УДК 004.056
DOI: 10.31673/2409-7292.2021.032428

Найман Г.Г., Гаркавенко Д.М.

МЕТОДИ ТА ЗАСОБИ УПРАВЛІННЯ ВРАЗЛИВОСТЯМИ КОРПОРАТИВНОЇ ІНФОРМАЦІЙНОЇ СИСТЕМИ НА ОСНОВІ МАШИННОГО НАВЧАННЯ

В статті висвітлено загальні поняття про IT-інфраструктуру, про корпоративну інформаційну систему, кібератаки і статистику кібератак. Так само розповідається про управління вразливостями, як процес зниження шкоди від реалізації загроз, з якими проблемами стикаються фахівці інформаційної безпеки і шляхи їх вирішення. Розглянуто можливість застосування штучного інтелекту в області інформаційної безпеки, переваги, недоліки, проблеми. Запропоновано гіпотетичне рішення шляхом розробки методів і засобів управління вразливостями в корпоративних інформаційних системах з використанням технологій машинного навчання.

Ключові слова: IT-інфраструктура, корпоративна інформаційна система, кібератака, вразливості, управління вразливостями, машинне навчання.

Вступ

В контексті кібербезпеки, минулий рік був з великою кількістю APT-атак, пошуку апаратних вразливостей і гучних витоків. За той час, поки керівники компаній приходили до усвідомлення необхідності вибудовувати дійсно ефективну систему інформаційної безпеки, злочинці міцно влаштувалися в кіберпросторі. Найбільш яскравим прикладом став ринок в дарквеб, де продається маса заборонених товарів і послуг, в тому числі хакерські утиліти і доступ до вже зламаної інфраструктури. Крім того, злочинці продовжують використовувати неграмотність користувачів в питаннях забезпечення власної безпеки. Співвідношення сил кіберзлочинців і захисників виявляється не на користь останніх: APT-угруповання активно використовують новітні уразливості, діють дуже швидко, а головне - часто змінюють інструментарій та тактики. Безпосередня загроза складних цілеспрямованих атак спонукає компанії по-новому поглянути на ефективність систем захисту. Настав час переглянути старі підходи і поговорити про новий тип інформаційної безпеки.

Проблеми машинного навчання в кібербезпеці.

Незважаючи на оптимістичні надії на машинне навчання, і прогнози, що вже скоро воно замінить всіх живих фахівців з інформаційної безпеки своїми автоматичними алгоритмами, в реальності говорити про це ще рано. Повної відмови від колишніх методів кібербезпеки на користь машинного навчання перешкоджають декілька причин.

По перше, нейромережеві моделі поводять себе не досить прозоро, яка не пояснює, чому з таких вхідних даних вийшов саме цей результат. Така відсутність безпосереднього зворотного зв'язку в когнітивному плані не дозволяє повністю відмовитися від людського контролю в таких важливих сферах, як інформаційна безпека.

По друге, відсутність достатнього для коректного машинного навчання моделей, кількості даних в усіх напрямках кіберзагроз, від комп'ютерних вірусів до прийомів соціальної інженерії.

Не можна ігнорувати і можливість специфічних атак на алгоритми машинного навчання, що може привести до невірних рішень, пропущеним атакам або помилкових спрацьовувань.

В свою чергу зловмисники теж використовують алгоритми штучного інтелекту для створення шкідливих програм, аналізу користувача поведінки, пошуку вразливостей, підбору паролів, підміни особистості, обходу систем захисту.

Також варто відзначити деякий конфлікт між вимогами генерального регламенту про захист персональних даних громадян і резидентів Євросоюзу General Data Protection Regulation та використанні цієї інформації в моделях машинного навчання кібербезпеки. Зокрема, GDPR передбачає наявність у користувача можливості «бути забутим», якщо він не дає згоди на збір своїх персональних даних або вирішив його відкликати. Ця вимога може

© Найман Г.Г., Гаркавенко Д.М., 2021

порушуватися, якщо імовірна модель автоматично аналізує поведінку користувача для попередження загроз [9].

Поки шо, машинне навчання не може замінити колишні методи кібербезпеки, проте вже значно доповнює і розширює їх. Зокрема, такі моделі можуть підвищувати точність сигнатурного аналізу, який швидко обробляє запити і не вимагає тривалого періоду навчання. Можна використовувати сигнатурний аналіз для виявлення запитів з явними ознаками атаки, а машинне навчання - для аналізу інших запитів. В результаті такого поєднання різних методів досягається висока швидкість роботи антивірусного ПЗ з мінімальною кількістю помилкових спрацьовувань і пропусків атак [10].

Вибір методу і постановка задачі. Виходячи з видів технологій машинного навчання і вирішуваних нею завдань, однією з найбільш відповідних для виявлення вразливостей є завдання класифікації за допомогою навчання з учителем - шляхом відповідної реалізації програмно-апаратної системи. При цьому, вхідними даними може бути опис ПЗ в одному з його уявлень (якоїсь стадії стану ПЗ, на яких буде проходити навчання), навчальною вибіркою буде безліч відомих і штучно згенерованих вразливостей в цьому уявленні, ознаками - особливості ПЗ, а результатом - клас вразливості або просто факт її наявності та методи вирішення. В свою чергу, управління уразливими за допомогою машинного навчання допоможе зменшити час на реагування та виправлення потенційних недообліку в архітектури інформаційних систем [11].

IT інфраструктура. У XXI столітті технології лежать в основі практично всіх аспектів сучасної людини, особливо сучасного підприємства - від організації роботи співробітників до операційної діяльності, виробництва товарів і надання послуг. Правильно налаштована мережева взаємодія дозволяє оптимізувати технології з метою поліпшення обміну інформацією, підвищення ефективності і продуктивності. Гнучка, надійна, а головне безпечна IT-інфраструктура допомагає підприємству досягти поставлених цілей і отримати конкурентну перевагу.

IT-інфраструктура включає в себе взаємопов'язані елементи і складається з двох базових груп компонентів - апаратного і програмного забезпечення.

До апаратних компонентів відносяться:

- настільні комп'ютери;
- сервери;
- центри обробки даних;
- концентратори;
- маршрутизатори;
- комутатори;
- об'єкти фізичної інфраструктури.

В свою чергу до програмних компонентів відносяться:

- системи управління контентом (CMS);
- системи управління взаємозв'язками з клієнтами (CRM);
- системи планування ресурсів підприємства (ERP);
- операційні системи;
- веб-сервери [1].

Корпоративна інформаційна система і її кібербезпека. Корпоративна інформаційна система - є складовою частиною IT-інфраструктури, що включає в себе бази даних, інформаційні центри, системи зв'язку, спільного доступу і роботи.

Корпоративні інформаційні системи, в першу чергу і піддаються кібератакам зловмисників. У світлі повідомлень про атаки, спрямованих на експлуатацію вразливостей IT-інфраструктури компаній, хочу привести статистику виявлених кіберзагроз за 2019-2020 роки (Рис. 1).

© Найман Г.Г., Гаркавенко Д.М., 2021

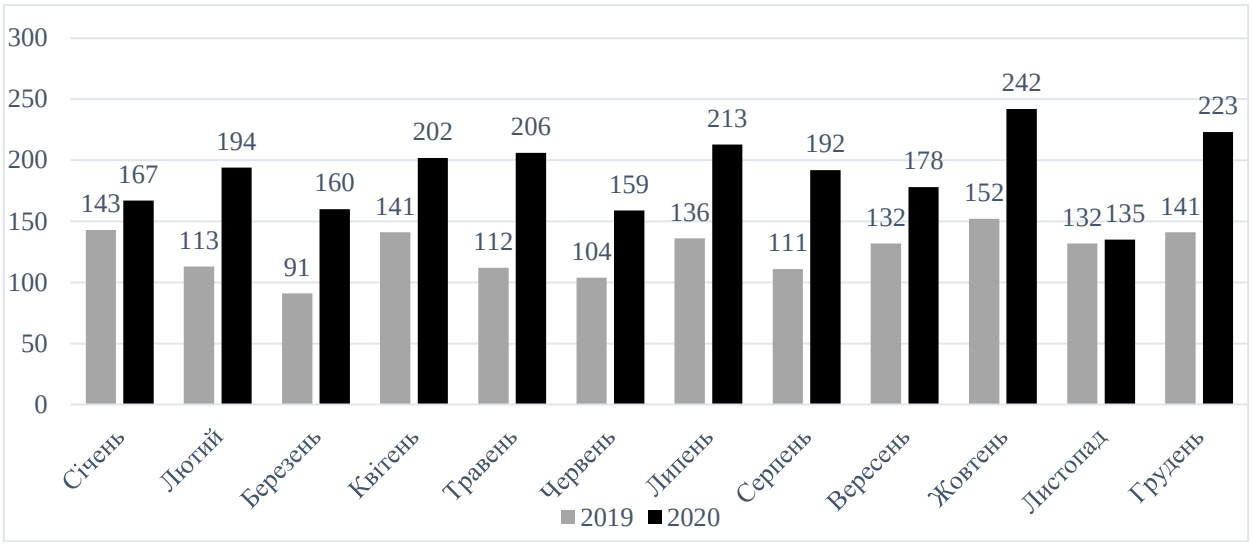

Рис.1. Кількість інцидентів в 2019 та 2020 роках

Свою роль зіграла і пандемія COVID-19, в зв'язку з нею був ініційований активний перехід співробітників на віддалену роботу, вивід внутрішніх сервісів компаній на мережевий периметр. Це вплинуло на ландшафт кіберзагроз у всьому світі: зловмисники без зволікання приступили до пошуку вразливостей в сервісах на периметрі компаній, в тому числі в рішеннях, які використовуються для організації віддаленої роботи, перевіряючи їх на міцність.

Наприклад, активно експлуатувалися проломи в Pulse Secure VPN, Citrix ADC і Citrix Gateway, в між мережевому екрані Cisco ASA. Оператори програм-вимагачів, зокрема Netwalker, Clop і REvil, користувалися вразливостями сервісів для поширення свого ПЗ [2].

Найбільш часто експлуатовані вразливості в 2020 році:

- CVE-2019-19781 (Citrix ADC і Citrix Gateway);
- CVE-2017-11882 (Microsoft Office);
- CVE-2019-11510 (Pulse Secure VPN);
- CVE-2020-11651 і CVE-2020-11652 (SaltStack Salt);
- CVE-2020-14882 (Oracle WebLogic);
- CVE-2019-0708 (RDP).

Протягом року спостерігалось збільшення числа атак, спрямованих на крадіжку корпоративних облікових даних співробітників. Для досягнення цієї мети хакери зламували веб-ресурси і викрадали бази з обліковими даними, підробляли форми аутентифікації, поширювали шпигунське ПЗ в корпоративній мережі і підбирали паролі для підключення до служб на мережевому периметрі. Як приклад, зловмисники відправляли жертвам посилання на офіційні документи, розміщені на підробленому ресурсі, що імітує інтерфейс хмарного сховища Google. Для перегляду документа було потрібно пройти процедуру аутентифікації, ввівши корпоративні облікові дані для Microsoft Office 365.

Інша тенденція - зростання ринків з продажу доступу до серверів компаній. На сьогоднішній день, навіть якщо хакери не змогли просунутися в атаці далі знайденої уразливості і отримання доступу до сервера, вони можуть легко продати цей доступ на форумі в дарквебі. Хочу зауважити, що постраждати може не тільки IT-інфраструктура компанії, але і її репутація в інтернеті - сайт. В 2020 році спостерігалось зміцнення інтересу до теми взлому сайтів. Пов'язати таку тенденцію можна з глобальним переходом організацій до роботи онлайн.

Вразливості і методи управління ними. Беручи до уваги все сказане, на мою думку, значної частини успішних кібератак можна було б уникнути, вдосконаливши методи

© Найман Г.Г., Гаркавенко Д.М., 2021

управління вразливостями. Саме по собі управління вразливостями - це процес, спрямований на зниження шкоди від реалізації загроз, обумовлених вразливістю інфраструктури.

Не всі вразливості несуть однаковий ризик. Ефективне управління вразливостями означає перехід від методу «виправляти все і завжди», до розумної розстановки пріоритетів щодо усунення загроз. Цей метод працює ефективно, поки компанія відносно невелика, досить мати один або кілька сканерів вразливостей і одного фахівця, який буде проводити періодичну перевірку всієї інфраструктури, закриваючи найбільш очевидні або прості в усуненні проблеми.

На практиці ж, коли компанія зростає, зростає число пристроїв в мережі, використовуються нові інформаційні системи, включаючи нестандартні, простого підходу вже не вистачає. Статистика, зібрана в 2020 році це підтверджує. Згідно ній, найбільше часу у фахівців інформаційної безпеки йде на те, щоб переконати IT-відділ в необхідності виправити вразливості (48%), і на аналіз результатів сканування (43%). Також до трудомістких завдань 31% фахівців віднесли перевірку усунення вразливостей. Зазначені труднощі характерні як для великих, так і для маленьких компаній. Однак простежується тенденція: чим більша компанія, тим важче фахівцям з ІБ домовитися з IT-відділом [3].

Саме по собі управління вразливостями - це досить складний процес виявлення вразливостей, оцінка ризику і вжиття відповідних заходів щодо їх усунення. Існує багато факторів, що впливають на прийняття рішень:

- високі ризики стати жертвою масової або таргетованої атаки, якщо вчасно не усунути вразливості (особливо це стосується зовнішнього периметру);
- висока вартість усунення для багатьох вразливостей, особливо, якщо немає готового рішення або вразливості піддається велика кількість розподілених пристроїв;
- коли за різні типи обладнання відповідають різні люди або навіть компанії, які не завжди володіють необхідною кваліфікацією для правильної оцінки виявлених вразливостей, процес усунення може вкрай затягнутися в часі або не розпочатися взагалі;
- якщо використовується спеціалізоване обладнання або SCADA-системи, то висока ймовірність відсутності необхідних оновлень від виробника або неможливість оновлення системи в принципі [4].

Через всі ці проблеми, процес управління вразливостями, реалізований безсистемно, не виглядає ефективним зовсім.

NIST Cybersecurity Framework. Унаслідок необхідності посилення кібербезпеки в 2013 році був підписаний указ №13636, який зобов'язав американський інститут по стандартизації NIST розробити базову модель захисту американських критичних інфраструктур (Cybersecurity Framework).

Cybersecurity Framework була опублікована в лютому 2014-го року, а останнє її оновлення відбулося в липні 2015. Закладена в ній ідея досить проста - уніфікувати підходи з безпеки промислових систем протягом усього їх життєвого циклу, спираючись на вже існуючі стандарти ІБ і передовий досвід.

В основу моделі було покладено модифікована тріада - доступність, цілісність і конфіденційність. Всі захисні заходи розбиті на п'ять великих блоків (функцій) - ідентифікація, захист, виявлення, реагування та відновлення. По суті, мова йде про життєвий цикл системи захисту будь-якого об'єкта - від корпоративної до промислової мережі [5].

З одного боку, для боротьби з конкретними вразливостями застосовуються відповідні алгоритми їх виявлення, при цьому, виходячи з суб'єктивності поняття «вразливість», в основу алгоритмів повинен бути закладений досвід експертів [6,7].

З іншого боку, вразливості досить швидко і просто вдосконалюються, що не дозволяє застосовувати застарілі правила.

Що б максимізувати ефективність управління вразливостями при мінімізації витрачених ресурсів, я пропоную розробити систему засобів і методів управління вразливостями в корпоративних інформаційних системах на основі машинного навчання. Вона зможе допомогти підвищити рівень захисту на кожному з п'яти рівнів.

© Найман Г.Г., Гаркавенко Д.М., 2021

Машинне навчання. Само по собі машинне навчання, це клас методів штучного інтелекту, характерною рисою яких є не пряме рішення задачі, а навчання за рахунок застосування рішень безлічі подібних завдань.

Виділяють три основні методи машинного навчання - контрольоване навчання, або навчання з учителем, неконтрольоване навчання або навчання без учителя, і навчання з підкріпленням.

Контрольоване навчання - цей метод навчання застосовується у випадках, коли є великі обсяги даних, база вразливостей з маркерами. Необхідно створити алгоритм, за допомогою якого машина могла б визначити вразливість. У ролі «вчителя» в даному випадку виступає людина, яка заздалегідь проставила маркери. Машина сама обирає ознаки, за якими вона розрізняє вразливості.

Завдання машини при неконтрольованому навчанні - знаходження зв'язку між окремими даними, виявлення закономірностей, підбирання шаблонів, упорядковування даних або опис їх структури, класифікація даних.

Навчання з підкріпленням є окремим випадком контрольованого навчання, але вчителем в даному випадку є «середовище». Машина або «агент» не має попередньої інформації про середовище, але має можливість здійснювати в ній будь-які дії. Середа реагує на ці дії і тим самим надає агенту дані, які дозволяють йому реагувати на них і вчитися. Фактично агент і середовище утворюють систему зі зворотним зв'язком [8].

Висновки

Виходячи з усього викладеного, а також існуючих напрацювань в області кібербезпеки, можливостей машинного навчання, застосування останньої для виявлення і управління вразливостями може виявитися вкрай затребуваним. Подальшим етапом повинна стати практична реалізація запропонованої системи, оцінка її результативності, вибір найкращих алгоритмів, ознак і властивостей машинного навчання, а також перевірка методів та засобів в реальному середовищі, прораховування ризиків її використання.

Перелік посилань

1. Что такое ИТ-инфраструктура года [Електронний ресурс] – Режим доступу: https://www.ibm.com/ru-ru/topics/infrastructure (29.04.2021)

2. Актуальные киберугрозы: итоги 2020 года [Електронний ресурс] – Режим доступу: https://www.ptsecurity.com/ru-ru/research/analytics/cybersecurity-threatscape-2020/ (01.05.2021)

3. Как выстроен процесс управления уязвимостями в российских компаниях [Електронний ресурс] – Режим доступу: https://www.ptsecurity.com/ru-ru/research/analytics/kak-vystroen-process-upravleniya-uyazvimostyami-v-rossijskih-kompaniyah/ (06.05.2021)

4. Как мы управление уязвимостями построили [Електронний ресурс] – Режим доступу: https://habr.com/ru/company/acribia/blog/460048/ (07.05.2021)

5. An Introduction to the Components of the Framework [Електронний ресурс] – Режим доступу до ресурсу: https://www.nist.gov/cyberframework/online-learning/components-framework (10.05.2021)

6. Буйневич М. В., Израилов К. Е., Мостович Д. И. Сравнительный анализ подходов к поиску уязвимостей в программном коде // Актуальные проблемы инфотелекоммуникаций в науке и образовании (АПИНО-2016): сборник научных статей V Международной научно-технической и научно-методической конференции. 2016. Т. 1. С. 256–260.

7. Израилов К. Е. Поиск уязвимостей в различных представлениях машинного кода // Информационная безопасность регионов России (ИБРР-2015): материалы IX Санкт-Петербургской межрегиональной конференции. 2015. С. 157.

8. Великое пробуждение исскуственного интелекта [Електронний ресурс] – Режим доступу: https://vc.ru/21767-the-great-ai-awakening (12.05.2021)

9. Artificial Intelligence vs. Machine Learning in Cybersecurity [Електронний ресурс] – Режим доступу: https://www.varonis.com/blog/ai-vs-ml-in-cybersecurity/ (13.05.2021)

10. Как машинное обучение помогает бороться с хакерскими атаками [Електронний ресурс] – Режим доступу: https://ib-bank.ru/bisjournal/post/822 (13.05.2021)

11. Буйневич М. В., Жуковская П. Е., Израилов К. Е., Покусов В. В. Применение машинного обучения для поиска уязвимостей в программном коде // Информационные технологии и телекоммуникации. 2019. Том 7. № 4. С. 59–65. DOI 10.31854/2307-1303-2019-7-4-59-65.

Надійшла: 04.07.2021 Рецензент: д.т.н., професор Гайдур Г.І.

© Найман Г.Г., Гаркавенко Д.М., 2021