F1
BEAP	0.29	0.88	0.44
CsFire	0.24	0.95	0.32
Mitch	0.79	0.65	0.71

Використовуючи зібрані дані про ціль, зловмисник може підключити жертву за допомогою спеціально створених фейкових новин. Інструменти ML допомагають ідентифікувати фальшиві новини, але для цього дослідники підтверджують, що найкращий спосіб для ML - навчитися створювати самі фальшиві новини . Таким чином, вони створили модель для керованого генерування тексту під назвою Гровер. У процесі дослідження використовувались чотири класи статей: людські новини, машинні новини, людська пропаганда та машинна пропаганда. Робітники на Amazon Mechanical Turk оцінив кожну статтю, включаючи загальну надійність. У випадку пропаганди оцінка зросла з 2,19 (з 3) на статті, створені вручну, до 2,42 статті, створені машиною.

SNAP_R був представлений у DEFCON 24. SNAP_R - це перший у світі автоматизований наскрізний генератор кампаній для підводного фішингу для Twitter [6]. Хоча попередні інструменти базувались на моделях з ланцюгами Маркова, SNAP_R базувався на періодичній NN з архітектурою LSTM. Використання Twitter як середовища пропонує деякі переваги для автоматичного створення тексту. Наприклад, обмеження тривалості повідомлення зменшує ймовірність граматичних помилок. Більше того,посилання на Twitter часто скорочуються, що дозволяє маскувати шкідливі домени. Це, у свою чергу, суттєво збільшило рівень успіху з 5–14% на інструментах, заснованих на ланцюжках Маркова, до 30–66%, що порівняно з 45% для ручного фішингу [9].

У більшості випадків зловмисники не знають алгоритму виявлення шкідливого програмного забезпечення, але можуть зрозуміти особливості, які він використовує, за допомогою ретельно розроблених тестових випадків в алгоритмі чорної скриньки. MalGAN - це генеративний алгоритм, що базується на суперницькій мережі, який генерує приклади суперницького шкідливого програмного забезпечення, які можуть обійти моделі виявлення на основі чорного ящика ML. Це може знизити рівень виявлення майже до нуля і ускладнити роботу методу оборони на основі перепідготовки проти змагальних прикладів [40]. Архітектура MalGAN показана на рис. 3 [4].

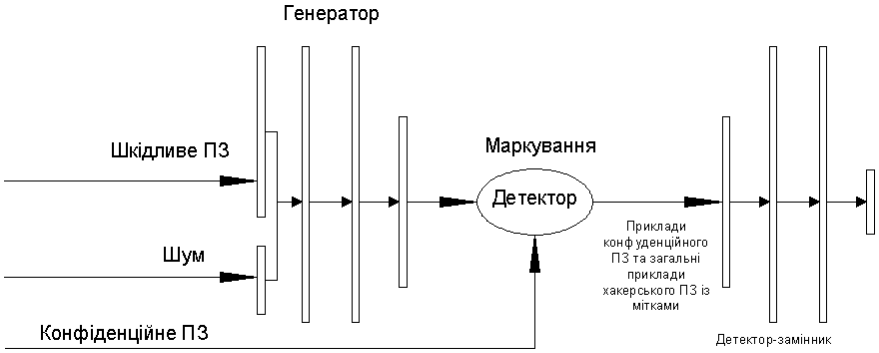

Рис. 3. Архітектура MalGAN

Генератор використовує вектор характеристик шкідливого програмного забезпечення та вектор шуму, щоб перетворити перший у свою загальну версію. Замінник-детектор використовується для встановлення детектора чорної скриньки та надання інформації про градієнт для навчання генератора. Обидві мережі представлені як багатошарові ANN-адреси з прямим поданням. До змагальних прикладів, протестованих проти детектора чорної скриньки, за різними методами ML, навченими 160-мірним двійковим векторам ознак, що представляють системні виклики API, логістична регресія, дерева рішень, векторні машини підтримки та багатошаровий персептрон, а також голосування ансамбль цих алгоритмів. Розробники програмних комплексів зловмисного програмного забезпечення перевчають детектори після вивчення таких невиявлених прикладів, але MalGAN потребує лише однієї епохи перекваліфікації, щоб отримати 0% справжнього позитивного показника. Каваї та ін. пізніше запропонував деякі покращення продуктивності [10].

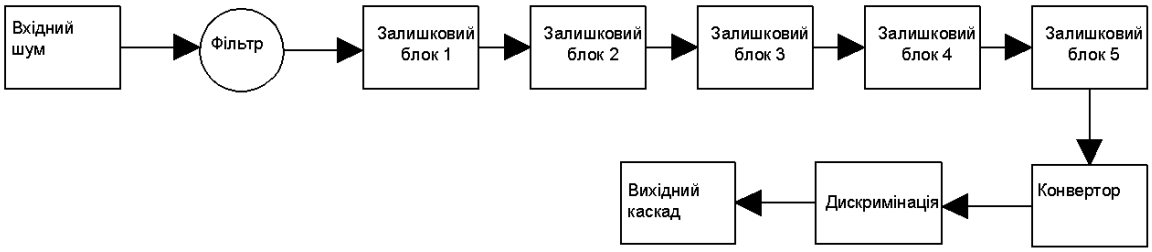

Рис. 4. Архітектура генератора PassGAN

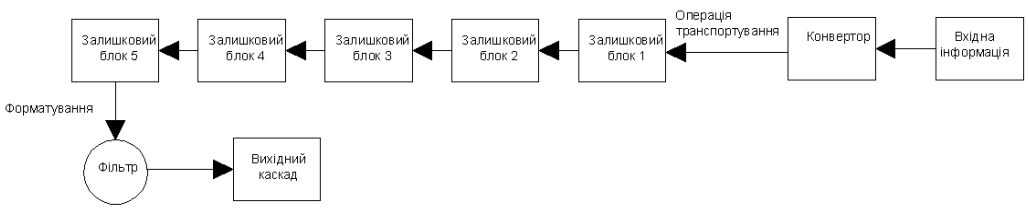

Рис. 5. Дискримінаційна архітектура PassGAN

Іншим прикладом використання GAN в кібербезпеці є атака вгадування пароля. Існує новий спосіб генерації припущень щодо паролів на основі DL та генеративних змагальних мереж відомий як PassGAN. Ключова відмінність у цьому підході полягає в тому, що DL не потребують апріорних знань про структуру паролів, на відміну від підходів, заснованих на правилах, моделях Маркова та FLA. PassGAN використовує вдосконалене навчання GAN Wasserstein (IWGAN) Гулраджані та ін. за допомогою оптимізатора ADAM. Генератор і дискримінатор у PassGAN побудовані на основі ResNets. Архітектура генератора та дискримінатора показана на рис. 4 та рис. 5, тоді як подання залишкового блоку показано на рис 6.

Рис.6. Архітектура залишкового блоку в PassGAN

Для максимальної ефективності зловмисники, швидше за все, використовують кілька інструментів злому паролів, такі як HashCat [7], John the Ripper [48], PCFG [49], OMEN [50] та FLA [5], щоб поєднати різні методи атаки. Наприклад, комбінуючи вихідні дані PassGAN з результатами HashCat Best64 [5], дослідники змогли вгадати від 51% до 73% додаткових унікальних паролів у порівнянні лише з HashCat [7].

Традиційні бот-мережі чекають команд від C&C, але зараз зловмисники використовують автоматизацію для самостійного прийняття рішень. Дослідники Fortinet прогнозували, що кіберзлочинці замінять ботнети на розумні кластери зламаних пристроїв, які називаються hivenets, тип атак, який здатний використовувати самонавчання на основі націлювання на вразливі системи з мінімальним наглядом [5].

На початкових етапах атаки зловмисники часто стикаються з проблемою обходу капчі. Suphannee та інші розробив недорогу атаку, яка використовує технології DL для семантичних анотацій зображень. Для вирішення завдань системі потрібно близько 19 секунд на виклик, з точністю 70,78% для reCaptcha [9] та 83,5% для captcha зображення Facebook. Система повинна автоматично визначити, які з поданих зображень семантично схожі на зразкові зображення. По-перше, система збирає інформацію для всіх зображень за допомогою Google Reverse Images Search (GRIS) [55]; Clarifai [6], який побудований на деконволюційних мережах [7]; TDL [8], в основі якої лежать глибокі машини Больцмана; NeuralTalk та Caffe. Lalі, якщо підказки не надано, система шукає зразок зображення в маркованому наборі даних для, щоб отримати його, якщо це можливо.

КІБЕРАТАКА З ПОВНОЮ ПІДТРИМКОЮ ML

Як зазначалось у попередньому розділі, кібератаки, що працюють на основі ML, не є гіпотетичною концепцією майбутнього. У цьому розділі описано, як можна здійснити автоматизовану кібератаку за допомогою ML. Ми розглянули два сценарії етапів озброєння та доставки: По-перше, у випадку вторгнення без людей зловмисники можуть використовувати подібний інструмент, але використовувати інформацію, надану Шоданом та Мітчем, замість функцій, отриманих за допомогою комп'ютерного зору. По-друге, зловмисники можуть використовувати соціальну інженерію, використовуючи інструменти для профілювання та фішингу, описані в попередньому розділі, та

створюючи посилання на байти, щоб заразити жертву. Для автоматизованої генерації експлоїтів противники можуть використовувати фреймворк angr з відкритим кодом, розроблений Shellphish, та об'єднати його з MalGAN в обхід оборонних систем. На етапі після експлуатації зловмисники можуть підбирати викрадені паролі, використовуючи PassGAN. Найновіший метод - використання інтелектуальних методів ухилення, запропонованих дослідниками Darktrace, і подальше самопоширення з низкою автономних рішень. Як демонструють ці приклади, ML може допомогти хакерам на кожному етапі атаки. З розвитком рівня інфраструктури кіберзлочинців, для поглибленої атаки не потрібні практичні роботи на клавіатурі, як це зараз.

ВИСНОВОК

Запроваджуючи систему на основі ML, ми повинні пам’ятати, що ML не є панацеєю. Жодна система не є безпечною. За певних умов ML захищає вразливі місця та створює нові прогалини. ML можна порівняти з собакою: "Машинне навчання може робити все, до чого можна навчити собаку, але ви ніколи не до кінця впевнені, до чого ви навчили собаку". Слід також зазначити наслідки, які може принести більш активне впровадження МЗ: по-перше, автоматизація та наслідком втрата людських робочих місць, по-друге, неминучий конфлікт із існуючою законодавчою базою, наприклад, при використанні технологій для запобігання кіберзлочинності чи кібертероризму.

У такій ситуації обвинувачений притягується до злочинів, які ще не вчинені, які не регулюються жодною правовою нормою. Більше того, частина інформації, яку дізнається ML, може бути приватною або конфіденційною, що порушує законодавство деяких країн. Подібним чином, низька якість або неадекватна кількість ML у кібербезпеці даних, що базуються на прогнозах, може призвести до неправильних рішень та непоправних помилок.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Ciaramella, P. D’Arco, A. De Santis, C. Galdi, R. Tagliaferri. (2006). Neural Network Techniques for Proactive Password Checking. *IEEE Transactions on Dependable and Secure Computing*, *3*(4), 327-339.
2. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin, A. Courville. (2017). Improved training of wasserstein GANs. In *Proc. of the 31st International Conference on Neural Information Processing Systems*, (pp. 5769-5779).
3. *Shodan search engine*. (б. д.). Shodan Search Engine. https://www.shodan.io/
4. Скрыпников, А. В., Денисенко, В.В., Саранов, И.А. (2020). Использование методов машинного обучения при решении задач информационной безопасности. *Воронежский государственный университет инженерных технологий*, 4, 69–79.
5. Le Roux, N., Bengio, Y. (2008). Representational power of restricted Boltzmann machines and deep belief networks. *Neural computation*, *20*(6), 1631-1649.
6. Sharma, B., Mangrulkar, R. (2019). Deep learning applications in cyber security: a comprehensive review, challenges and prospects. *International Journal of Engineering Applied Sciences and Technology*, *4*(8), 148-159
7. Ranzato, M.A., Boureau, Y.L., Cun, Y.L. (2008). Sparse feature learning for deep belief networks. *Advances in neural information processing systems*, (pp. 1185-1192).
8. Mirkin, B. G. (2011). *Core concepts in data analysis: Summarization, correlation and visualization*. Springer Science & Business Media.
9. https://www.theverge.com/2018/1/12/16882408/google-racist-gorillas-photo-recognition-algorithm-ai
10. https://www.kaspersky.com/enterprise-security/wiki-section/products/sandbox
11. Kazennov, A. M. (2010). Basic concepts of CUDA technology. *Computer Research and Modeling*, *2*(3), 295–308. https://doi.org/10.20537/2076-7633-2010-2-3-295-308
12. Лоскутов, А. (Ред.). (2003). *Нейросетевые алгоритмы прогнозирования и оптимизации систем*. Наука и Техника.

КІБЕРБЕЗПЕКА: освіта, наука, техніка
CYBERSECURITY: EDUCATION, SCIENCE, TECHNIQUE
№ 4 (12), 2021
ISSN 2663 - 4023

Yevhen V. Ivanichenko
Candidate of Technical Sciences, Associate Professor of chair of Computer Science and Mathematics
Borys Grinchenko Kyiv University, Kyiv, Ukraine
ORCID ID: 0000-0002-6408-443X
y.ivanichenko@kubg.edu.ua

Mylana A. Sablina
Lecturer of Chair of Computer Science and Mathematics
Borys Grinchenko Kyiv University, Kyiv, Ukraine
ORCID ID: 0000-0001-9452-1867
m.sablina@kubg.edu.ua

Kateryna V. Kravchuk
Student of the Faculty of Information Technology and Management
Borys Grinchenko Kyiv University, Kyiv, Ukraine
ORCID ID: 0000-0002-3589-8784
kvkravchuk.fitu16@kubg.edu.ua

USE OF MACHINE LEARNING IN CYBER SECURITY

Abstract. The urgency of the topic is the integration of machine learning technologies into cybersecurity systems. After getting acquainted with the technical literature, the main technologies of machine learning that are implemented in the organization of cybersecurity were formulated. Acquainted with the main type of artificial neural network used in the prevention and detection of cyber threats and found that the main to consider the general application of machine learning technologies are artificial neural networks based on a multilayer perceptron with inverse error propagation. It is proposed to use indicators of compromise cyberattacks as initial information for automatic machine learning systems. Emphasis is placed on the main types of data that can be used by surveillance subsystems for information security and cybersecurity to perform tasks and prevent, classify and predict cybersecurity events. According to the results of the analysis, the main problem areas for their implementation in information security systems are identified. The problem of using machine learning (ML) in cybersecurity is difficult to solve, because advances in this area open up many opportunities, from which it is difficult to choose effective means of implementation and decision-making. In addition, this technology can also be used by hackers to create a cyber attack. The purpose of the study is to implement machine learning in information security and cybersecurity technology, and to depict a model based on self-learning

Key words: machine learning; cybersecurity; neural networks; cyberattack; cybersecurity using machine learning.

REFERENCES

1. Ciaramella, P. DArco, A. De Santis, C. Galdi, R. Tagliaferri. (2006). Neural Network Techniques for Proactive Password Checking. IEEE Transactions on Dependable and Secure Computing, 3(4), 327-339.
2. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin, A. Courville. (2017). Improved training of wasserstein GANs. In Proc. of the 31st International Conference on Neural Information Processing Systems, (pp. 5769-5779).
3. Shodan search engine. (b. d.). Shodan Search Engine. https://www.shodan.io/
4. Skrыpnykov, A. V., Denysenko, V.V., Saranov, Y.A. (2020). Yspolzovanye metodov mashynnoho obuchenyia pry reshenyy zadach ynformatsyonnoi bezopasnosty. Voronezhskyi hosudarstvennыi unyversytet ynzhenernыkh tekhnolohyi, 4, 69–79.
5. Le Roux, N., Bengio, Y. (2008). Representational power of restricted Boltzmann machines and deep belief networks. Neural computation, 20(6), 1631-1649.
6. Sharma, B., Mangrulkar, R. (2019). Deep learning applications in cyber security: a comprehensive review, challenges and prospects. International Journal of Engineering Applied Sciences and Technology, 4(8), 148-159
7. Ranzato, M.A., Boureau, Y.L., Cun, Y.L. (2008). Sparse feature learning for deep belief networks. Advances in neural information processing systems, (pp. 1185-1192).

8. Mirkin, B. G. (2011). Core concepts in data analysis: Summarization, correlation and visualization. Springer Science & Business Media.
9. https://www.theverge.com/2018/1/12/16882408/google-racist-gorillas-photo-recognition-algorithm-ai
10. https://www.kaspersky.com/enterprise-security/wiki-section/products/sandbox
11. Kazennov, A. M. (2010). Basic concepts of CUDA technology. Computer Research and Modeling, 2(3), 295–308. https://doi.org/10.20537/2076-7633-2010-2-3-295-308
12. Loskutov, A. (Red.). (2003). Neirosetevыe alhorytmы prohnozyrovanyia y optymyzatsyy system. Nauka y Tekhnyka.