КІБЕРБЕЗПЕКА: освіта, наука, техніка
CYBERSECURITY: EDUCATION, SCIENCE, TECHNIQUE
№ 2 (18), 2022
ISSN 2663 - 4023
DOI 10.28925/2663-4023.2022.18.187196
УДК 004.8

Сукайло Ігор Олександрович
аспірант кафедри інформаційної та кібернетичної безпеки
імені професора Володимира Бурячка
Київський університет імені Бориса Грінченка, Київ, Україна
ORCID ID 0000-0003-1608-3149
i.ukailo.asp@kubg.edu.ua

Коршун Наталія Володимирівна
доктор технічних наук, професор,
професор кафедри інформаційної та кібернетичної безпеки
імені професора Володимира Бурячка
Київський університет імені Бориса Грінченка, Київ, Україна
ORCID ID: 0000-0003-2908-970X
n.korshun@kubg.edu.ua

ВПЛИВ NLU І ГЕНЕРАТИВНОГО ШІ НА РОЗВИТОК СИСТЕМ КІБЕРЗАХИСТУ

Анотація. Поєднання систем кібербезпеки та штучного інтелекту є логічним кроком на даному етапі розвитку інформаційних технологій. На сьогоднішній день багато постачальників засобів забезпечення кібербезпеки впроваджують машинне навчання та штучний інтелект у свої продукти або послуги. Разом з тим, ефективність інвестицій у роботи над передовими технологіями машинного навчання та глибокого навчання в контексті створення значних вимірних результатів цих продуктів є предметом дискусій. При розробці таких систем виникають проблеми з досягненням точності та масштабуванням. В статті розглянуто класифікацію систем штучного інтелекту, моделі штучного інтелекту, які використовуються продуктами безпеки, їх можливості, наведено рекомендації, які слід враховувати під час використання генеративних технологій штучного інтелекту щодо систем кіберзахисту. Можливості NLP ChatGPT можна використовувати для спрощення налаштування політик в продуктах безпеки. Доцільним є підхід, який враховує як короткострокові, так і довгострокові показники для вимірювання прогресу, диференціації та надання цінності споживачам за допомогою ШІ. Також розглянуто питання використання генеративного ШІ на базі платформних рішень, що дозволяє агрегувати різноманітні дані користувачів, обмінюватися ідеями та досвідом між великою спільнотою, а також опрацьовувати високоякісні телеметричні дані. Завдяки мережевому ефекту з'являється можливість донавчати моделі ШІ та покращувати результативність кіберзахисту для всіх користувачів. Ці переваги призводять до віртуального циклу підвищення залученості користувачів та покращення результатів кіберзахисту, що робить рішення безпеки на основі платформи привабливим вибором як для бізнесу, так і для приватних осіб. При проведені аудиту кіберзахисту будь-якої ІТ-інфраструктури засобами ШІ встановлюються обмеження та глибина аудиту з урахуванням попереднього досвіду.

Ключові слова: штучний інтелект; генеративна модель; кібербезпека; машинне навчання.

ВСТУП

Зважаючи на еволюцію систем штучного інтелекту та вражаюче зростання кіберзагроз, проглядаються тенденції, які декілька років тому здавалися фантастичними, але, нажаль, вони від цього не стають менш загрозливими. Наш світ стає все більше пов'язаним з технологіями та інтернетом, що робить роль кібербезпеки надзвичайно

важливою, якщо не ключовою. Причому кібербезпека стоїть не лише на захисті нашої приватності чи фінансової інформації, але й тісно пов'язана з фізичною безпекою.

Постановка проблеми. У 2004 році DARPA Grand Challenge [1] поклала початок фундаментальним дослідженням технологій безпілотних автомобілів з метою прискорення розробок у військовій сфері. 13 березня 2004 року жодна команда не пройшла маршрут DARPA Grand Challenge від Барстоу, штат Каліфорнія, до Прімма, штат Невада, але це спровокувало інтенсивні дослідження безпілотних автомобілів у наукових колах та автомобільних компаніях світу і, як наслідок, сьогодні на дорогах ми маємо «майже» безпілотні автомобілі. Вони керуються ШІ та потребують для захисту систем, що працюють на базі ШІ. Ми покладаємося на технології для вирішення як персональних задач, так і стратегічних проблем: від прохання до Siri про призначення зустрічі до лікування раку та боротьби зі зміною клімату. Однак, чим більше технології інтегруються з нашим життям, тим більша ймовірність того, що зловмисники спробують цим скористатися. Тому поєднання систем кібербезпеки та ШІ є невідворотним кроком.

Використання статистичних методів і методів машинного навчання для виявлення спаму почалося ще в середині 90-х років, але реалізація повністю автоматизованого центру безпеки на підприємствах залишається віддаленою можливістю, незважаючи на значний прогрес.

Аналіз останніх досліджень і публікацій.

В [2] систематизовано основні напрямки досліджень у галузі штучного інтелекту. В роботі [3] обґрунтовано можливість використання методів машинного навчання, глибокого навчання та інтелектуального аналізу даних у кібербезпеці задля виявлення вторгнень, аналізу шкідливих програм та виявлення спаму. Інновації штучного інтелекту в Gartner Hype Cycle™ for Artificial Intelligence 2022 [4] відображають пріоритети в чотирьох основних категоріях: ШІ, орієнтований на дані, модельно-орієнтований ШІ, ШІ, орієнтований на програми та людиноцентричний ШІ. В [5] серед переваг використання ШІ в кібербезпеці приводяться можливість його застосування для виявлення кіберзагроз та шкідливих дій, боротьби з ботами, прогнозування ризиків порушення кібербезпеки та для покращення захисту кінцевих пристроїв. Окремої уваги заслуговує інтеграція технологій машинного навчання в системи кібербезпеки [6]. В роботах [7], [8] розглянуті структури систем автоматичного розпізнавання мовлення, гібридних і наскрізних, плюси та мінуси кожної з них, включаючи порівняння даних навчання та вимог до обчислювальних ресурсів, та розглянуто основні підходи до розпізнавання мовлення. Модель NLU ChatGPT відображає значний прогрес у обробці природної мови та її розумінні, що відкриває нові можливості для використання ШІ у різних галузях, включаючи безпеку.

З огляду на сучасні темпи розвитку та значні прориви у галузі штучного інтелекту за останній час, можна вважати, що настав час для концентрації зусиль, спрямованих на більш глибоку інтеграцію ШІ у галузі кібербезпеки.

Метою статті є висвітлення поточного стану кібербезпеки з використанням ML/AI, ключових факторів впливу генеративного штучного інтелекту у галузі безпеки та кроків для використання переваг ГШІ як ефективного напрямку розвитку кібербезпеки.

ТЕОРЕТИЧНІ ОСНОВИ ДОСЛІДЖЕННЯ

Є багато означень штучного інтелекту, але якщо їх узагальнити, то можна сказати, що ШІ - це галузь комп'ютерних наук, яка досліджує і розробляє обчислювальні підходи і методи, що дозволяють машинам виконувати завдання, які зазвичай вимагають певного рівня людського інтелекту. Іншими словами, зробити машини розумними. За різними

джерелами можна зробити декілька класифікацій систем ШІ. Для прикладу розглянемо розподіл на чотири типи, виходячи з можливостей обчислювальної техніки порівняно з людським інтелектом [9], [10], [11].

Artificial Narrow Intelligence (ANI)

Також відомий як "слабкий ШІ". Такі системи характеризуються підходами, зосередженими на вирішенні дуже специфічних завдань у межах сфери, для якої вони були розроблені. ANI дуже добре виконує завданнями, що повторюються, і в багатьох випадках працює набагато краще, ніж людина. Прикладами є Siri, Google Translate і Watson від IBM.

Artificial Broad Intelligence (ABI)

Також відомий як "широкий ШІ". За своєю природою є інтеграцією двох або слабших систем, або методів ШІ, які приймають рішення для виконання завдання або процесу. Підприємства можуть використовувати дані, навчаючи системи для вирішення задач конкретних бізнес-процесів, наприклад, самокерованих транспортних засобів, аналізу інвестиційних стратегій для корпоративних клієнтів у банківській сфері.

Artificial General Intelligence (AGI)

Також відомий як "глибокий ШІ". У цих системах використовуються підходи, які дозволяють машинам виконувати інтелектуальні завдання на тому ж рівні, що і людина. Очікується, що системи AGI володітимуть теорією розуму, а також будуть самосвідомими, здатними розуміти переконання, думки, емоції та очікування людей і здатними до соціальної взаємодії. Як і люди, системи AGI можуть міркувати, розробляти стратегії та плани, спираючись на емоції та попередні знання. Однією з практичних реалізацій, що максимально наблизилася до систем цього класу, є ChatGPT.

Artificial Super-Intelligence (ASI)

Вважається, що системи ASI будуть використовувати підходи, які, гіпотетично, будуть володіти здібностями та інтелектом, що перевершують людські.

Окремо слід звернути увагу на одну з підсистем AGI, а саме на генеративний ШІ (GenAI) - це напрямок розвитку систем штучного інтелекту, які можуть генерувати усі види даних, включаючи аудіо, код, зображення, текст, 3D-об'єкти, відео тощо. Вона використовує для навчання існуючі дані, але в результаті генерує нові та унікальні результати, що призводить до виявлення нових підходів у багатьох галузях науки та у різних технологіях. GenAI отримав поштовх у розвитку значною мірою завдяки нещодавнім проривам у галузі AGI, таким як ChatGPT та Midjourney.

Використання МН/ШІ в існуючих системах кіберзахисту

Використання статистичних методів і машинного навчання в комерційних системах кіберзахисту бере свій початок ще з 90-х років і сьогодні майже всі постачальники засобів безпеки впроваджують МН та ШІ у свої продукти або послуги, оскільки це стало поширеною тенденцією в індустрії. Багато компаній, незалежно від розміру, мають спеціальні групи з обробки даних, які працюють над передовими технологіями машинного навчання та глибокого навчання. Однак ефективність цих інвестицій у створенні значних вимірних результатів є предметом дискусій.

МН/ШІ вже використовується в усіх сферах безпеки:

- виявлення шкідливих програм, спаму та фішингових атак;
- визначення та пріоритизація вразливостей для розробки та впровадження автоматизованих систем, що реагують на виявлені вразливості;
- реагування на інциденти для аналізу даних мережевого трафіку та системних журналів, виявлення потенційних інцидентів безпеки, розробка та впровадження систем автоматизованого реагування на виявлені інциденти;
- категоризація та фільтрація доступу до інтернету та електронної пошти;

• виявлення та запобігання вторгнень, блокування несанкціонованого доступу або підозрілої активності;

• аналіз мережевого трафіку;

• аналіз даних, як-то журнали активності користувачів та автентифікації, для виявлення шаблонів підозрілої активності та покращення контролю доступу. Найпоширенішими сферами використання є управління безпекою хмарних сервісів SaaS і рішення CNAPP.

• формування каталогу загроз для конкретної IT інфраструктури;

• виявлення внутрішніх загроз і поведінкова аналітика для аналізу даних з різних джерел, таких як активність користувачів та системні журнали, щоб виявити патерни підозрілої активності та потенційні загрози безпеці.

Основні підходи, прийняті для вирішення питань безпеки, включають використання статистичних методів, машинного навчання та алгоритмів глибокого навчання. Ці підходи ґрунтуються на принципах навчання без вчителя або з учителем, коли набори даних групуються разом на основі подібності та представлені в моделі, або де розмічені дані використовуються для навчання моделей безпеки, які згодом можуть бути використані для класифікації невідомих подій або об'єктів. Розмічені дані означають, що дані, які використовуються для навчання моделі штучного інтелекту, повинні бути позначені як «доброякісні» або «зловмисні» чи будь-якою іншою категорією, котру потрібно виявити, коли модель визначає щось як «хороше» або «погане».

Проблеми, що лежать в основі машинного навчання та систем глибокого навчання в усіх цих сферах, полягають у двох аспектах:

• По-перше, проблеми з досягненням точності. Це пов'язано з відсутністю великих обсягів розмічених даних у галузі кібербезпеки, які потрібні для навчання та перенавчання моделей. Крім того, багатовимірний та нестаціонарний характер даних і динамічна зміна векторів атак ускладнюють розробку надійних і точних моделей.

• По-друге, проблеми з горизонтальним та вертикальним масштабуванням цих систем глибокого навчання у реальних системах кіберзахисту. Це зазвичай збільшує витрати та ускладнює впровадження.

РЕЗУЛЬТАТИ ДОСЛІДЖЕННЯ

Основний вплив GenAI/ChatGPT на системи кіберзахисту

Отже, що саме змінилося у кіберзахисті з появою ChatGPT та моделей GenAI на основі Transformer? Щоб перейти до цього моменту, розглянемо дві моделі штучного інтелекту, які використовувалися продуктами безпеки.

Дискримінаційні моделі ШІ: ці моделі навчаються розрізняти або класифікувати дані за різними категоріями. Наприклад, деяку програму можна класифікувати як доброякісну або шкідливу. Ці моделі знайшли широке використання у сфері безпеки. Їм потрібні значні обсяги розмічених навчальних даних для початкового навчання нової моделі штучного інтелекту, яку потім можна використовувати для виявлення шкідливих програм або мережевих загроз.

Генеративні моделі штучного інтелекту: ці моделі вивчають ймовірнісний розподіл даних і можуть генерувати нові вибірки даних, подібні до вхідних. Їх можна використовувати, щоб визначити, чи призведе певна послідовність команд або подій в інфраструктурі до хорошого чи поганого результату (тобто до порушення).

GPT-3 (Generative Pre-trained Transformer 3) являє собою генеративну модель штучного інтелекту, яка в основному використовується для завдань NLU, таких як

генерація тексту, узагальнення, переклад, і не потребує розмічених даних на етапі попереднього навчання, але потребує їх для точного донавчання під конкретні завдання. Точне донавчання GPT-3 на меншому наборі розмічених даних може допомогти моделі покращити продуктивність для конкретних завдань, наприклад, класифікація тексту, розпізнавання іменованих об’єктів тощо. У контексті безпеки ці моделі мають три важливі можливості:

1. Ці моделі можуть сприймати послідовність подій або об'єктів, а потім з високою достовірністю передбачати, якою буде наступна подія або об'єкт.
2. Маючи великий обсяг даних логів і протоколів подій, ці моделі можуть вивчати складні загальні зв'язки і закономірності між подіями, що містяться в терабайтах логів, щоб передбачити, що станеться при заданій послідовності або наборі подій.
3. Маючи набір текстових правил, ці моделі можуть перекладати ці правила в інший формат або у мову програмування з високим ступенем точності.

«Можливість 1» може бути використана для прогнозування загроз: якщо зловмисник ввів послідовність команд у цільову систему - зловмисних чи безпечних, - модель передбачить наступні можливі дії зловмисника, а «Можливість 2» допоможе визначити, чи є цей результат атакою/порушенням безпеки. Також ці «можливості» можуть бути використані, щоб генерувати нові дані про загрози та диференціювати нові типи загроз для подальшого відпрацювання реакції на них.

«Можливість 3» може мати значний вплив на створення моделей політик безпеки, а саме - на концепцію політики безпеки і доступу на основі намірів. Хоча концепція намірів вже давно існує, ChatGPT дозволяє реалізувати її в простий спосіб – коли користувачі отримують зрозумілий інтерактивний інтерфейс для створення багаторівневих брандмауерів і веб-політик, просто виклавши у текстовій формі свої наміри. Можна вважати, що в кінцевому підсумку замкнута система оцінки ризиків в інфраструктурі, виявлення загроз, реагування на загрози і управління політиками відкриє можливість створення автономної хмари безпеки для підприємства.

Рекомендації щодо систем кіберзахисту, які слід враховувати під час використання генеративних технологій ШІ/ChatGPT

Ведення журналів і телеметрії в продуктах і сервісах безпеки

Хоча в багатьох продуктах кіберзахисту і немає недоліку у телеметрії, проблема полягає в отриманні значущих сигналів, які можна застосувати для навчання моделей безпеки. Йдеться не про великі дані, а про якісні дані.

Створення надійних механізмів прийняття рішень на основі ШІ

Команди SOC (Security Operations Center) інвестують у дорогі «технології МН», але часто відмовляються від них через відсутність довіри до рішень, що приймає система. Оскільки моделі штучного інтелекту стають більш ефективними в прийнятті рішень, їхня складність також зростає, що ускладнює розуміння того, як вони дійшли конкретних висновків. Наприклад, у сфері безпеки механізми прийняття рішень на основі штучного інтелекту, які пропонують обмежити доступ до системи через підвищений ризик або пом’якшити контроль доступу через низьку оцінку ризику того чи іншого середовища, повинні мати методи для уточнення своїх оцінок. Щоб підвищити довіру до рішень і рекомендацій моделей штучного інтелекту, інвестиції в доступні та зрозумілі інтерфейси для цих систем та розвиток людей, що з ними працюють, є критично важливими.

Безпека на основі намірів

Можливості NLP ChatGPT можна використовувати для спрощення налаштування політик в продуктах безпеки. Хоча постачальники послуг безпеки завжди прагнуть спростити конфігурацію політик, моделі на основі ChatGPT і NLP можуть допомогти досягти повноцінного єднання людської мови та мови конфігурації пристроїв при

створенні політик на основі намірів. Вони можуть допомогти у перекладі намірів користувачів, виражених природною українською мовою, у мову конфігурації систем кіберзахисту.

Наступним кроком буде включення в сервіси кіберзахисту механізму рекомендацій щодо політик. Цей механізм повинен аналізувати конфігурацію існуючої інфраструктури, щоб надати рекомендації щодо політик і правил іншим клієнтам. Наприклад, "Багато великих банків в Україні мають такі конфігурації політик… Чи хотіли б ви використати їх як базові?"

Робота зі зворотними відгуками від користувачів

Підходи до кіберзахисту на основі штучного інтелекту відрізняються: одні є практичними, а інші заглиблюються в фундаментальні концепції машинного навчання. Однак при використанні останніх важко встигати за сучасними загрозами та технологіями, що стрімко розвиваються. Тому часто лунають звинувачення у браку даних або у недосконалих алгоритмах. Тим часом інженери з безпеки, яким бракує наукового розуміння штучного інтелекту, можуть використовувати швидкі рішення МН, але вони швидко застарівають в міру динамічного розвитку загроз. Доцільним видається гібридний підхід, який враховує як короткострокові, так і довгострокові показники для вимірювання прогресу, диференціації та надання цінності споживачам за допомогою ШІ.

Підтримка ChatGPT/NLP для створення інтерфейсу користувача у продуктах

ChatGPT і NLP слід використовувати для створення більш просунутих та інтуїтивно зрозумілих користувацьких інтерфейсів для продуктів безпеки. Використовуючи можливості моделі генерувати природну мову, ці інтерфейси можна зробити більш інтуїтивно зрозумілими та легшими для навігації, що підвищить шанси на успішне впровадження та використання засобів кіберзахисту кінцевими користувачами. Крім того, слід переглянути процес взаємодії з клієнтами, щоб покращити час відгуку служби безпеки та якість реагування на ескалацію конфліктних ситуацій.

Очікується, що зростаюча популярність розмовної моделі ChatGPT в пошукових системах Google продовжиться, оскільки вона зменшує когнітивне навантаження на користувачів під час пошуку та вибору між результатами. Крім того, інтелектуальний і розмовний формат ChatGPT має тенденцію до спрощення взаємодії користувача з системою, який раніше вимірювався на основі кількості та послідовності кліків на інтерфейсі користувача.

Використання генеративного ШІ на базі платформних рішень

Під платформним підходом мається на увазі наявність "мережевого ефекту". Мережевий ефект - це явище, яке полягає в тому, що цінність системи безпеки, продукту або послуги зростає, коли все більше людей або організацій приймають і використовують їх. Це допомагає збільшенню цінності безпеки, пропонованої користувачам, в міру зростання кількості користувачів. Пропозиція безпеки стає більш цінною, коли більше користувачів беруть участь у мережі. Наприклад, платформа безпеки, якою користується багато користувачів, є більш цінною, ніж та, якою користується лише кілька користувачів. Це пояснюється тим, що велика мережа користувачів може обмінюватися інцидентами, подіями, конфігураціями та інформацією про ризики, що дозволяє донавчати моделі ШІ та покращувати результативність кіберзахисту для всіх користувачів. Чим більше користувачів, тим ціннішою стає платформа і тим привабливішою вона стає для потенційних користувачів. Це створює замкнений цикл підвищення залученості користувачів і покращення можливостей системи кіберзахисту. Постачальники послуг безпеки з великою кількістю користувачів, які приділяють особливу увагу високоякісній телеметрії для навчання моделей і є прозорими щодо алгоритмів ШІ, що використовуються в їхніх рішеннях, з більшою

ймовірністю допоможуть іншим досягти успіху. Чим ширший простір використання таких систем, тим вища якість рішень, що приймаються генеративними алгоритмами ШІ, і тим кращий результат.

Загалом, використовуючи рішення безпеки на основі платформи, постачальники послуг кіберзахисту мають унікальну можливість створювати кращі рішення на основі штучного інтелекту. Цей підхід дозволяє агрегувати різноманітні дані користувачів, обмінюватися ідеями та досвідом між великою спільнотою, а також опрацьовувати високоякісні телеметричні дані для навчання моделей. Ці переваги призводять до віртуального циклу підвищення залученості користувачів, покращення результатів кіберзахисту, що робить рішення безпеки на основі платформи привабливим вибором як для бізнесу, так і для приватних осіб. Слід зауважити, що вищевказані платформи можуть бути як глобальними, так і гібридними чи приватними у рамках однієї держави чи певної спільноти в індустрії, що призведе до використання для донавчання даних, притаманних саме цій спільноті.

ВИСНОВКИ ТА ПЕРСПЕКТИВИ ПОДАЛЬШИХ ДОСЛІДЖЕНЬ

Будь-яка система кіберзахисту на основі систем ШІ тим ефективніша, чим більша кількість користувачів її використовує. Будь-яка замкнена, відокремлена від інтернету, захищена ІТ-інфраструктура ризикує опинитися у становищі, коли зловмисні програми у інтернеті розвиваються значно швидше і ефективніше під час боротьби систем ШІ, що постійно удосконалюють засоби атак та засоби захисту від них. І саме ці зловмисні програми тим чи іншим шляхом, найчастіше завдяки людському фактору, можуть потрапити у відокремлену інфраструктуру і вивести її з ладу.

При проведені аудиту кіберзахисту будь-якої ІТ-інфраструктури засобами ШІ встановлюються обмеження та глибина аудиту, враховуючи попередній досвід, і досить обмежено прогнозування результату аудиту. При цьому завжди існує ймовірність, що ШІ настільки заглибиться у процес аудиту, що при пошуку нових вразливостей знайде можливість виконати і виконає команду на кшталт «format c:», чим повністю виведе з ладу систему, що перевіряється.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1 Defense Advanced Research Projects Agency - The Grand Challenge https://www.darpa.mil/about-us/timeline/-grand-challenge-for-autonomous-vehicles
2 Савченко В.А., Шаповаленко О.Д. (2020) Основні напрями застосування технологій штучного інтелекту у кібербезпеці. Сучасний захист інформації, 4(44), 6-11.
3 Azzah Kabbas , Atheer Alharthi, Asmaa Munshi (2020) Artificial Intelligence Applications in Cybersecurity. International Journal of Computer Science and Network Security, V.20 №.2, 120-124.
4 J. Wiles (2022) What's New in Artificial Intelligence from the 2022 Gartner Hype Cycle https://www.gartner.com/en/articles/what-s-new-in-artificial-intelligence-from-the-2022-gartner-hype-cycle
5 Gaurav Belani. The Use of Artificial Intelligence in Cybersecurity: A Review https://www.computer.org/publications/tech-news/trends/the-use-of-artificial-intelligence-in-cybersecurity
6 Ivanichenko, Y., Sablina, M., & Kravchuk, K. (2021). Використання машинного навчання в кібербезпеці. Електронне фахове наукове видання «Кібербезпека: освіта, наука, техніка», 4(12), 132–142. https://doi.org/10.28925/2663-4023.2021.12.132142
7 O. Iosifova, I. Iosifov, V. Sokolov, O. Romanovsky and I. Sukaylo (2021). Analysis of Automatic Speech Recognition Methods. CEUR Workshop Proceedings, 2923, 252 – 257.

8 I. Iosifov, O. Iosifova, V. Sokolov, P. Skladannyi and I. Sukaylo (2021). Natural Language Technology to Ensure the Safety of Speech Information. CEUR Workshop Proceedings, 3187, 216 – 226.
9 Goertzel, B., & Pennachin, C. (2007). Artificial general intelligence (Ser. Cognitive Technologies). Springer. https://doi.org/10.1007/978-3-540-68677-4
10 IBM Services. (2018). Beyond the hype: A guide to understanding and successfully implementing artificial intelligence within your business. https://www.ibm.com/downloads/cas/8ZDXNKQ4
11 O'Carroll, B. (2020, January 31). What are the 3 types of AI? A guide to narrow, general, and super artificial intelligence. https://codebots.com/artificial-intelligence/the-3-types-of-ai-is-the-third-even-possible

КІБЕРБЕЗПЕКА: освіта, наука, техніка
CYBERSECURITY: EDUCATION, SCIENCE, TECHNIQUE
№ 2 (18), 2022
ISSN 2663 - 4023

Igor O. Sukaylo
Ph.D. student of the Department of Information and Cybersecurity
named after Professor Volodymyr Buriachok
Borys Grinchenko Kyiv University, Kyiv, Ukraine
ORCID ID 0000-0003-1608-3149
i.ukailo.asp@kubg.edu.ua

Nataliia V. Korshun
Doctor of Technical Sciences, professor,
Professor of the Department of Information and Cybersecurity
named after Professor Volodymyr Buriachok
Borys Grinchenko Kyiv University, Kyiv, Ukraine
ORCID ID: 0000-0003-2908-970X
n.korshun@kubg.edu.ua

THE INFLUENCE OF NLU AND GENERATIVE AI ON THE DEVELOPMENT OF CYBER DEFENSE SYSTEMS

Abstract. The combination of cyber security systems and artificial intelligence is a logical step at this stage of information technology development. Today, many cybersecurity vendors are incorporating machine learning and artificial intelligence into their products or services. However, the effectiveness of investments in advanced machine learning and deep learning technologies in terms of generating meaningful measurable results from these products is a matter of debate. When designing such systems, there are problems with achieving accuracy and scaling. The article considers the classification of artificial intelligence systems, artificial intelligence models used by security products, their capabilities, recommendations that should be taken into account when using generative artificial intelligence technologies for cyber protection systems are given. ChatGPT's NLP capabilities can be used to simplify the configuration of policies in security products. An approach that considers both short-term and long-term metrics to measure progress, differentiation, and customer value through AI is appropriate. The issue of using generative AI based on platform solutions, which allows aggregating various user data, exchanging ideas and experience among a large community, and processing high-quality telemetry data, is also considered. Thanks to the network effect, there is an opportunity to retrain AI models and improve the effectiveness of cyber defense for all users. These benefits lead to a virtual cycle of increased user engagement and improved cyber security outcomes, making platform-based security solutions an attractive choice for businesses and individuals alike. When conducting a cyber security audit of any IT infrastructure using AI, the limits and depth of the audit are established taking into account previous experience.

Keywords: Artificial Intelligence; generative model; cyber security; machine learning.

REFERENCES

1 Defense Advanced Research Projects Agency - The Grand Challenge https://www.darpa.mil/about-us/timeline/-grand-challenge-for-autonomous-vehicles
2 Savchenko V.A., Shapovalenko O.D. (2020) The main areas of application of artificial intelligence technologies in cybersecurity. Modern information protection, 4(44), 6-11.
3 Azzah Kabbas , Atheer Alharthi, Asmaa Munshi (2020) Artificial Intelligence Applications in Cybersecurity. International Journal of Computer Science and Network Security, V.20 №.2, 120-124.
4 J. Wiles (2022)What's New in Artificial Intelligence from the 2022 Gartner Hype Cycle https://www.gartner.com/en/articles/what-s-new-in-artificial-intelligence-from-the-2022-gartner-hype-cycle
5 Gaurav Belani. The Use of Artificial Intelligence in Cybersecurity: A Review https://www.computer.org/publications/tech-news/trends/the-use-of-artificial-intelligence-in-cybersecurity

6 Ivanichenko, Y., Sablina, M., & Kravchuk, K. (2021). Using machine learning in cybersecurity. Cybersecurity: Education, Science, Technique, 4(12), 132–142. https://doi.org/10.28925/2663-4023.2021.12.132142

7 O. Iosifova, I. Iosifov, V. Sokolov, O. Romanovsky and I. Sukaylo (2021). Analysis of Automatic Speech Recognition Methods. CEUR Workshop Proceedings, 2923, 252 – 257.

8 I. Iosifov, O. Iosifova, V. Sokolov, P. Skladannyi and I. Sukaylo (2021). Natural Language Technology to Ensure the Safety of Speech Information. CEUR Workshop Proceedings, 3187, 216 – 226.

9 Goertzel, B., & Pennachin, C. (2007). Artificial general intelligence (Ser. Cognitive Technologies). Springer. https://doi.org/10.1007/978-3-540-68677-4

10 IBM Services. (2018). Beyond the hype: A guide to understanding and successfully implementing artificial intelligence within your business. https://www.ibm.com/downloads/cas/8ZDXNKQ4

11 O'Carroll, B. (2020). What are the 3 types of AI? A guide to narrow, general, and super artificial intelligence. https://codebots.com/artificial-intelligence/the-3-types-of-ai-is-the-third-even-possible