Міністерство освіти і науки України
Прикарпатський національний університет імені Василя Стефаника
Кафедра комп'ютерної інженерії та електроніки
(повна назва кафедри)

Карпенко Роман Андрійович
Karpenko Roman

УДК 004:681.5

Спеціальність 123-Комп'ютерна інженерія
(шифр та назва спеціальності)

Кваліфікаційна робота
на здобуття освітньо-кваліфікаційного рівня магістр
(бакалавр, спеціаліст, магістр)

Розробка системи протидії фроду у фінансових організаціях. Автоматичне виявлення загроз у системі за допомогою комплексної системи машинного навчання
Development of a fraud prevention system in financial organizations. Automatic detection of threats in the system using a complex machine learning system

Науковий керівник:
кандидат фіз.-мат. наук, доцент
Запухляк Р.І.

Рецензент:
кандидат фіз.-мат. наук, доцент

директор Центру дистанційного навчання Івасюк І.Я.

Івано-Франківськ

2022

Формат	Поз.	Позначення	Найменування	Кільк.	Прим.
		123.КІ(м)-21.11	Пояснювальна записка	1	

Змн.	Арк.	№ докум.	Підпис	Дат	123.КІ(м).21.11		
Розробив		Карпенко Р.А.			Специфікація	Літ.	Арк. 2 / Аркушів 1
Перевірив		Запухляк Р.І.					
Н. Контр.							
Затвердив							

АНОТАЦІЯ

Карпенко Р.А. Розробка системи протидії фроду у фінансових організаціях. Автоматичне виявлення загроз у системі за допомогою комплексної системи машинного навчання : магістер. робота : спец. 123 «Комп'ютерна інженерія» / Карпенко Роман Андрійович ; Прикарпатський національний університет імені Василя Стефаника ; каф. комп'ютерної інженерії та електроніки ; наук. керівник Р.І. Запухляк Р.І. к.ф.-м.н., доц. – Івано-Франківськ, 2022. – 101 с.

Зміст роботи (анотація): визначено загальні принципи кібербезпеки; детально розглянуто проблематику кібербезпеки вебресурсів, різновиди атак на вебресурси та інструменти моніторингу стану безпеки вебресурсів; проаналізовано кібератаки, здійснених на вебресурси фінансових установ різних форм власності в Україні; розглянуто психологічні аспекти поведінки користувачів, що використовуються як для виявлення, так і для побудови поведінкової моделі бота; розкрито сутність категорії «система протидії фроду» в контексті в контексті заходів з кібербезпеки вебресурсів; систематизовано існуючі практики використання машинного навчання для перевірки веб-ресурсів на вразливості; здійснення робота зі створення облікових записів користувачів з використанням моделей машинного навчання; експериментальним шляхом доведено вразливість та недосконалість антифрод мереж; розроблено план подальшого дослідження системи протидії фроду з урахуванням отриманих результатів.

Ключові слова: системи та методи протидії фроду, рекурентна нейронна мережа, моделі машинного навчання, етапи ідентифікації користувача, рівні ідентифікаторів користувача, перевірка вебресурсів на вразливості, поведінкова модель бота, кібератака.

Змн.	*Арк.*	*№ докум.*	*Підпис*	*Дата*			
Розробив		Карпенко Р.А.			Анотація	*Літ.*	*Арк.* 3 / *Аркушів* 1
Перевірив		Запухляк Р.І.					
Н. Контр.							
Затвердив							

ABSTRACT

Karpenko R.A. Development of a fraud prevention system in financial organizations. Automatic detection of threats in the system using a complex system of machine learning: master's degree. work: spec. 123 "Computer engineering" / Karpenko Roman Andreyovych ; Prykarpattia National University named after Vasyl Stefanyk; department computer engineering and electronics; of science head R.I. Zapuhlyak R.I. Ph.D.-M.Sc., Assoc. – Ivano-Frankivsk, 2022. – 101 p.

The content of the work (abstract): the general principles of cyber security are defined; the issues of cyber security of web resources, types of attacks on web resources and tools for monitoring the state of security of web resources are considered in detail; analyzed cyber attacks carried out on the web resources of financial institutions of various forms of ownership in Ukraine; the psychological aspects of user behavior are considered, which are used both to identify and to build a behavioral model of the bot; the essence of the "fraud prevention system" category is revealed in the context of cyber security measures of web resources; existing practices of using machine learning to check web resources for vulnerabilities are systematized; implementation of work on creating user accounts using machine learning models; experimentally proved the vulnerability and imperfection of anti-fraud networks; a plan for further research of the anti-fraud system was developed, taking into account the results obtained.

Keywords: fraud prevention systems and methods, recurrent neural network, machine learning models, stages of user identification, levels of user identifiers, checking web resources for vulnerabilities, bot behavioral model, cyber attack.

Міністерство освіти і науки України
Прикарпатський національний університет імені Василя Стефаника
Кафедра комп’ютерної інженерії та електроніки

Пояснювальна записка

до магістерської кваліфікаційної роботи на тему:

«Розробка системи протидії фроду у фінансових організаціях. Автоматичне виявлення загроз у системі за допомогою комплексної системи машинного навчання»

ЗМІСТ

ВСТУП

.

РОЗДІЛ 1 КІБЕРБЕЗПЕКА ВЕБРЕСУРСІВ: АКТУАЛЬНІ ПРОБЛЕМИ ТА ВИКЛИКИ

Для того щоби розбіратися з проблематикою моєї магістерської роботи потрібно спершу відповісти на декілька запитань. А саме:

- Що таке кібербезпека?
- Чому кібербезпека важлива?
- Як кібербезпека пов’язана з ботами для автоматизації акаунтів?

Отже, що таке кібербезпека, - це сукупність усіх реєстраційних, превентивних та архітектурних заходів, які направлені на збереження та захищення інформаційних систем від неправомірного доступу. Тоді як інформаційна система - це сукупність засобів для збереження та обробки інформації.

Завдяки технологічному розвитку люди отримують все більше можливостей, на даний момент ми здатні робити таку кількість речей, які не були доступні людям ще 30 років тому. Люди здатні контактувати одне з одним за тисячі кілометрів, використовують роботів для своєї роботи, їздити в машинах з автопілотом та багато іншого. Однак з розвитком, приходить і небезпека. Через невпинну цифровізацію ефективність і наслідки кібератак на інформаційні системи значно зросли. За даними дослідження (Haris Uddin Sharif et al. [5]) за останні 12 років з 2010 по 2021 рік кількість різноманітних кібер інцидентів збільшилася в 3.9 разів, а втрати через кібератаки в одних тільки Сполучених Штатах Америки виросли до рекордних 6.9 мільярда доларів [5].

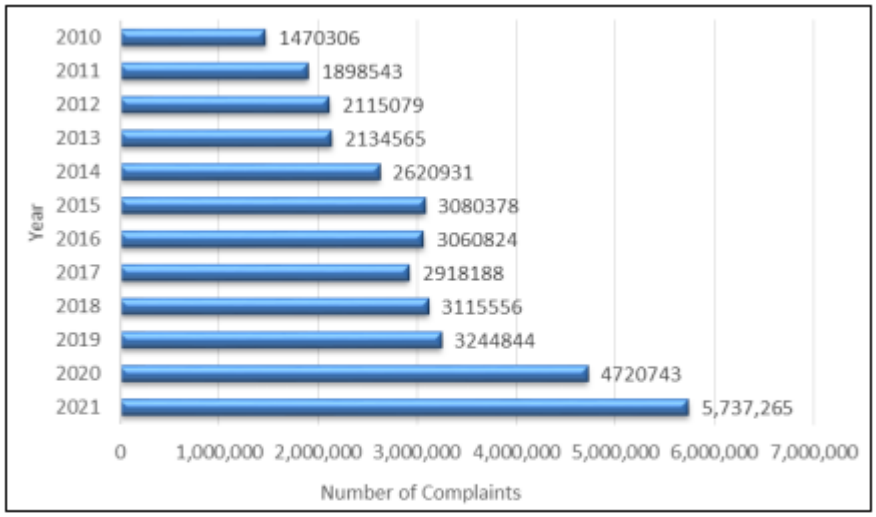


Figure 1 Number of Complaints of Cyberattacks last twelve years (2010-2021), U.S.A.

Мал. 1.1. Число скарг на кібератаки за останні 12 років.

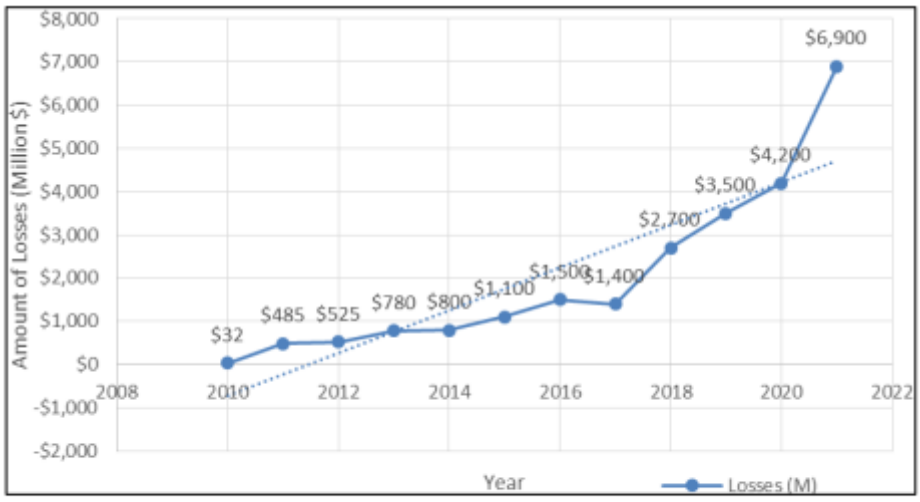


Figure 2 Amount of Losses of Cyberattacks last twelve years (2010-2021), U.S.A.

Мал. 1.2. Фінансові втрати від кібератак за останні 12 років (2010 -2022), Сполучені Штати Америки.

Потрібно також врахувати що багато компаній дуже часто взагалі не повідомляють про свої проблеми з кібербезпекою, щоби уникнути штрафів з боку регуляторів чи зменшити свої репутаційні втрати (46) або вони навіть не знають про те що на них була здійснена атака. Це означає, що насправді кількості кібер інцидентів і втрат від них може бути значно більше.

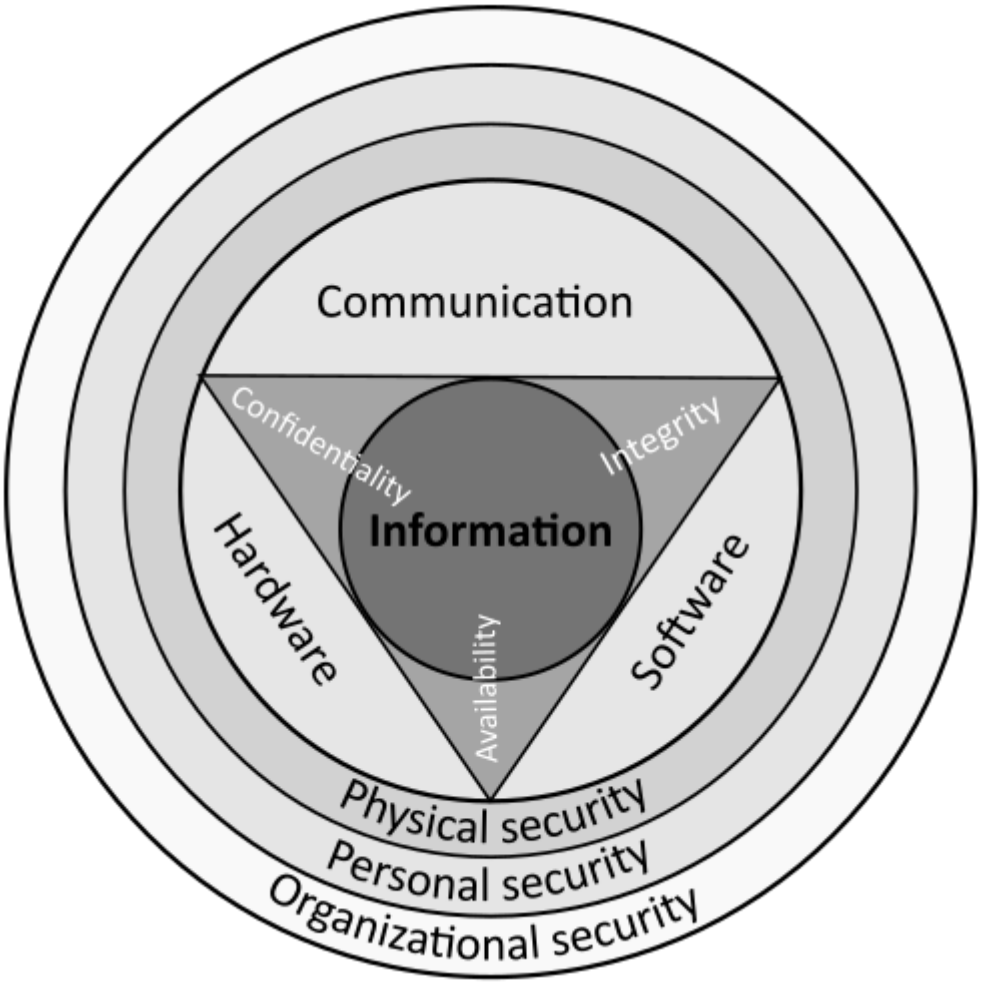
Отже, проаналізувавши можливі загрози які здатні завдати кібератаки, ми з впевненістю можемо констатувати що кібербезпека, на сьогоднішній день є запорукою безпечного функціонування організацій. При чому організацій будь-якого рівня, від малого бізнесу до державних органів.

1.1. Загальні принципи кібербезпеки

Кібербезпека - надзвичайно широка наука. Це все обумовлено тим що на світі існує безліч різних програмних рішення, а також різного обладнання. Саме через це розмаїття можливих варіантів ПЗ, ми маємо широкий спектр знань, які необхідні для роботи в кібербезпеці. Однак незважаючи на це, в своєму фундаменті вона має декілька основоположних принципів побудови захищеної системи.

Один з них - це тріада CIA, який розшифровується як:

- Confidentiality
    - Це забезпечення конфіденційність для важливих, не публічних даних
    - Тобто це збереження інформації, яка є таємною, а її розкриття може спричинити шкоду, в секреті
- Integrity
    - Це здатність системи зберігати дані в цілісності і захищати їх від пошкодження або навмисної зміни
- Availability
    - Це забезпечення доступності системи до використання.



Мал 1.3. Тріада CIA. Взяті до уваги всі можливі аспекти захисту інформації, а також всі їх рівні, від фізичного захисту до організаційного [7].

Якщо кожен з цих принципів достатньо дотримується тоді система є відносно захищеною (6). Відносно, тому що ніколи не можна виключати ризику злому та проникнення.

Тепер, знаючи основні принципи кібербезпеки, ми можемо зробити висновок що боти можуть порушувати усі три компоненти. Проте найбільше за всіх страждає саме Конфіденційність. Це відбувається тому що, боти це програми автоматичного доступу до закритої частини вебсайту. Вони створені спеціально для того, щоби обходити певні обмеження цих веб сайтів і успішно автентифікуватися в них. (47) Відповідно такі системи, можуть бути використані для автоматичного викрадення чутливих даних з акаунтів користувачів.

З приводу інших аспектів моделі безпеки, ми про них поговоримо у наступному розділі.

Отже, автоматизовані програми для роботи з аккаунтами, при використанні в неправомірних цілях можуть завдати великої шкоди інформаційній безпеці. Ось чому ми повинні розуміти загрози, які несуть ці програми, вміти правильно виявляти їх активність та вчасно зупиняти такі атаки.

1.2. Кібербезпека вебресурсів

З усе більшим розвитком Інтернету - все більше людей використовують різні інтернет ресурси. Соціальні мережі, магазини, державні послуги, послуги банків та багато чого іншого, уже давно доступне для своїх користувачів в онлайн режимі. За ростом числа користувачів та активності використання слідує і збільшення кількості кібератак, як на користувачів цих інтернет платформ, так і на самі платформи. В даному розділі, я хотів би розглянути деякі різновиди атак, які можуть здійснюватися на веб ресурси, переглянути деякі наочні приклади атак попередніх років, а також більш детально зануритися в ті методи, які використовує моє рішення.

1.2.1. Різновиди атак на вебресурси

Найбільш розповсюджені атаки на вебресурси описані в проекті OWASP top 10 [2]. OWASP - це організація створена для того щоби допомагати спеціалістам в інформаційних технологіях боротися з проблемами та вразливостями на в їх програмному забезпеченні [3]. Тоді як OWASP top 10 - це проект цієї організації направлений на виділення найбільш часто розповсюджених проблем з безпекою саме на веб-ресурсах.

Статистика 2021 року показує наступні десять найпоширеніших вразливостей веб-сайтів:

1. A01:2021-Порушений контроль доступу

a. Отримання неправомірного доступу до закритої інформації

2. A02:2021-Криптографічні збої

a. Проблеми з криптографією і шифруванням на веб-сайті

2. A03:2021-Ін'єкції

a. Вбудовування зловмисного коду або команди в легітимні додатки. Найбільш поширені види ін’єкцій - SQL ін’єкція, Міжсайтовий скриптинг, вбудовування JavaScript коду в легітимний додаток, а також Command line ін’єкція.

2. A04:2021-Незахищений дизайн

a. Погано продуманий з точки зору безпеки дизайн. Наприклад, персональні дані прямо у відкритій частині сайту

2. A05:2021- Некоректна конфігурація безпеки

a. Порушення норм і правил конфігурації веб-сайтів їх розробниками.

2. A06:2021- Вразливі та застарілі компоненти

a. Невчасно оновлене програмне забезпечення - це потенційний ризик для безпеки ресурсу

2. A07:2021 - Помилки в ідентифікації та автентифікації

a. Проблеми з ідентифікацією користувача на веб-сайті під час його використання ресурсу

2. A08:2021 - Збої цілісності програмного забезпечення та даних

a. Зловмисна зміна програмного забезпечення під час його створення або при атаці на джерело розгортання додатків, атака на CI/CD цикли.

2. A09:2021-Помилки реєстрації подій та моніторингу безпеки

a. Проблеми зі достатнім моніторингом небезпечних інцидентів та подій, які були викликані зловмисними діями

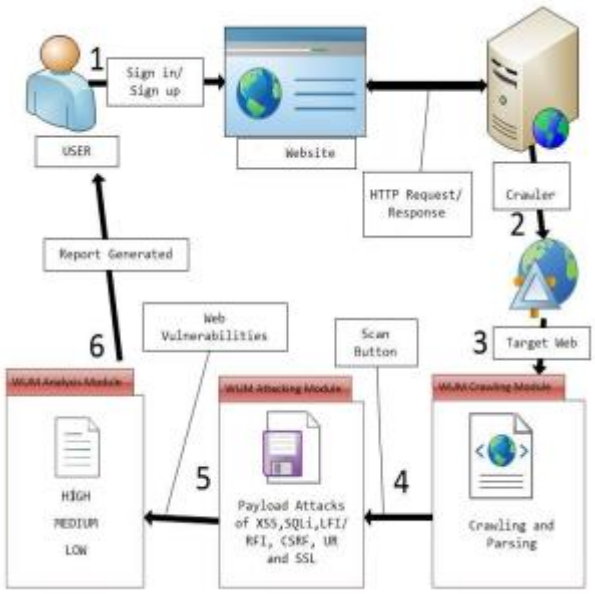
2. A10:2021 - Підробка запиту на стороні сервера

a. Виконання зловмисного запиту, без достатньої його перевірки. Може спричинити отримання закритої інформації або доступу до внутрішніх сервісів веб-сайту.

Серед цих найпоширеніших ми зосередимо увагу на перших трьох вразливостях та розглянемо сучасні технології пошуку цих вразливостей.

1.2.2. Автоматичні сканери пошуку веб вразливостей як інструмент моніторингу стану безпеки вебресурсів

На сьогоднішній день існує безліч автоматичних сканерів для пошуку веб вразливостей. Їх принцип роботи проілюстрований на малюнку 1.4. [4].



Мал. 1.5. Ілюстріція роботи сканеру на веб вразливості.
1.Етап - користувач, створивши аккаунт на цільовому веб сайті,
2. Передає сесію сканеру на вразливості.
3. Сканер запускає свій модуль пошуку по всьому веб-сайту займаючись парсингом і кровлінгом по ньому.
4. Знайшовши потенційні місця виконання вразливостей сканер запускає свій модуль для атак, тестуючи проблемні місця веб-сайту найпоширенішими варіантами веб вразливостей
5. Після закінчення тесту атакуючий модуль передає ці дані, модулю для аналізу вразливостей
6. Модуль аналізу створює звіт на основі отриманої інформації і передає його користувачу.

Сканери на вразливості часто використовують як в захисних цілях дослідники на вразливості так і зловмисники, які шукають вразливості для їх злочинного використання. Найчастіше зловмисники в своїй роботі шукають саме 3 вразливість з OWASP Top 10 [2], тому що через неї найпростіше згодом отримати неправомірний доступ і закріпитися на атакованому веб ресурсі (48).

Проте одними з найважливіших моментів в роботі сканерів, є не тільки їх здатність до пошуку вразливостей, але також їх всебічна доступність та спроможність до інтеграції з іншим програмним забезпеченням.

Саме тому станом на сьогодні, сканери є незамінними утилітами для роботи з тестуванням на вразливості. Одним з прикладів такої роботи може бути їх інтеграція з програмами-вимагачами, які використовують їх для пошуку вразливостей всередині мережі і для латерального просування по системі. (49)

1.2.3. Використання ботів для неправомірного доступу до цільового серверу

Однак ми сконцентруємо свою увагу на використання сканерів саме в ботах. Так як боти це спеціальне програмне забезпечення, воно може бути розширене будь-якими засобами для роботи з вебсайтами. Виходячи зі своєї специфіки, боти можуть бути обладнані сканерами на вразливості або навіть зловмисними утилітами віддаленого доступу, які вони можуть залишати після успішної атаки. Можливості цих програм, по суті, обмежені лише можливостями команди розробників. З цього слідує, що, так як боти по своїй логіці навчені обходити аутентифікацію та ідентифікацію вебресурсів, то вони можуть бути також використаними в зловмисних цілях.

Наприклад, після проходження аутентифікації бот може запустити свій вбудований сканер на вразливості, який буде шукати вразливості на вебсайті, а потім, при знаходженні потрібних, виконати корисне навантаження і отримати неправомірний доступ до цільового сервера. Ці загрози існують уже досить давно, прикладом можна назвати створення автоматизованої ботнет мережі під назвою Mirai (50). Ботнет - це сукупність заражених сервісів, які віддалено керуються з контрольного серверу використовуючи вірусні агенти.

Отже, окрім не прямих загроз через маніпулювання аудиторією та порушення конфіденційності веб ресурсу автоматизовані боти можуть виконувати також і інші зловмисні атаки на цільовий вебресурс. Це дозволяє у повній мірі зрозуміти, які загрози може нести таке програмне забезпечення та підійти до вирішення проблеми комплексно.

Щоби захиститися від таких неправомірних дій, необхідно розглянути, яким чином були вирішені такі кібер інциденти раніше. Для цього було обрано декілька наочних прикладів кібер атак на великі технологічні компанії, а також створення найбільшої ботнет у світі (за різними даними вона нараховувала близько 250 тисяч захоплених девайсів по всьому світу).

1.3. Аналіз кібератак, здійснених на вебресурси фінансових установ різних форм власності в Україні

1.3.1. Використання експлойтів 0-го дня для здійснення кібератак

Розглянемо декілька прикладів нещодавніх атак на вебресурси. В першу чергу потрібно згадати російську атаку на українську інфраструктуру та органи державної влади в 2017 під назвою NotPetya. Ця атака вийшла далеко за межі України і спричинила збої в таких міжнародних компаніях, як Maersk, Mondelez та багато інших [8]. Програма працювала наступним чином:

1. Перші зараження були проведені через атаку на ланцюги поставок та зміну файлів оновлення для програми MEDoc, яка використовувалася для обліку та подання звітності до контролюючих органів усіма компаніями в Україні та для обміну юридично значущими первинними документами між контрагентами в електронному вигляді. Атака на ланцюг поставок - це атака в якій атакуючі компрометують мережу, яка належить одному з менш захищених частин в загальній системі. Наприклад, у випадку з NotPetya - хакери атакувавши одну компанію, яка розробила програмне забезпечення для багатьох інших, змогли скомпрометувати велику кількість інших компаній.

2. Згодом потрапивши всередину комп'ютерної системи програма перехоплювала адміністративний контроль над системою.
3. Викрадала всю необхідну інформацію для Smb протоколів
4. Реплікувала себе
5. Поширювалася латерально по мережі (тобто отримувала доступ до інших систем всередині однієї мережі)
6. Для того, щоби отримати адміністративний доступ до інших мереж, вона запускала експлойт EternalBlue.
7. Згодом програма наперед встановлювала в системі певні дату і час виконання свого шифрувального модуля
8. Запускала шифрування
9. Змінювала програму для запуску системи на свій код
10. При запуску системи відкривалось повідомлення про необхідність сплатити 300 $ в криптовалюті біткойн.

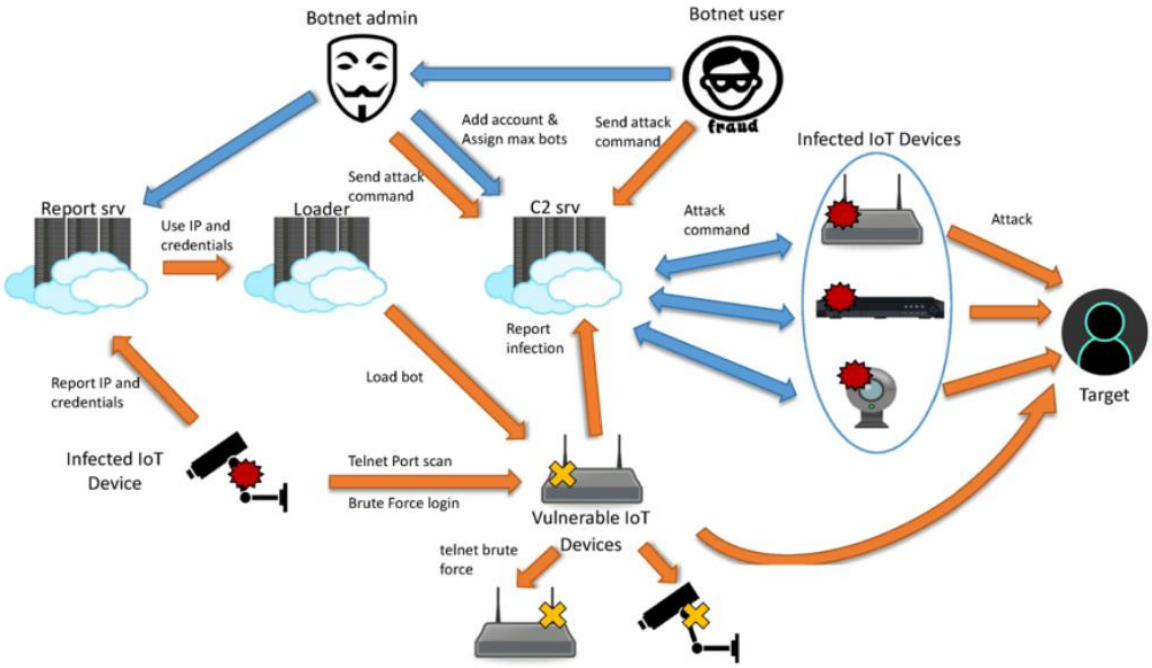
В результаті атаки була зупинена робота більше ніж 300 державних підприємств, установ та приватних компаній в Україні та у світі. Збитків було завдано на більше як 10$ мільярдів доларів [9). Дослідники виявили, що ця шкідлива програма була направлена не так на збір викупу, як на знищення критичної інфраструктури України. Проаналізувавши їх звіти, можна впевнено зробити висновок, що використання для атаки експлойтів 0-го дня або таких шкідливих програм, що ще не були використані ніде (тобто спеціально створені під конкретну задачу чи об'єкт враження) значно ускладнюють захист “жертви” атаки. Проте, оскільки кібербезпека все більше розвивається, створення таких експлойтів стає все більш затратною справою в грошовому і часовому еквіваленті (51). Саме тому, більшість кібер злочинців не мають можливостей шукати або купувати такі експлойти. Найчастіше вони використовуватимуть доступні уже відомі експлойти і будуть направляти свої атаки на застарівше програмне забезпечення.

Однак на попередження таких атак, можна також використовувати системи розумного моніторингу, які будуть реєструвати підозрілу активність всередині комп’ютерної мережі і не тільки зупиняти потенційно небезпечну активність, а й встановлювати та локалізувати її першоджерело. Цей тип аналізу називається - поведінковий, і на даний момент він застосовується у більшості антивірусних програмних продуктах. Більш детально ми поговоримо в наступному розділі про використання ML для перевірки веб-ресурсів на вразливості.

У підсумку, можна з упевненістю сказати, що лише своєчасне оновлення програмного забезпечення і системні та систематичні заходи з виявлення загроз, можуть врятувати систему від пошкодження або від втрати важливих даних.

1.3.2. Використання систем з центром управління для здійснення кібератак

NotPetya - це програма черв. Вона самостійно реплікується та розповсюджується через всі вразливі системи допоки не виконає свою зловмисну функцію. В неї немає централізованого управління, вона не отримує команди ззовні і діє за чітко сформульованими та наперед визначеними задачами. Однак не всі злочинні програми такі - існують також і інші, що здатні виконувати будь-які команди, які надійшли від центрального серверу управління. Прикладом такого, зловмисного програмного забезпечення є безумовно найбільша мережа ботів, яка коли-небудь була створена. Це ботнет - Mirai. Під час апогею свого поширення ця мережа досягала до 300 тисяч заражених IOT девайсів і могла здійснювати атаки на такі цілі як провайдери Інтернету та багато інших [10].



Малюнок 1.6. Принцип роботи системи Mirai [11].

Сама система складалася з таких основних компонентів, як: контролюючий сервер, сервер завантаження шкідливого програмного забезпечення, сервер збору інформації, а також саме шкідливе програмне забезпечення [10].

Принцип роботи наступний:

1. Заражений перший девайс

a. Передає інформацію про зараження контролюючому серверу

b. Запускає сканер на пошук інших вразливих девайсів в мережі інтернет

c. При їх знаходженні через брутфорс атаки підбирає пароль до наступного девайсу за протоколом телнет.

d. Передає пароль та IP адресу наступного девайсу одному з серверів збору інформації

2. Сервер збору інформації

a. Передає ці дані визначеним серверам завантаження

2. Сервер завантаження

a. Отримує дані з серверу збору інформації

b. Отримує доступ до девайсу через логін і пароль

c. Завантажує шкідливе програмне забезпечення на девайс.

d. Оновлює шкідливе програмне забезпечення

2. Сервер управління

a. Отримує інформацію про нові зараження

b. Передає інформацію про наступні цілі зараженим девайсам

c. Запускає команди атакувати або будь-яку іншу команду.

2. Зловмисник

a. Керує сервером управління

b. Запускає команди для атаки

2. Клієнт зловмисника

a. Оплачує гроші за проведення атак на вигідні для нього цілі

Отже, така система дає набагато більше гнучкості для її розробників. Тут вони можуть використовувати її в довгостроковій перспективі, оновлювати своє програмне забезпечення, а також вести свою діяльність довгий час без того щоби бути поміченими правоохоронними органами.

На даний момент основний спосіб боротьби з таким програмним забезпеченням - це активний моніторинг усіх вразливих девайсів, зміна паролів на більш захищені, створення навмисно вразливих систем-пасток (ханіпотів) для виявлення та ідентифікації злочинної діяльності [10].

Однак шкідливе програмне забезпечення дуже швидко еволюціонує і підвищує свої можливості захисту від ідентифікації та активно протидіє заходам направленим на зупинку її роботи.

Наприклад, згідно з дослідженням (52) шкідливе програмне забезпечення може маскувати свій трафік через навмисне створення штучних веб запитів, що нагадують легітимні запити. Також вони здатні ідентифікувати ханіпоти через використання штучного інтелекту (цитати). Через такий стрімкий розвиток шкідливого програмного забезпечення дослідникам з кібербезпеки також необхідно розвивати свої технології з виявлення таких загроз. Цим розвитком є створення систем для виявлення зловмисних програм на основі штучного інтелекту для проведення поведінкового аналізу.

1.4. Принципи роботи систем контролю над ботами та їх використання у роботі з соціальними мережами

Тепер коли ми проаналізували принципи роботи основних програм, які можуть використовувати зловмисники для своєї діяльності, можна зрозуміти

принципи роботи систем для контролю над ботами, а також ті загрози, які вони несуть у разі їх використання.

Проте, слід зазначити, що основна задача таких систем для автоматичного контролю полягає у роботі з соціальними мережами. А саме: обхід аутентифікації і ідентифікації користувачів, виконання замовлених медійних задач, а також ведення інформаційної війни.

Показовим прикладом таких операцій безперечно є використання таких ботоферм під час російсько-української війни. Термін ботоферма - в даному випадку використовується, як визначення сукупності ботів, що об’єднані в одну мережу.

З 2014 року і по сьогоднішній день росія активно використовує соціальні додатки, такі як Твіттер, Фейсбук, Телеграм та інші, для ведення своєї агресивної медіа кампанії проти України [12]. Багато з цих ботоферм, також були зупинені внаслідок правоохоронної діяльності Служби Безпеки України на території України [12]. Основні методи роботи цих систем - це використання СІП фонів, тобто серверів для роботи з телекомунікаціями, а також автоматизованих клієнтів - ботів.

Малюнок 1.6. Базова схема роботи ботоферми.

Базовий принцип роботи стандартної системи ботів:

1. Кластер клієнтських ботів

a. Зазвичай це або програмне забезпечення або смартфони на основі операційної системи андроїд низької якості

b. Кожен бот, отримує свою сім карту і відповідно свою IP адресу для конекту до мобільної мережі

c. З цієї мобільної мережі здійснюється робота клієнта

d. Виконує автоматичні дії з розвитку аккаунту (лайки, поширення, коменти)

e. Отримують нові акаунти для роботи від центрального сервера

2. Сіп фон сервер

a. Сервер для використання мобільної мережі

b. Надає доступ в інтернет для клієнтського кластеру

2. Центральний сервер

a. Надсилає команди про виконання клієнтам
b. Отримує нові акаунти від провайдера акаунтів
c. Створює нових клієнтів
d. Передає їм нові IP адреси
e. Керує роботу СІП фон сервера

2. Провайдер акаунтів

a. Третя сторона, яка виконує роботу з реєстрації акаунтів в соціальних мережах

Такі системи, часто використовують відкрите API деяких соціальних мереж, системи автоматизації роботи веб браузерів або низько-якісні смартфони андроїд з отриманими рут правами на них, для автоматизації роботи з соціальними мережами. (53)

Найчастіше використовуються саме два останніх варіанти, тому що загалом соціальні мережі, як у випадку з Facebook не надають свої API в публічний доступ. (54).

Отже, саме такі системи і використовуються для поширення дезінформації. Службам технічної підтримки соціальних мереж добре відома ця проблема, в деяких випадках вона може сильно впливати на вартість акцій на біржі і рівень довіри суспільства до цього бізнесу. Як, наприклад, у випадку Твіттера і його купівлі Ілон Маском (55). Тому кожна компанія докладає великих зусиль для створення систем протидії ботам. Проте про системи боротьби з недоречним використанням медійних платформ, що не передбачені умовами роботи з ними, йтиметься окремо при детальному аналізі системи протидії фроду, у наступному розділі роботи які нестандартні методи протидії злочинним системам застосовуються на сьогоднішній день. А зараз я спробую розглянути які психологічні аспекти та патерни поведінки користувача слід врахувати щоб побудувати найбільш дієву систему протидії фрогу у фінансових організаціях.

1.5. Психологічні аспекти, що використовуються як для виявлення, так і для побудови поведінкової моделі бота

У сучасному світі майже кожен користувач інтернету має облікові записи в різних соціальних мережах. Люди, проводячи своє дозвілля в інтернеті, поглинають багато інформації: як корисної, так і тієї, що завдає шкоду різного характеру (матеріальні збитки від імпульсивних покупок, пригнічений моральний стан від негативних новин, заздрість від слідкування за життям успішних людей тощо). Але інтернет, як і будь-яке інше джерело інформації, може також використовуватися для поширення різних ідей: політичних (агітація за певного політичного діяча), релігійних (схиляння до слідування за ідеями будь-якої релігійної групи), маркетингових (навіювання потреби в тому чи іншому товарі), та інших.

Далі під терміном зацікавлені особи розумітимемо людину чи групу людей, яка бажає просунути ідею чи ідеї.

До появи інтернету ідеї від зацікавлених осіб та організацій до людей долинали через газети, телебачення, а також через спеціально навчених спеціалістів. Але з появою інтернету, у міру збільшення часу, що проводиться в ньому звичайним користувачем, вищезазначені засоби донесення інформації стали втрачати силу впливу. З тих пір зацікавленим особам доводиться підлаштовуватися під сучасні умови та просувати свої ідеї за допомогою інтернет-ресурсів.

Нам відомо, що інтернет є величезною мережею, і контактувати з кожною людиною особисто майже неможливо - це вимагає великих фінансових і людських ресурсів. Тому для економічно виправданого, ефективного просування ідей зацікавленим особам доводиться використовувати засоби масового впливу, такі як інтернет-ресурси новин, відомі інтернет-персоналії, а також доводиться автоматизувати роботу з акаунтами людей. Прикладом останнього є боти, які дозволяють автоматично взаємодіяти з акаунтами користувачів у соціальних

мережах: робити масове розсилання повідомлень, відзначати під фотографіями в Instagram, додавати до груп у Facebook тощо.

Актуальність використання роботів дуже висока. Це обґрунтовується кількома причинами.

Першою причиною є те, що різні групи людей використовують різні джерела інформації в інтернеті. Наприклад, підлітки та молоді люди воліють проводити час у соціальних мережах, де переважає розважальний контент, наприклад, Youtube, Instagram, TikTok і Telegram, а канали новин і сайти переглядають нечасто; старше покоління людей віддає перевагу більш звичним їм джерелам інформації, таким як новинні сайти, Youtube-канали відомих їм з часів телебачення журналістів, аналітиків, культурних діячів.

Навіть при використанні однієї і тієї ж платформи, наприклад, Youtube, канали, що переглядаються у людей різної вікової категорії можуть відрізнятися за тематикою, жанром, середньою тривалістю відео, якістю зйомки і так далі. Отже, щоб охопити максимально широку аудиторію, що складається з людей різного віку, статі, які мають різні інтереси, доведеться задіяти багато різних джерел інформації та впливових особистостей, провівши перед цим їхній аналіз вручну. Таким чином, процес просування ідеї вимагає чималих фінансових, людських та тимчасових витрат.

При автоматизації вищезгаданих процесів за допомогою ботів, маючи набір даних облікових записів всіх людей (імена або id облікових записів, адреси електронних пошт), донести інформацію до різних груп людей стає простіше - при розсилці інформації бот може не враховувати вік, стать, інтереси конкретного користувача. Таким чином, ідеї поширюються одразу серед усіх людей.

Друга причина актуальності використання роботів також пов'язана з економічною практичністю. Зацікавлені особи можуть мати на меті розповсюджувати ідеї лише серед певної групи людей. Розглянемо гіпотетичний приклад політичної агітації. Ми знаємо, що в Україні лише повнолітній

громадянин має право голосувати під час виборів президента. Тому, якщо один із учасників перегонів за місце президента вирішить просувати свою кандидатуру з використанням інтернет-ресурсів, його цікавитимуть лише повнолітні громадяни України – це його цільова аудиторія.

Однак не всі користувачі залишають персональну інформацію про себе. При взаємодії з акаунтом може виявитися так, що його власник не входить до цільової аудиторії (неповноліття, громадянство іншої країни, інші причини). Подібних облікових записів може бути багато. Автоматизація, яку забезпечує використання бота, мінімізує витрати на взаємодію з людьми, які не є цільовою аудиторією.

Нам зрозуміло, що бот – це ефективний інструмент, який робить процес просування ідей серед користувачів інтернету легким та доступним. Проте, щоб ідею успішно поширити, необхідно враховувати психологію впливу на користувача як на людину – ефективність донесення ідеї залежить від цього, як саме бот реалізує цей процес.

1.5.1. Принципи переконання людей

Протягом минулих десятиліть було проведено багато досліджень, які вивчають процеси соціального впливу на дії та звички інших людей. Robert B. Cialdini, Roselle L. Wissler і Nicholas J. Schweitzer [29] ідентифікували 6 фундаментальних принципів переконання: Принцип симпатії, Принцип авторитету, Принцип дефіциту, Принцип сталості, Принцип взаємності та Принцип Соціального схвалення.

Принцип симпатії. Очевидно, що люди більше зазнають впливу тих, хто їм подобається. Це грунтується на різних чинниках: фізична привабливість, компліменти, спільні зусилля, і найголовніше – схожість.

Robert B. Cialdini et al [29] продемонстрували, що схожість імен одержувача та відправника змушує людей частіше відповідати на листи з опитуваннями.

Схожість імен – це незначний чинник, вагомішими чинниками можуть бути загальний інтерес, приналежність до загальної групи, наявність спільної мети. Більш вагомі чинники призводять до появи сильнішої симпатії.

Принцип авторитету. На людей впливають ті, кого вони вважають авторитетами. Це пояснюється тим, що авторитет, ймовірно, має більше досвіду та знань. Наприклад, наявність дипломів та нагород у кабінеті лікаря на 34% підвищують шанс того, що пацієнт прислухається до рекомендацій, і виконуватиме їх [29]. Важливо відзначити, що авторитетність має бути встановлена до спроб впливу.

Принцип дефіциту. Предмети стають бажанішими, якщо вони менш доступні. Дослідження показують, що переконання, засновані на потенційній втраті у разі невикористання шансу, більш переконливі, ніж переконання, що наголошують на тому, що буде отримано при використанні того ж шансу. Прикладом успіх продажів автомобіля Oldsmobile, який з'явився після анонсу припинення випуску даної моделі виробником [29].

Принцип сталості. Люди прагнуть зберігати постійність щодо своїх думок, тверджень та дій. Щоб переконати людину щось зробити, її угоду із зобов'язанням потрібно зробити публічною, наприклад, зробити запис обіцянки (особливо, якщо цей запис побачать інші люди). Позитивний результат цієї дії описаний у дослідженні Robert B. Cialdini et al [29], де зміна фрази реєстратора з "Будь ласка, зателефонуйте, якщо у вас зміняться плани" на "Ви зателефонуєте, якщо у вас зміняться плани, будь ласка?" знизила відсоток неявки по броні до ресторана з 30% до 10%.

Принцип взаємності. Найчастіше люди бажають віддати те, що отримали: це може бути обмін як матеріальними цінностями, так і нематеріальними (увага, інформація, поступки, повага). Звичний, ранковий обмін посмішкою з колегою є прикладом цього принципу.

Принцип соціального схвалення. Люди бажають отримати схвалення в особі інших людей щодо їх рішень: «Якщо багато людей, таких як я, роблять це так само, значить, я все роблю правильно». У ситуаціях, коли людина відчуває невпевненість, цей принцип проявляється найсильніше.

Одночасне застосування кількох принципів призводить до більш вираженого впливу прийняття рішень. Важливо розуміти, що перераховані вище переконання можуть використовуватися як з благими, так і з руйнівними намірами [29].

Основна причина, через яку ці принципи ефективні, полягає в тому, що вони служать евристиками (розумовими спрощеннями) – способом обробки інформації, на який зазвичай покладаються, – люди не завжди мають можливість або мотивацію до дуже активного і ретельного мислення [32]. Крім того, ці принципи також ефективні при багатьох онлайн-взаємодіях [33].

1.5.2. Роль соціальних мереж у переконанні людей

Людьми легко маніпулювати через соціальні мережі, особливо при застосуванні психологічних тактик [30]. Цей тип впливу належить до соціального впливу. Соціальний вплив — це зміна поглядів, переконань чи поведінки людини, що відбувається внаслідок реального чи уявного тиску [31]. Простота підключення до інтернету, а також залежність від нього зробила людей сприйнятливими до цього впливу, а значить, що люди можуть бути використані з корисливою метою з меншим зусиллям.

Використовуючи соціальні мережі, люди не схильні до ясного мислення [32]. Людина схильна довіряти джерелу інформації, яке йому подобається [30] - це може бути відома інтернет-персона, знаменитість, або будь-яка інша людина з приємною зовнішністю.

Дослідження Regan [34] показало, що люди схильні купувати лотерейні квитки у людини із приємнішою зовнішністю. Даний принцип поширюється і на соціальні мережі – якщо людина рекомендує той чи інший товар, або ділиться своїми ідеями, закликає до певної дії, то слухачі готові прислухатися і наслідувати заклик, якщо персона їм подобається.

У соціальний мережах існує багато способів маніпулювання людьми. Найчастіше вони є методиками переконання людей з ціллю фінансового обману.

Протягом багатьох років у різних мережах використовується фішинг-метод «мандрівник, що застряг» - лист, надісланий нібито від друга, який переконує одержувача у фінансовій біді друга, з проханням матеріально допомогти. Друг, як людина, яка вам подобається, викликає довіру. Тому отримувачі нерідко відкривають такі листи, і вивчають їх, потрапляючи у пастку шахраїв.

Ще один спосіб маніпуляції, до яких вдаються зацікавлені особи, заснований на симпатії, що виникає у жертви. Halper [35] описує шахрайство в соціальній мережі, в процесі якого шахрай, користуючись нагодою Дня Валентина, зв'язується з потенційною жертвою, викликає у неї романтичний інтерес, і через деякий час спілкування переконує надіслати гроші.

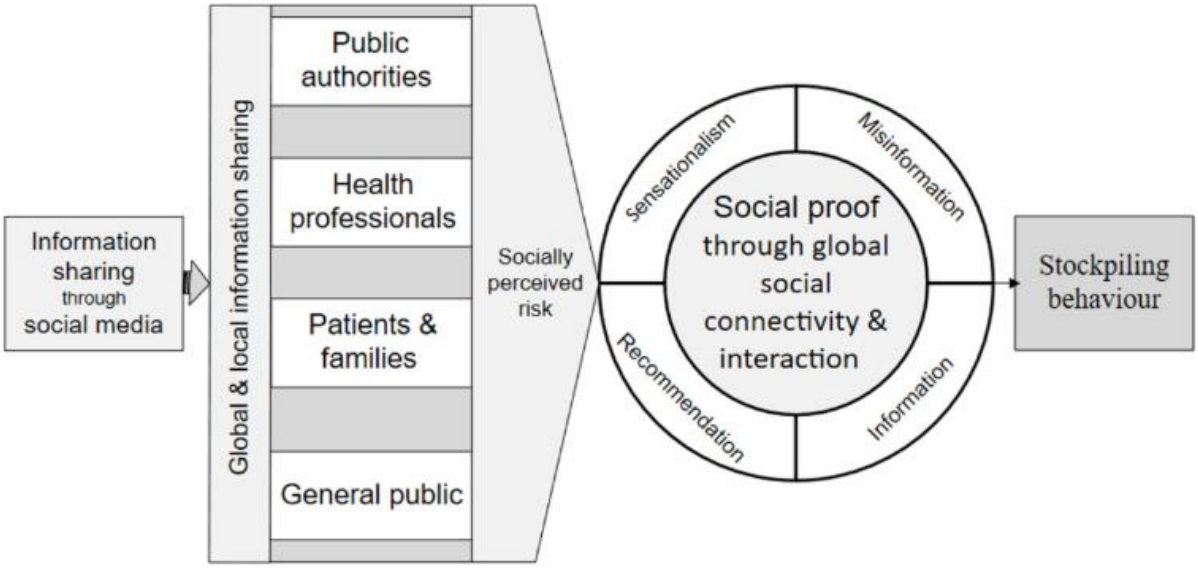
Соціальна мережа відіграє велику роль у цьому шахрайстві: жертва не може перевірити, хто сидить по той бік монітора - чи відповідають фотографії, опис профілю, наміри людині? Соціальна мережа дозволяє шахраям відчувати безкарність за свої дії, які здійснюються без розкриття його особистості.

І навпаки, в соціальній мережі людина, яка не приховує свою особистість, може прикрашати її, або навішувати на себе неіснуючі ярлики – називати себе лікарем, науковцем, політичною фігурою. Для інших людей ця людина може стати авторитетом, а отже, думка такої інтернет-персони впливатиме на них [29].

Як я вище зазначив, люди схильні до формування підвищеного інтересу до дефіцитних продуктів [29]. Зацікавлені особи, які бажають переконати людей перейти за посиланням та провести інші дії, можуть створити штучний ажіотаж за

допомогою фраз на кшталт «пропозиція обмежена», «встигни зареєструватися» та подібних, хоча насправді обмеженості часу чи кількості немає, або обмеженість перебільшена.

У соціальній мережі інформація розповсюджується дуже швидко – в один клік користувач може поділитися інформацією з десятками, сотнями та більше людей. А значить, інформація, що здається важливою і корисною певному користувачеві, може бути швидко поширена. Цим часто користуються зацікавлені особи, які переконують людей у корисності інформації чи ідей, і просять поділитися ними з іншими. У такому разі виходить, що роботу з поширення беруть на себе самі люди, а не зацікавлена особа.



Мал. 1.7. Концепція формування поведінки переконання у потребі створення запасів через соціальні мережі [30].

На Мал. 1.7. відображено концепцію, яка використовується зацікавленими особами в соціальних мережах для переконання користувачів у наявності потреби того чи іншого продукту. Використовуючи одночасно декілька принципів переконання (принцип авторитету, принцип соціального схвалення, принцип симпатії, принцип дефіциту), зацікавлені особи нав'язують інформацію

користувачеві, спонукаючи його до дій, саме, до запасання товаром, наголошуючи на його потенційному дефіциті.

1.5.3. Вплив ботів на соціальні мережі

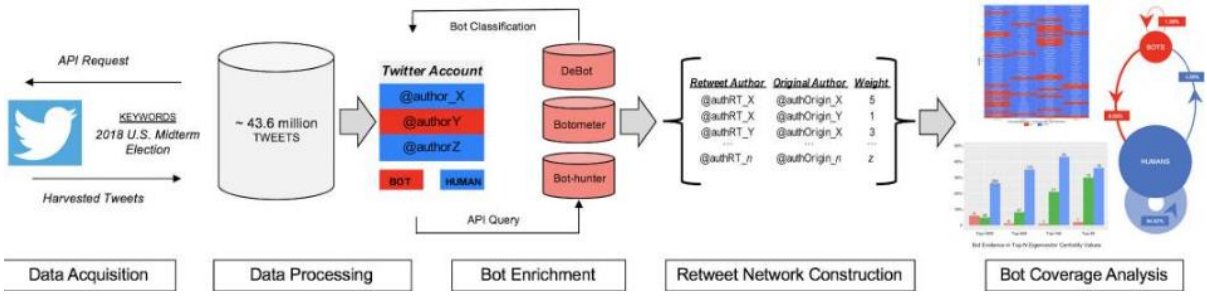
Бот у соціальній мережі є програмним агентом, який автономно спілкується у соціальних мережах. Повідомлення (наприклад, твіти в Twitter), які він поширює, можуть бути простими та працювати в групах та різних конфігураціях з частковим контролем людини (гібридна версія) за допомогою алгоритму. Боти також можуть використовувати штучний інтелект для поширення повідомлень у звичайному людському діалозі.

Творці роботів переслідують певну мету, яка приносить їм вигоду. Це може бути як поширення ідей з отриманням матеріальної вигоди у перспективі, так і поширення шкідливої інформації для отримання миттєвої фінансової вигоди.

Ілюзію соціального схвалення в соціальних мережах допомагають створити програми-боти, що генерують штучні позначки «мені подобається», коментарі під фото та записами на сторінці людини, що підвищують кількість передплатників та друзів обраної сторінки Facebook, Twitter, Instagram та Youtube. Створення фальшивої підтримки викликає більший інтерес та підтримку персонажа, і у багатьох випадках спонукає людей витрачати гроші [30]. Таким чином, люди з більшою ймовірністю відвідають сторінку, натиснуть на відповідні посилання і поставлять «лайк» або «підпишуться» на сторінку, якщо побачать, що, ймовірно, багато інших роблять те саме.

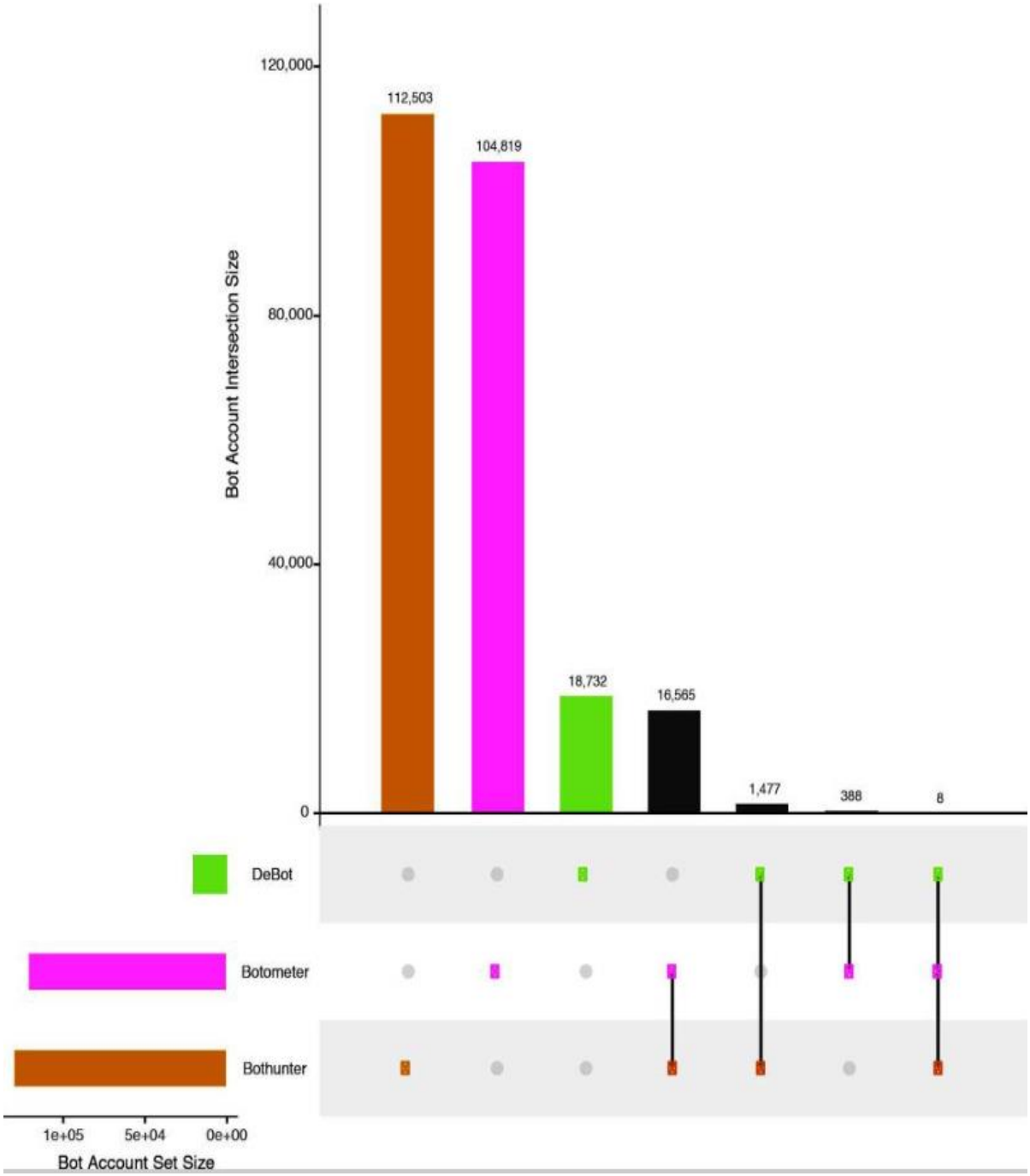
Іноді це може й не бути явною спробою зловмисного впливу (тобто, соціальний інженер – створювач боту, не краде особисту інформацію користувача безпосередньо), але в інших випадках це може дозволити іншим отримати доступ до конфіденційної інформації, яка може бути використана зловмисно.

Помилково створюючи видимість того, що багато інших людей натискають на посилання або що сторінка їм подобається, можна вплинути на більшу кількість людей, щоб вони зробили те саме, тому що вони часто дивляться на поведінку інших, не замислюючись про свої дії.



Мал. 1.8. Фреймворк аналізу соціальних ботів, який використовує кілька платформ виявлення ботів [41].

За оцінками, 9–15% активних облікових записів у Twitter можуть бути соціальними ботами. Аналіз наявності соціальних роботів відображено на Мал. 1.8. Щонайменше 400 000 ботів написали близько 3,8 млн твітів, що становить приблизно 19% від загального обсягу. Кількість ботів, що цілеспрямовано беруть участь у агітації, можна побачити на Мал. 1.9.



Мал. 1.9. Аналіз охоплення аудиторії ботами, виявленими під час проміжних виборів у США у 2018 році у Twitter, з використанням платформ виявлення ботів Botometer, Bot-hunter та DeBot [41].

Використання соціальних роботів суперечить умовам обслуговування багатьох платформ, таких як Twitter і Instagram, хоча певною мірою дозволено іншими, такими як Reddit та Discord. Навіть для платформ соціальних мереж, які

обмежують використання соціальних ботів, певний ступінь автоматизації призначений для забезпечення доступності API-інтерфейсів соціальних мереж.

Платформи соціальних мереж також розробили власні автоматизовані інструменти для фільтрації повідомлень, що надходять від роботів, хоча вони недостатньо просунуті, щоб виявляти всі повідомлення ботів [42].

1.5.4. Вплив ботів на вибори в різних країнах світу

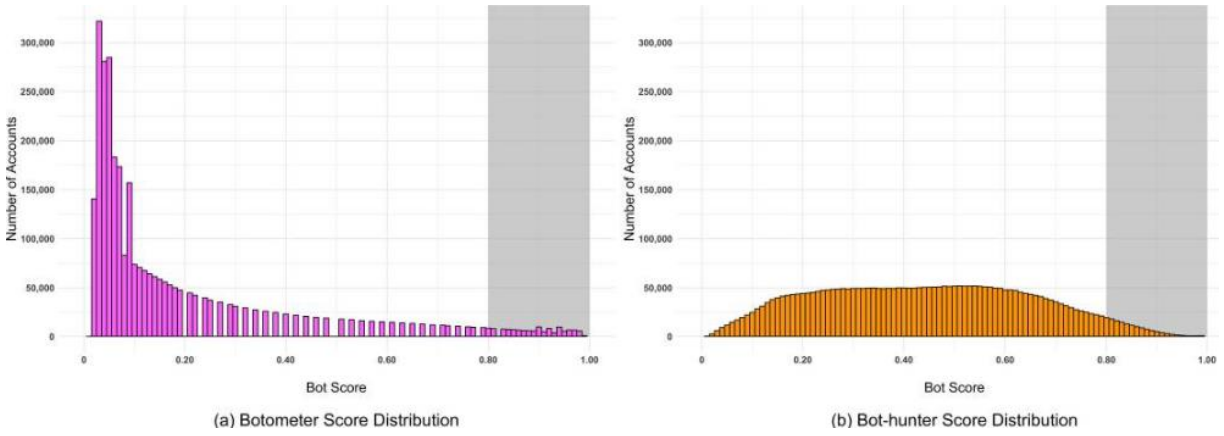
Люди прагнуть віддавати свій голос людині, яка їм подобається, і якій вони вірять. Потрібний образ кандидата легко побудувати у соціальній мережі – при цьому соціальні інженери використовують різні психологічні техніки впливу на групи людей.

Сайти соціальних мереж можуть швидко поширювати інформацію, включаючи контент новин, відео та вбудовані посилання, що призвело до значних змін у моделях спілкування та використання контенту соціальних мереж [36]. Наприклад, американці все більше займаються активним навчанням та пошуком інформації в соціальних мережах на додаток до пасивного навчання та пошуку інформації, які неминуче відбуваються у мережевих соціальних середовищах [37]. Це включає також вивчення політичної ситуації, кандидатів на виборах.

У соціальних мережах є можливість створювати сенсації та неправильно тлумачити інформацію, що надходить від різних офіційних осіб, наприклад, від державних установ та політичних діячів [38]. Для цього можуть використовуватися боти, які розповсюджують через коментарі у Youtube, Instagram, Facebook тощо. інформацію, вигідну тому чи іншому кандидата. Намір часто полягає в тому, щоб ввести аудиторію в оману та змусити її повірити в те, що цільовий громадський діяч сказав щось (часто суперечливе чи провокаційне) [39].

Portland Communications у своєму звіті How Africa Tweets зазначили, що боти Twitter становлять понад 20% впливових осіб у таких країнах Африки, як Лесото та Кенія.

Боти також використовувалися у політичних перегонах у США за місце президента у 2016 році у Twitter. 15% усіх акаунтів у Twitter, які брали активну участь в обговоренні президентських виборів у США, були ботами. Після перемоги Дональда Трампа, боти припинили виявляти активність. Однак у квітні 2022 року, напередодні виборів у Франції, ці ж боти відновили роботу, підтримуючи Марін Ле Пен, вкрай правого кандидата на виборах у Франції. Деякі боти навіть перейшли на французьку мову [40].



Мал. 1.10. Розподіл балів облікових записів Twitter, представлених у сукупності твітів про проміжні вибори в США у 2018 році, з використанням (a) Botometer (рожевий) та (b) Bot-hunter (помаранчевий) платформ виявлення ботів [41].

Келлер та Клінгер [43] продемонстрували, що частка соціальних ботів серед підписників політичних партій Німеччини у Twitter зросла з 7,1% раніше до 9,9% під час виборчих кампаній. Частка активних соціальних ботів не зросла під час виборчих кампаній; партія AfD не мала більше ботів-фоловерів, ніж інші партії – навпаки, найменшу частку. Боти, яких вони виявили, майже не поширювали хештегів, пов’язаних із німецькою політикою.

Ймовірно, що за величезною кількістю ботом стоїть та сама група зловмисників, що прагне вплинути на вибори в країні за країною [40].

1.5.5. Використання ботів для вчинення протиправних дій у фінансовому секторі

Хоча більшість нещодавніх атак ботів були пов’язані з його активністю в соціальних мережах, ці платформи не єдині, які можуть бути об'єктами атак шахраїв. Як приклад, ми можемо представити цю проблему на ринку фінансових установ, таких як біржі криптовалют, банківські рахунки та фондові ринки. Тут можна виділити кілька проблем: а) безпека цінностей; б) ціна акцій.

Безпека цінностей - це безпека депозитів, які знаходяться на рахунках користувачів або всієї системи фінансової установи. Це можуть бути як прямі цінності, як гроші чи конвертовані цінності, так і ті, які пізніше можна перетворити на прямі цінності. Сьогодні, оскільки цифрові товари набувають все більшого поширення, доступ до них має бути дуже добре захищений.

Як приклад можна згадати крадіжку 281 мільйона доларів у BTC або ETH, яка отримала назву «Порушення безпеки KuCoin» і відбулася у 2021 році [44]. Це небезпечне порушення безпеки призвело до значних збитків для мільйонів користувачів криптовалютної біржі KuCoin.

Ціна акцій – це цінний папір публічної компанії або цифрового товару, такого як криптовалюта, який продається на біржі криптовалют. На ціни акцій можуть негативно вплинути кілька способів: зловмисні маркетингові кампанії проти компанії, порушення громадської безпеки та погана репутація керівника компанії. Те ж саме стосується криптовалюти, оскільки за кожною з них стоїть команда, а в більшості випадків юридична особа, яка опублікувала цю криптовалюту для громадськості, і на її ціну може вплинути негативна інформаційна маркетингова кампанія.

Автоматизовані загрози, такі як мережа ботів, можуть вплинути на безпеку як цінностей, так і курсу акцій. Оскільки бот — це програмне забезпечення, створене так, щоб виглядати схожим на реального користувача, він може вчиняти будь-які дії з веб-службою, яку використовує. Наприклад, ці мережі ботів можуть проникнути в обліковий запис користувача, використовуючи його облікові дані, і спробувати вилучити всі цінні речі з облікового запису користувача. З іншого боку, він може почати тестування на проникнення, намагаючись отримати доступ до конфіденційної інформації або знайти слабке місце в системі. Крім того, ці системи стають ще розумнішими завдяки використанню всіх можливих проривів у машинному навчанні (56).

Тому мережі ботів є реальними загрозами для фінансового сектору економіки, з якими слід боротися. Щоб вирішити цю проблему, я хотів би представити своє рішення бот-мережі, яке можна використовувати для перевірки вебплатформ на шахрайство.

Висновки

1. Боти можуть порушувати усі три компоненти системи кібербезпеки вебресурсів. Оскільки вони створені саме для того щоб обходити певні обмеження цих вебсайтів і успішно автентифікуватися в них, бо є програмами автоматичного доступу до закритої частини вебсайту, то найбільшу загрозу вони складають для збереження конфіденційності інформації.
2. Сканери є незамінними утилітами для роботи з тестуванням на вразливості. Але, коли мова йде про неправомірний доступ до цільового сервера за допомогою бота, боти використовують сканери на вразливості або зловмисні утиліти віддаленого доступу, які вони можуть залишати після успішної атаки. Після проходження аутентифікації бот може запустити свій вбудований сканер на вразливості, який буде шукати вразливості на

вебсайті, а потім, при знаходженні потрібних, виконати корисне навантаження і отримати неправомірний доступ до цільового сервера.

3. Використання для атаки експлойтів 0-го дня або таких шкідливих програм, що були спеціально створені під конкретну задачу чи об'єкт враження значно ускладнюють захист “жертви” атаки. Проте, оскільки кібербезпека все більше розвивається, створення таких експлойтів стає все більш затратною справою в грошовому і часовому еквіваленті. Саме тому, більшість кіберзлочинців використовуватимуть доступні уже відомі експлойти і будуть направляти свої атаки на застарівше програмне забезпечення.
4. Зловмисна система з центром управління дає набагато більше гнучкості для її розробників, дозволяє використовувати її в довгостроковій перспективі, оновлювати своє програмне забезпечення, а також вести свою діяльність довгий час без того щоби бути поміченими правоохоронними органами. Тому на даний момент основний спосіб боротьби з таким програмним забезпеченням - це активний моніторинг усіх вразливих девайсів, зміна паролів на більш захищені, створення навмисно вразливих систем-пасток (ханіпотів) для виявлення та ідентифікації злочинної діяльності.
5. Оскільки соціальні платформи поширюють інформацію серед величезної аудиторії та мають великий вплив на громадську думку майже третини населення в усьому світі, все більша частина зловмисників намагатиметься використовувати їх як зброю в інформаційних війнах. Для автоматизації взаємодії, зменшення витрат та максимального охоплення аудиторії, використовуються боти, які дозволяють автоматично взаємодіяти з акаунтами користувачів у соціальних мережах.
6. Бот – це ефективний інструмент, який робить процес просування ідей серед користувачів інтернету легким та доступним. Проте, щоб ідею успішно поширити, необхідно враховувати психологію впливу на користувача як на людину – ефективність донесення ідеї залежить від цього, як саме бот реалізує цей процес. Одночасне застосування кількох принципів

переконання людей призводить до більш вираженого впливу прийняття рішень. Основна причина, через яку ці принципи ефективні, полягає в тому, що вони служать евристиками (розумовими спрощеннями) – способом обробки інформації, на який зазвичай покладаються.

Використовуючи одночасно декілька принципів переконання (принцип авторитету, принцип соціального схвалення, принцип симпатії, принцип дефіциту), зацікавлені особи нав'язують інформацію користувачеві, спонукаючи його до дій фінансового характеру. Іноді це може дозволити іншим отримати доступ до конфіденційної інформації, яка може бути використана зловмисно.

7. Ілюзію соціального схвалення в соціальних мережах допомагають створити програми-боти, що генерують штучні позначки «мені подобається», коментарі під фото та записами на сторінці людини, що підвищують кількість передплатників та друзів певного суб’єкта. В подальшому таке соціальне схвалення монітезується з використанням різних інструментів.
8. Мережі ботів є реальними загрозами для фінансового сектору. Особливо вразливими є біржі криптовалют, банківські установи та фондові ринки. Крім прямої крадіжки фінансових ресурсів та інформації, існують репутаційні ризики, що можуть штучно створюватись мережею ботів, та напряму впливають на фінансові рейтинги, вартість бізнесу, акцій, тощо

РОЗДІЛ 2 СИСТЕМИ ТА МЕТОДИ ПРОТИДІЇ ФРОДУ ЯК СКЛАДОВА ЧАСТИНА ЗАХОДІВ КІБЕРБЕЗПЕКИ ВЕБРЕСУРСІВ

2.1. Концепція використання машинного навчання для перевірки вебресурсів на вразливості

Як було зазначено раніше, зі збільшенням числа загроз і їх еволюціонуванням, розвивається і сектор безпеки. Одним з таких рішень є використання новітніх систем на основі штучного інтелекту для виявлення і знищення шкідливого програмного забезпечення.

Для цього використовують такі алгоритми як, SVM, KNN, логічна регресія та інші. Ми розглянемо загальну концепцію, того як відбувається аналіз на загрози.

Support Vector Machine — це надійний алгоритм машинного навчання. У ньому діє як класифікація так і регресія. Класифікація підтримуючого вектора складається на основі суміжних рішень, які розділяють екземпляри з різними мітками класів в окремих групах. Також він визначає оптимальну роздільну гіперплощину. Функція ядра для SVM обчислює значення параметрів, що стосуються продуктивності детектора який визначає час і місце вторгнення, на репрезентуючому наборі даних. Комбінуючи SVM із отриманням інформації та BPSO, можна отримати два методи зменшення параметрів функцій та використати їх для створення системи виявлення на вторгнення. Запропонований метод демонструє хорошу ефективність у випадку атак «Відмова в обслуговуванні» (DoS), атаки U2R і R2L. Також деякі дослідники представили алгоритми фільтрації на основі класифікатора SVM для ідентифікації кількох класів вторгнення. Інші використовували кластеризацію нечітких C-середніх для збору неоднорідних навчальних даних в однорідній підмножині. Згодом ANN

навчаються на отриманих підмножинах в однорідних даних, а класифікатор SVM використовується для виявлення вторгнень. Класифікатор KNN вимірює подібність між двома примірниками за допомогою функції відстані.

Також автори іншого дослідження експериментують з різними типами атак і різними значеннями k за допомогою індексованого часткового пошуку відстані k-найближчого сусіда. Ще одне дослідження поєднало K-Means і kNN для створення системи виявлення вторгнень. Згодом було докладено великих зусиль, щоб запропонувати деякі нові методи для покращення продуктивності класифікації вторгнень через kNN. Крім цього був представлений інтелектуальний фільтр тривог на основі багатокласового класифікатора kNN і фільтрування несправжніх тривог. Для побудови моделей для виявлення вторгнень використовували алгоритм Ant Colony Optimization. Проте, крім KNN, також можна використовувати дерево рішень — це алгоритм машинного навчання, який широко використовується для вирішення проблем класифікації та регресії. Внутрішні вузли представляють характеристики або незалежні змінні даних, гілки представляють правила прийняття рішень і кожен листовий вузол представляє результат. Беручи до уваги привабливу структуру дерева рішень і його хорошу точність, дослідники доклали значних зусиль, щоб створити системи пошуку вторгнень (СПВ), на основі алгоритмів дерева рішень, таких як ID3, C4.5 і CART а також великий набір методів обрізання та вибору ознак.

Логістична регресія — це алгоритм, який моделює умовну ймовірність певного значення, яке вираховується через використання набору вхідних даних. Найпоширенішим її використанням є пошук бінарного рішення. Наприклад, коли відповідь виражена як правда/неправда або так/ні тощо. Мультиноміальна логістична регресія є узагальненням моделі логістичної регресії, коли існує більше двох можливих відповідей. Для виявлення загроз часто використовується також і логістична регресія. В контексті її використання, автори запропонували

алгоритм порогового значення коефіцієнта посилення (CST-GR) для вибору відповідних особливостей. Етап вибору функції проводиться на трьох типах кібератак. Зменшена кількість вибраних ознак дозволяють побудувати швидку і не важку, в плані навантаження на систему, СПВ на основі таких машин, як логістична регресія.

Глибоке навчання є окремою групою машинного навчання. Той підхід який використовується глибоким навчанням алгоритми при обробці та аналізу даних, щоб виявити цікаві закономірності та виконати задачі регресії, кластеризації або класифікації, ґрунтується в першу чергу на імітації будови людського мозку. Тобто цей алгоритм складається з серії нейронних мереж, структурований багатьма рівнями, щоб імітувати функціонування людського мозку. Алгоритми глибокого навчання так само широко використовується в кібербезпеці для боротьби з розподіленими атаками на відмову в обслуговуванні.

У підсумку хочеться підкреслити, що станом на сьогодні існують ефективні технології виявлення загроз що використовуються такими компаніями, як Google, Amazon та Microsoft. Всі вони уже досить довгий час використовують їх у своїх продуктах. Також кожний з них створили ефективні системи протидії зловмисного використання.

Проте ми сфокусуємося виключно на такій системі яка працює для соціальних мереж та фінансових організацій.

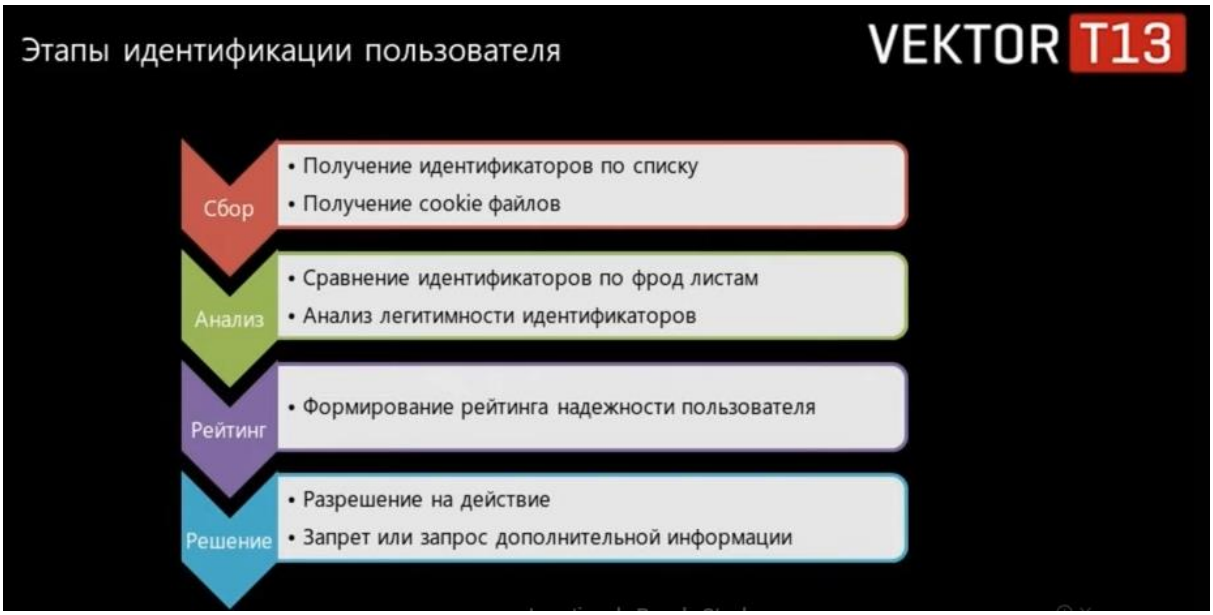
2.2. Етапи ідентифікації користувача при застосуванні системи протидії фроду

Загалом за своєю суттю системи протидії фроду це комплексні системи, які направлені на успішну ідентифікацію та унікалізацію своїх користувачів, з подальшим присвоєнням їм свого власного фрод рейтингу. Фрод рейтинг - це шкала за якою визначають вірогідність того що новий

користувацький акаунт є не справжнім і відповідно він направлений на те щоби здійснювати діяльність, яка не встановлена правилами веб-платформи та є потенційною загрозою безпеці усієї системи. Ідентифікація - це визначення користувача під час його реєстрації в системі, найчастіше через певні унікальні ідентифікатори. Унікалізація - це в свою чергу процес визначення користувача за певною інформацією, яка належить лише цьому користувачу. Способів унікалізації може бути безліч від збору інформації по IP до відбитків процесорів на андроїд девайсах і їх ідентифікаторів. Кожен з цих ідентифікаторів може бути використаний вебсервісом для поділу користувачів та виділення нецільового користувача. Нецільові користувачі це, в першу чергу, ті, хто не приносять компанії яка розробляє відповідне веб рішення жодного фінансового прибутку. Адже, якщо звернутися до “Правил і положень” будь-якого популярного веб сервісу, особливо таких що відносяться до соціальних мереж, там буде написано, що якщо ви приймаєте дані правила і положення, тоді ваші дані можуть бути використані для представлення персоналізованої реклами. (57) Таким чином, розробники соціальної мережі монетизують свій бізнес.

В даному розділі ми сфокусуємося саме на антифрод системах, що працюють з користувачами персональних комп'ютерів: на операційних системах для персональних комп'ютерів з використанням такого програмного забезпечення як веб браузер. Об'єктом нашої уваги буде операційна система Windows 10 (власність компанії Microsoft) і браузери на основі ядра Chromium (виробництва компанії Google, Google Chrome та Microsoft Edge). Предметно буде розглянуто, як в працюють більшість фрод систем, на основі веб браузера і яку інформацію про систему вони здатні зібрати.

Отже, сама робота антифрод системи складається з наступних кроків, що представлені на малюнку.



Малюнок 2.1. Етапи ідентифікації користувача. (взятий з Youtube каналу VektorT13 - українського розробника рішень для обходу систем ідентифікації).

Робота системи починається зі збору інформації, її подальшого аналізу, формування рейтингу користувача і згодом прийняття рішення про надання дозволу на використання, запит додаткової інформації або заборона на вчинення будь-яких дій у системі. Розглянемо поетапно те яким чином проводиться збір інформації та як аналізується зібрана інформація.

2.3. Рівні ідентифікаторів користувача

Отже, яку інформацію здатен збирати про користувачів веб браузер та відповідно веб сервіс який цей браузер використовує.

Інформація яка доступна браузера може бути представлена наступною діаграмою:

Малюнок 2.2. Рівні Ідентифікаторів (13).

Тобто в загальному для збору інформації цільовому веб-сервісу доступні наступні рівні ідентифікаторів [13]:

1. Мережевий рівень

a. Рівень IP адрес і протоколів

2. Браузерний рівень

a. Рівень ідентифікаторів які генеруються під час використання браузерних технологій

2. Рівень операційної системи або програмний рівень

a. Ідентифікатори згенеровані операційною системою

b. Ідентифікатори додаткового програмного забезпечення, що встановлене на персональний комп’ютер

2. Апаратний рівень

a. Рівень використання обладнання на персональному комп’ютері

Почнемо з найнижчого апаратного рівня ідентифікації користувачів.

а) Апаратний рівень. За допомогою цього рівня, веббраузер здатний отримувати основну інформацію про систему в якій він працює. Основні ідентифікатори, які можуть бути цікаві вебсервісам наступні:

1. Розмір монітору

a. Його роздільна здатність

2. Інформація про процессор

a. Кількість ядер

b. Марка процессора

2. Інформація про відеокарту

a. Яка марка відеокарти

2. Оперативна пам’ять

a. Кількість пам’яті

Чому саме ці ідентифікатори найбільш цікаві для вебсервісів? Тому що вони дозволяють визначити аномального користувача, тобто їх нецільового користувача, який не принесе заробітку для компанії. Майже кожен з цих ідентифікаторів, може показати наступні речі: чи використовує користувач звичайний персональний комп’ютер чи він використовує систему віртуалізації або серверні технології. Наприклад, антифрод система зможе зрозуміти, що система яку використовує користувач на даний момент аномальна, проаналізувавши кількість оперативної пам’яті та кількість ядер процесора. Якщо припустити що оперативної пам’яті в користувача 128 ГБ, а кількість ядер перевищує 32 (58), тоді обидва параметри занадто великі для персональної машини, а отже фрод рейтинг даного користувача суттєво знизиться і скоріше за все він не зможе виконувати цільові дії на сайті або він буде змушений пройти додаткові перевірки.

Так само антифрод система працює з монітором та відеокартою. Встановлено що стандартна роздільна здатність монітору становить 16/9 (59). Відповідно більшість користувачів будуть мати саме таку роздільну здатність свого монітору під час використання веб ресурсу. Якщо роздільна здатність складає 3/4, що є стандартною конфігурацією для віртуальної машини (60) тоді це також буде сприйняте антифрод системою, як дещо підозріле і, відповідно, загальний фрод рейтинг погіршиться [14].

Щодо відеокарти то в даному випадку веб сайт здатний через отриману марку відеокарти отримати інформацію про те чи дійсно у користувача справжня відеокарта чи він використовує програми віртуалізації.

б) Рівень операційної системи. На цьому рівні вебсервіси можуть отримувати про систему наступну інформацію [15]:

1. Тип та версія операційної системи

a. Інформація з заголовку

i. Ідентифікатори браузера, які надаються при його інсталяції

b. Інформація з канвасу

i. Браузерна технологія для відображення малюнків

b. Інформація з шрифтів

i. Деякі шрифти присутні виключно в певній операційній системі

2. Ім’я комп’ютера

a. Деякі веб ресурси, такі як Google, можуть отримувати ім’я комп'ютера і зберігати його для подальшого аналізу

2. Геолокація

a. Отримання геолокації через IP адресу

b. Отримання геолокації через визначення найближчих Wi-Fi мереж (where.am.I)

i. Не використовується без дозволу користувача

b. Windows Server Location

i. використовує геолокацію з локальної операційної системи

2. Мови системи

a. Веб-сервіс інформацію відповідно до мови вибраної в системі, що визначена як основна

2. Відкриті порти

a. Які програми працюють на системі користувача

2. Часовий пояс

a. Отримує часовий пояс для того, щоби передавати інформацію про час за яким можна звертатися до користувача або який час відображати у додатку

2. Локальна IP

a. Яка використовується інфраструктура в користувача

Кожні з цих параметрів можуть бути використані для виявлення використання засобів приховування свого справжньої особи. Розглянемо кожен пункт по черзі.

Тип та інформація про операційну систему може бути отримана 3 різними способами: через браузерну інформацію (відбиток браузера або інша назва UserAgent), через технологію Canvas або через шрифти операційної системи. Canvas - це технологія яка була включена в Html5, для рендерингу зображень не на стороні сервера, а на стороні клієнта (16). Ця технологія пришвидшує завантаження зображень і анімації в браузері користувача. Однак кожен браузер має свою систему створення графічних зображень та, відповідно, має свій ідентифікатор. Цей ідентифікатор не є унікальним проте, маючи доступ до статистичної інформації, тобто до інформації наскільки часто такий ідентифікатор зустрічається на інших системах, можна зрозуміти з яким саме браузером працює користувач. Отримання інформації про операційну систему служить в першу чергу, для надання кращих послуг користувачам. Наприклад, при завантаженні програмного забезпечення вебсайт сам запропонує користувачу скачати програму

для своєї операційної системи. Однак, ця технологія також може бути використана для виявлення фроду.

Наведемо наступний приклад: припустимо, що відбиток браузеру відображає систему Windows 10 та браузер Google Chrome, а ідентифікатор канвасу повідомляє, що цей ідентифікатор зустрічається найчастіше в Mozilla Firefox де операційна система Linux, а шрифти підходять для шрифтів Linux. Відповідно, проаналізувавши всі ці дані, антифрод-система може побачити невідповідності, та зрозуміє, що має місце спроба приховування особи і на підставі цього аналізу приймає рішення про зниження рейтингу користувача.

Ім’я комп'ютера використовується не так часто, більш актуальне для встановлення кінцевого бенефіціара. Корпорація Google має технологію, що використовується у випадку створення багатьох облікових засобів з одним і тим же ім’ям користувача. Інформацію про алгоритми, що застосовуються при її реалізації корпорація не розголошує [15].

Геолокація використовується в першу чергу для надання кращих послуг користувачам в регіоні їх перебування. Для встановлення більш точної геолокації користувача може також використовуватися технологія отримання доступу до найближчих точок WI-FI (17), що має назву WPS - wi-fi position service. Вона дозволяє геолокувати користувача з похибкою до одного метра використовуючи всі наявні мережі до яких WI-FI адаптер локального комп’ютера відправляє запити. Незважаючи на свою першочергову мету, технологія геолокація також може бути використана при виявленні фроду. Наприклад, у випадку якщо геолокація не збігається з даними переданими користувачем при реєстрації або якщо користувач ніколи не заходив в вебдодаток через цю IP адресу, з цієї геолокації - може спрацювати фрод система.

Мови та часовий пояс використовуються веб-сервісами для того щоби правильно вибирати мову і показувати сповіщення відповідно до часового поясу користувача. В даному випадку може використовуватися як перевірка чи справді користувач знаходиться в тому ж самому часовому поясі що і його геолокація. Якщо присутнє не співпадіння тоді, скоріше за все, користувач використовує системи приховування особистості. Щодо використаної мови то цей параметр може бути справді використаний, для запуску антифрод діяльності, в тому випадку якщо користувач пробує зробити імперсонацію свого акаунту, як людини яка проживає в іншій країні.

Відкриті порти - ця інформація може бути використана при підключенні до стрімінгу для веб-ресурсів або для роботи з якимись специфічними задачами. В основному, це перевірка може показати що на публічній айпі адресі користувача відрито багато портів і, відповідно, запущено багато програм, які звичайними користувачами дуже часто не використовуються.

Внутрішня айпі адреса може бути використана, щоби визначити побудову внутрішньої мережі користувача. Якщо ця мережа нагадує серверну, тоді також може бути запущена антифрод система.

в) *Браузерний рівень.* На цьому рівні збирається інформація яка належить саме браузеру. Тут можна виділити декілька основних параметрів які може використовувати антифрод система для визначення неправомірного використання. Насамперед - це властивості браузера, його функціонал, а також відбитки браузера [18].

Відбитки браузера ми також аналізували коли говорили про програмний рівень ідентифікації. Найважливішими при браузерній ідентифікації, крім відбитку канвасу та UserAgent, є також аналіз таких технологій як WebRTC, AudioContext та WebGl. Розглянемо кожну з них по порядку.

WebRTC - це технологія для роботи зі стрімінгом відеодзвінків між двома різними девайсами прямо в браузері. Окрім того що ця технологія дозволяє робити відеодзвінки, вона також дає можливість отримати справжню IP адресу користувача в тому випадку якщо він її приховує через методи проксифікації свого мережевого трафіку. Проксифікація - це загортання веб запитів з однієї машини в запити іншої через протокол http або socks, для зміни оригінальної адреси яку використовує користувач. (61) Технологія WebRTC побудована таким чином, щоби створювати з'єднання між двома девайсами напряму без залежності від 3-ї сторони. Однак для того щоби ця технологія працювала обом девайсам потрібно знати IP адреси один одного. Саме тому, незважаючи на спроби маскування свого мережевого трафіку, браузер все одно зможе отримувати справжню IP адресу користувача.

AudioContext - технологія для роботи з аудіофайлами. Також може бути використана як ідентифікація віртуальної машини або віддаленого сервера, через те що у віртуальних машинах дуже часто немає можливості програвати аудіо. Це призводить до того, що через цю технологію можна отримати інформацію про те що аудіо не може програватися на даній машині і, відповідно, ця машине не є користувацькою.

WebGl - технологія для роботи з 3D графікою для рендерингу онлайн ігор. Для ідентифікації користувачів з високим рівнем фроду може використовуватися запит на рендеринг якогось зображення, яке може бути перетворене в ідентифікатори системи WebGl, які використовує цей браузер.

Тепер коли ми розглянули ці три основні технології, ми можемо перейти до властивостей браузера які можуть бути використані антифрод системами, як спосіб ідентифікації користувачів.

Основними з цих властивостей є: загальна потужність браузера, апаратне прискорення графіки та потужність рендерингу зображень.

Апаратне прискорення графіки - це здатність використовувати потужність обладнання для генерації зображення. В даному випадку ця технологія теж може використовуватися для ідентифікації аномальної активності, так як більшість віртуальних машин не підтримують обладневе прискорення графіки, адже на віртуальних машинах просто немає доступу до відеокарти. І тому вони використовують програмне прискорення графіки, а в основних ідентифікаторах у WebGl або у Canvas для віртуальних машин буде відображатися однакова інформація, та яка найчастіше використовується у віртуальних машинах.

Потужність рендерингу - це здатність системи рендерити зображення у високій якості і досить швидко. Так як серверні системи, направлені на виконання в основному технічних завдань вони не створені для того щоби на них рендерили графіку. Саме тому при проведенні тесту на рендеринг такі системи їх не пройдуть і будуть показувати мінімальні значення.

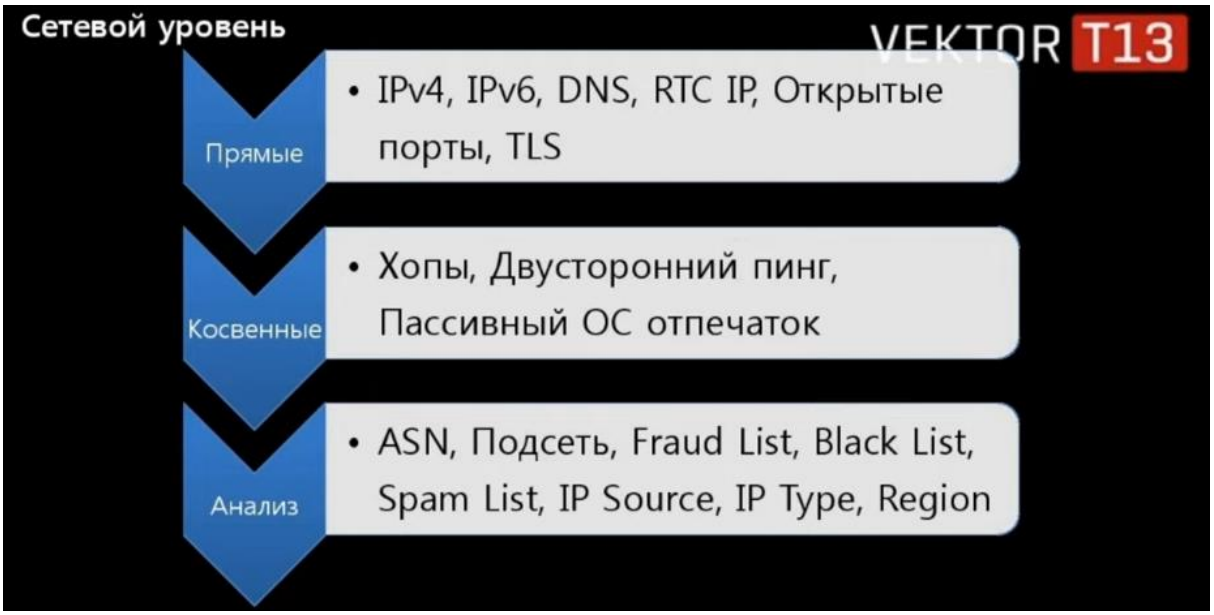
Потужність браузера - це швидкість завантаження контенту і виконання задач у браузері. Тут, на відміну від інших параметрів, чим більш посередня можливість роботи тим краще. Тому що, якщо браузер виконує дуже швидко всі операції це означає що процесор, який виконує ці задачі, потужніший за той який є у звичайних користувачів. Це можна бути ознакою використання віддаленого серверу, що негативно впливає на загальний фрод рейтинг.

До функцій браузера ми можемо віднести підтримку CSS та підтримка API підключення до USB приладів. Однією з можливостей браузерної ідентифікації це визначення, які саме CSS технології він може підтримувати. В залежності від версії браузерів, в них може бути відсутня підтримка деяких технологій. Відповідно, якщо користувач підміняє свої ідентифікатори таким чином, що в

нього зображений інший браузер, тоді веб-платформа дуже просто здатна виявити підміну, перевіривши підтримку тих API, які мають бути доступні для цільового браузера. У випадку не співпадінь фрод рейтинг може знижуватися. У випадку роботи з API USB тут потрібно згадати проект Fugu запущений компанією Google в 2016 році.(62) Ця технологія направлена на те щоби ще більше з'єднати браузер з фізичними елементами системи. Наприклад, отримувати доступ до системи контролю підсвітки клавіатури чи управління вібрацією ігрового джойстику. Проте крім цього ця технологія здатна визначати, які пристрої на даний момент підключені до системи, а також отримувати детальну інформацію про них. У випадку, якщо до системи не буде підключено фізичної клавіатури або комп'ютерної мишки, це може знову ж таки призвести до зниження фрод рейтингу.

г) *Мережевий рівень.* Останній рівень в цій моделі роботи антифрод систем. Цей рівень направлений на збір інформації про зовнішнє мережеве оточення цільового користувача з подальшим його аналізом через свої бази даних фрод активності або відношення до серверної, а не домашньої інфраструктури.

Узагальнена кількість інформації, яка може бути зібрана будь-якою антифрод системою показана на наступній ілюстрації [19].



Малюнок 2.3. Список інформації, яку може збирати антифрод система

Прямі ідентифікатори - цю інформацію система може отримати напряму проаналізувавши мережеве оточення системи користувача. Непрямі ідентифікатори - це така інформація, яку зможе отримати антифрод система, при прямій взаємодії з пристроєм користувача. Наприклад, при обмінами запитів з користувацькою системою.

Аналітичні ідентифікатори - ці дані зібрані з використанням аналітичної інформації про IP адресу. Наприклад, при використанні Fraud List бази даних, яка вміщує в собі інформацію про те чи була ця IP адреса використана для фродулентної активності.

Розглянемо найважливіші з прямих ідентифікаторів. Це Ipv4, Ipv6 та DNS.

Ipv4 - це інтернет протокол четвертого покоління. Загальна величина доступних в цьому адресному просторі Ip адрес складає 32-біта або 2^32. Так, як цей адресний простір обмежений цим числом на його зміну був запропонований адресний простір Ipv6. Отримавши цей ідентифікатор антифрод система може використати його для подальшого аналізу.

Ipv6 - інтернет протокол шостого покоління в собі він розміщує 128-бітний адресний простір він може включати в себе 2^128 степені різних варіацій Ip адрес (20). Цей новий протокол, через деякий час повністю замінить собою протокол 4-го покоління. Особливість цього протоколу також в тому, що при його передачі в ньому можна закодувати MAC - адресу користувацької мережевої карти. MAC-адреса - це унікальний ідентифікатор, який назначається всім мережевим платам. Відповідно при аналізі Ipv6 можна отримувати додаткові дані про мережеву карту, що сприятиме додатковій ідентифікації користувача.

DNS (Сервер доменних імен) - ця технологія використовується для пов'язування нечитабельної для людей адреси інтернет протоколів до читабельних доменних імен. Сервери доменних імен залежать від місця розташування кожного користувача, адже ця технологія в основному побудована в ієрархічному стилі. На найнижчому рівні є регіональний ДНС сервер до якого звертається по запит на знаходження айпі адреси користувач, якщо у нього немає інформації по певному доменному імені, тоді він звертається до доменного сервера вищої ланки ітд. Через цю особливість, може бути таке що при підміні своєї айпі адреси користувач залишив незмінним свій ДНС сервер. Потрапляючи до антифрод системи ця інформація може бути використана, для ідентифікації зміни своєї особистості людиною.

Після збору інформації, антифрод система проводить аналіз цієї інформації використовуючи свої або орендовані бази перевірки даних по айпі адресах. Перевіряється на використання спаму, фроду чи належиться ця IP адреса до серверних систем, чи є вона проксі сервером, або впн сервером. На основі зібраної інформації, можуть робитися відповідні рішення по підвищенню чи пониженню фрод рейтингу користувача.

Тепер проаналізувавши більшість найбільш важливих ідентифікаторів, які збирають антифрод системи, ми можемо зрозуміти, яким чином нам потрібно

будувати свою систему для подальшого їх обходу. В наступному розділі ми присвятимо час тому яким чином генерується інформація для створення нових акаунтів в соціальних мережах. Це робиться з метою зрозуміти, яким чином створюються ці акаунти і як працюють системи автогенерації акаунтів. Взагалі антифрод системи не потрібно аналізувати, лише за одними аргументами. Кожний з аргументів, які отримуються фрод системою аналізуються окремо і формують загальний фрод рейтинг користувача. На основі якого і приймається згодом рішення про використання санкцій до цього користувача, його додаткової перевірки або дозволу на роботу в мережі. Наступним етапом, за статичним аналізом йде також і поведінковий аналіз. Соціальні мережі, а також фінансові організації, такі як банки, наприклад, розробляють свої власні системи поведінкового аналізу, які вирізняються досить високою здатністю до ідентифікації фроду.

Висновок

1. Використання новітніх систем на основі штучного інтелекту для виявлення і знищення шкідливого програмного забезпечення є одним з найперспективніших рішень у сфері безпеки та використовується для виявлення та знищення шкідливого програмного забезпечення. Комбіннуючи SVM із отриманням інформації та BPSO, можна отримати два методи зменшення параметрів функцій та використати їх для створення системи виявлення на вторгнення, особливо у випадку атак «Відмова в обслуговуванні» (DoS), атаки U2R і R2L. Перспектиними є використання алгоритмів фільтрації на основі класифікатора SVM для ідентифікації кількох класів вторгнення. Вартим уваги є дослідження метою якого є створення системи виявлення вторгнень та запропоновано нові методи для покращення продуктивності класифікації вторгнень через kNN. Крім того для створення системи пошуку вторгнень досить часто використовують

алгоритм машинного навчання – дерево рішень, зважаючи на його структуру та точність. Для виявлення загроз використовується також і логістична регресія.

2. Глибоке навчання є окремою групою машинного навчання. Алгоритми глибокого навчання так само широко використовується в кібербезпеці для боротьби з розподіленими атаками на відмову в обслуговуванні.
3. Загалом за своєю суттю системи протидії фроду це комплексні системи, які направлені на успішну ідентифікацію та унікалізацію своїх користувачів, з подальшим присвоєнням їм свого власного фрод рейтингу.
4. Головні етапи ідентифікації користувача при застосуванні системи протидії фроду - збір інформації про користувача, її подальшого аналізу, формування рейтингу користувача і згодом прийняття рішення про надання дозволу на використання, запит додаткової інформації або заборона на вчинення будь-яких дій у системі.
5. Аналіз рівнів ідентифікаторів користувача, які збирають антифрод системи (мережевий рівень, браузерний рівень, рівень операційної системи та апаратний рівень) дозволяють зрозуміти, яким чином нам потрібно будувати свою систему для подальшого їх обходу.
6. Антифрод системи не потрібно аналізувати, лише за одними аргументами. Кожний з аргументів, які отримуються фрод системою аналізуються окремо і формують загальний фрод рейтинг користувача. На основі якого і приймається згодом рішення про використання санкцій до цього користувача, його додаткової перевірки або дозволу на роботу в мережі. Наступним етапом, за статичним аналізом йде також і поведінковий аналіз. Соціальні мережі, а також фінансові організації, такі як банки, наприклад, розробляють свої власні системи поведінкового аналізу, які вирізняються досить високою здатністю до ідентифікації фроду.

РОЗДІЛ 3 МЕТОДОЛОГІЯ СТВОРЕННЯ ОБЛІКОВИХ ЗАПИСІВ З ВИКОРИСТАННЯМ МОДЕЛЕЙ МАШИННОГО НАВЧАННЯ

Тепер коли ми розглянули усі особливості роботи антифрод системи ми можемо розглянути те яким чином відбувається створення цілісної системи, яка створена для того, щоби оминати системи захисту. Також ми на практиці побудуємо одну з її частин.

3.1. Генеративна частина мережі ботів

Проте зараз почнемо ми зі створення облікових записів. Створення облікових записів для професійних користувачів соціальних мереж складається з декількох етапів:

1. Створення антидетект системи у себе на персональному комп'ютері

a. Цей процес включає застосування всієї наявної інформації про антифрод систему конкретних соціальних мереж і подальшим налаштуванням своєї системи таким чином, щоби обійти ці обмеження

2. Використання коректних телефону або пошти

a. Для реєстрації в соціальних мережах найчастіше треба саме ця інформація

b. Як телефон так і пошта мають бути відносно довгий час у використанні, а також не знаходитися в базах даних фродулентної активності.

2. Реєстрація акаунту

3. Наповнення даними

a. Використання зображення, яке не належить справжнім користувачам мережі. Адже зображення можуть бути проаналізованими мережею та можуть бути знайдені співпадіння, що викличе спрацьовування антифрод системи.

b. Створення правдоподібної біографії

i. Біографія має співпадати з регіоном де "проживає" акаунт

b. Наповнення акаунту фотографіями і правдоподібними даними

c. Активізація діяльності акаунту через підписки на лідерів думок або на спільноти, лайки та репости контенту. Спільноти - це групи людей, які об’єдналися в соціальній мережі за своїми інтересами.

i. Ця робота направлена на те щоби обходити поведінковий аналіз антифрод систем. Адже якщо профіль не буде виконувати жодних дій антифрод система також може запуститися

ii. Для створення таких акаунтів також важливо зразу не бути підписаним на спільноти, які близько пов’язані з діяльністю, яку планують вести ці акаунти. Тому що таким чином можна досить швидко відслідкувати групу пов’язаних між собою акаунтів за їх вподобаннями.

Усі зібрані тут рекомендації були отримані шляхом експериментів з роботою над соціальними мережами. Згодом в експериментальній частині, ми розглянемо які саме дані та які системи використовувалися. На даний момент, ми сфокусуємося на тому, яким чином можна зібрати готовий акаунт користуючись сучасними методами.

3.1.1. Використання генеративно-змагальної нейронної мережі для генерації зображень користувачів

Генеративно-змагальна нейронна мережа (GAN - generative adversarial network) - це система машинного навчання направлена на генерацію зображення за допомогою двох систем машинного навчання, які конкурують між собою і таким чином здатні знаходити, копіювати і змінювати варіації всередині датасету. Через свою новаторську структуру цей алгоритм машинного навчання здатний справлятися з генерацією абсолютно нових патернів на зображеннях використовуючи наявні дані. Таке новаторство призвело до того що тепер за допомогою GAN мережі можна генерувати будь-які зображення людей, як тих що уже існують, так і таких, що ніколи не існували.

Однак крім того, що ці мережі можуть генерувати красиві і правдоподібні зображення вони також можуть генерувати обличчя і фотографії начебто справжніх людей для обходу системи захисту від фроду. Для цього використовуються, як свої рішення так і рішення написані власноруч.

На наступній картинці я продемонструю обличчя згенероване за допомогою цієї технології і взяте з сайту https://thispersondoesnotexist.com/.

Мал. 3.1. Демонстративна картина, яка зображує неіснуючу людину. Зображення згенеровне за допомогою StyleGan. Джерело https://thispersondoesnotexist.com/.

Причому ця технологія абсолютно відкрита будь-хто може перейти на джерело на сайті Github цього веб-ресурсу, отримати доступ до коду та на його основі створити своє рішення [21].

На даний момент було створено багато різних моделей машинного навчання, які займаються визначенням таких зображень. Проте мало з них є дійсно ефективними в боротьбі проти створення подібних зображень (63). Однак інколи навіть людське око може відрізнити зображення створене штучним інтелектом і справжнє. Ось приклад такого зображення:

Малюнок 3.2. Зображення дівчинки, створене штучним інтелектом.

Крім очевидного нереального фону на цій картинці дуже добре видно що зіниці у цієї дитини неправильної форми, а отже і саме зображення не реальне.

Тобто можна зробити наступний висновок - незважаючи на те, що ця технологія досить добре справляється з генерацією зображень, вона поки що далека до створення ідеального зображення. Проте прогрес йде вперед і деякі компанії уже зараз надають своє програмне забезпечення в аренду для створення таких нереальних персонажів (64). Це може призвести до того що заповнення профілів соціальних мереж може бути дуже простою і швидкою справою. Зловмисникам буде варто лише створити відповідну систему для створення зображень і вони в режимі реального часу будуть отримувати всі необхідні для себе матеріали.

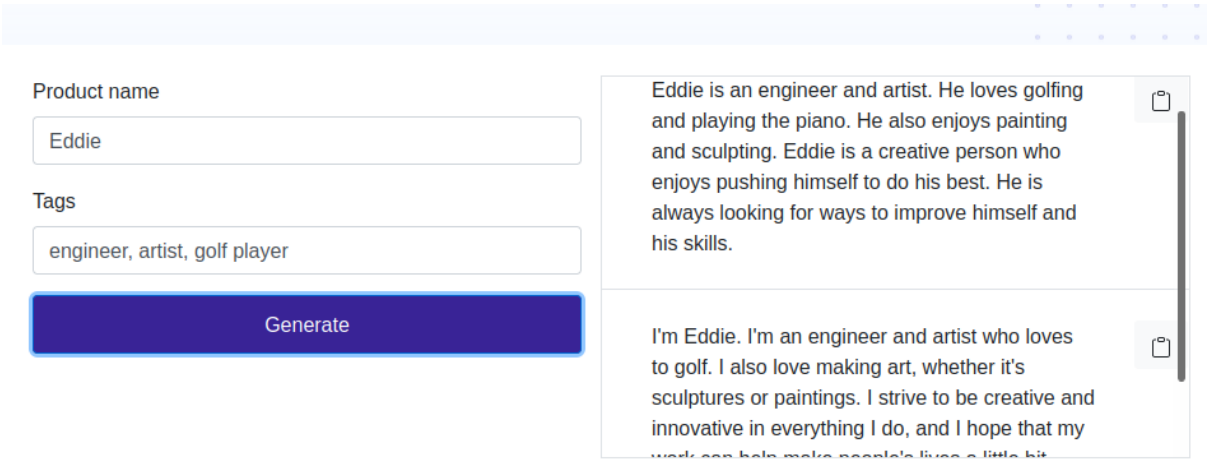
Наступним етапом роботи над профілем буде створення короткої біографії для нового акаунту

3.1.2. Створення біографії користувача за допомогою рекурентної нейронної мережі

В створенні нової біографії нам допоможе така галузь машинного навчання як natural language processing. Ця технологія дозволяє алгоритму розуміти введені слова, так само як би це зробила людина. Завдяки цьому він може створювати правдоподібні тексти у відповідь на введену інформацію. Для цієї роботи була створена архітектура нейромереж під назвою RNN або рекурентні нейронні мережі. Цей алгоритм був розроблений для вирішення задач зі створення послідовного виводу з такого послідовного вводу. Тобто по суті мати змогу генерувати текст. Звичайні нейромережі не здатні такого робити, тому що вони працюють тільки з теперішнім вводом даних і не запам’ятовують попередньо введений стан. Саме тому вони не можуть бути використані для

створення тексту [22]. Тоді як RNN запам’ятовує кожен введений ввід і на основі його створює новий.

Тепер коли ми розібралися з тим як працює генерація тексту ми маємо можливість перевірити, як вона працює в реальному житті. Зображення взяте з веб-сайту https://www.neuraltext.com/ai/personal-bio-generator. Було задіяно декілька спроб генерації тексту і створено декілька правдоподібних біографій.



Малюнок 3.3. Згенерований текст з веб-сайту https://www.neuraltext.com/ai/personal-bio-generator.

В даному випадку було використано лише ім’я “Eddie” та додаткові параметри, які нейромережа використає при створенні нового тексту. Вийшли короткі правдоподібні тексти англійською мовою. Під правдоподібністю тут мається на увазі те що це послідовна розповідь, яка має зміст. Таку модель так само можна натренувати для інших мов і для генерації більших текстів. Крім цього можна використати введену інформацію про персонажа і отримати повністю згенерований для нього профіль.

Ось приклад з ще одного сайту який спеціалізується на генерації драматичних сюжетів біографії. Ввівши лише інформацію про ім'я, стать, вік, соціальний статус, освіту і позитивні або негативні типи характеру - можна отримати повний профіль згенерованих даних про цю вигадану особистість. (малюнок 3.4.).

Згенерована особистість в повній мірі може бути використана, як справжня цією інформацією можна заповнити профіль в соціальній мережі. Ще окрім цього зловмисники, можуть зберігати за потреби цю інформацію щоби на її основі генерувати тексти, які буде писати їх бот. Багато ботів, в тому числі і російських, реагують на певні ключові слова через, які вони активуються і починають писати свої згенеровані тексти.

Basic Information

Name:	Ilon Musk Wright
Nickname:	Lazy Ilon
Reason for nickname:	Descriptive
Date of birth:	Wednesday, 7th Jan 1998 (Age 24)
Star sign:	Capricorn
Nationality:	British
Ethnicity:	Caucasian
Social class:	Upper class
Religion:	Agnostic
Sexuality:	Straight
Education:	Masters / doctorate
Course:	Chemistry
Political views:	Apathetic
Relationship status:	In a relationship with Spenser Raleigh Humphreys
Career path:	Artist

Physical characteristics

Height:	very short
Shape:	overweight
Build:	very fine build
Hair colour:	ginger
Eyes:	blue

Малюнок 3.4. Повністю розписаний персонаж вигаданої особистості

Отже, проаналізувавши декілька варіантів, як можна використовувати генерацію тексту ми можемо зробити висновок, що теперішні системи для управління ботами можуть бути дуже добре розвинуті та діяти майже повністю автоматично. Якщо моделювати ідеальну систему для роботи з ботами, тоді найімовірніше це буде виглядати наступним чином:

1. Генерація профілю для цільової групи людей
   a. Цільова група людей - це різні за віком і статтю представники населення в певному цільовому регіоні
   b. Створення профілю з нуля за випадковими даними. Головне щоби він співпадав з цільовою групою людей.
2. На основі заготованого профілю генерується зображення нової особистості
   a. Декілька профільних зображень
2. З інформації про ім’я та вік особи створюються електронна пошта
   a. На такій системі має стояти функції антидетекту, адже звичайна пошта не зможе бути створена без них, адже в більшості поштових провайдерів налаштовані системи боротьби з фродом.
   b. Або інший варіант використовуються сервери сіп телефонії з назначенням нової сім карти кожному користувачу.
2. Відбувається реєстрація користувача
3. Профіль автоматично заповнюється даними
4. Акаунт починає підписуватися на інші акаунти і створювати свою активність в мережі.

Як ми можемо побачити це запропонована мною спрощена модель роботи комплексної системи зі створення автоматичних акаунтів. Ця модель спрощена,

тому що в ній не розписане налаштування системи для роботи антидетекту, а також немає інформації з приводу того якими чином буде відбуватися активність на веб-ресурсі. Однак разом ця компонента системи називається “Генеративною”. В неї така назва тому що вона генерує нову інформацію про особистість. Далі я буду говорити про цю частини саме в такій інтерпретації.

Отже розглянувши повністю алгоритми генерації зображень та тексту, ми маємо перейти до останньої проте найважливішої частини - це поведінкова модель. Саме вона відповідатиме за тривалість використання акаунту і за допомогою неї можна буде приховувати негативну діяльність від будь-яких антифрод систем.

3.1.3. Побудова поведінкової моделі персонажа

Після проведення дослідження я визначив, що поведінкова модель може бути побудована двома способами. Це RNN - модель де замість тексту ми генеруємо новий список завдань, які має виконати бот. Або це Reinforcement learning модель. На мою думку в рамках дослідження варто розглянути обидва варіанти.

РНМ з генерації не лише тексту використовувалася і раніше (65). В даному випадку можна побудувати модель на основі зібраної інформації від звичайних користувачів. Такі дії вже зараз виконуються шкідливим програмним забезпеченням для збору поведінкової статистики з користувачів мережі інтернет для кращого маскування своєї злочинної діяльності (66). Так само можуть робити і автоматизовані боти. Для цього вони можуть або найняти людей на таку роботу або зібрати дані через встановлення шкідливого програмного забезпечення на персональні комп’ютери цілей без їх відома. Перший варіант безперечно найбільш простий і дешевий у використанні. Зібравши достатню кількість

інформації від різних груп людей можна побудувати на їх основі нову генеративну модель, яка буде працювати над створенням поведінки бота.

Reinforcement learning - це технологія машинного навчання, яка знаходиться на стику наглядових і не наглядових систем машинного навчання. В основі її лежить так робота з підкріпленням позитивних результатів моделлю у випадку її позитивної дії і негативним покаранням у випадку дії яка була виконана погано. Також програмі дається можливість виконувати певні задачі чітко закладені в алгоритмі. Наприклад, потрібно правильно припаркувати машину (67). Алгоритм працює з тими обмеженнями, які в нього є намагаючись виконати задачу якомога краще. Ця технологія досить складна в розробці і використанні, адже для тренування успішної дії для моделі їй потрібно зробити дуже велику кількість повторів. В нашому випадку головна ціль моделі згенерувати подібний до людського набір дій які зможуть пройти наші параметри ідентифікації автоматизованих акаунтів. Відповідно нападаючий, приблизно знаючи як працює антифрод система цільового веб-ресурсу, може налаштувати такі обмеження експерименту, що в кінцевому результаті наша модель зможе прийти до правильного висновку про те як обійти захист. Переглянувши, як працює навчання з підкріпленням в інших дослідженнях я зробив висновок, що це дуже дорогий та можливо неточний метод на який не варто покладатися (68). Можливо в майбутньому тренування моделі таким чином буде виправдана, проте, на мою думку, це однозначно не сьогодні.

Отже, тут як ми бачимо у нас вимальовується суцільна частина однієї системи під назвою Генеративна. Вона складається з моделі для створення зображень, тексту та поведінки. Тепер коли ми розглянули усі ці частини нашої системи ми можемо перейти до другої її частини - аналітичної.

3.2. Аналітична частина мережі ботів

Аналітична частина мережі ботів відповідає за те, як кожна програма бот збирає інформацію та те, що відбувається на сторінці веб браузера. Аналізуючи доступну інформацію, я прийшов до висновку, що бот, створений на основі програми для персональних комп'ютерів, може аналізувати інформацію декількома наявними засобами. Це аналіз html сторінки та штучний зір. Зараз ми розглянемо детально обидві варіації і як вони працюють.

Аналіз html сторінки можна зробити, отримавши її при роботі зі своїми проксі серверами або при роботі з браузером в автоматичному режимі. При автоматизації роботи браузера відбувається його запуск в тестовому режимі і за допомогою фреймворків для тестування роботи веб сторінок такі як Puppeteer [25] можна звертатися до завантаженої веб сторінки прямо під час роботи браузера. Ця технологія набагато легша у використанні, однак браузер може знати про те що з ним відбувається така робота. Відповідно він здатний передати цю інформацію своїй антифрод системі, що може застосувати санкції проти користувача. Я також тестував цей варіант реалізації, проте вирішив в рамках дослідження провести більш складний експеримент про який я напишу трохи згодом.

Для тестування роботи проксі сервера, я використовував Titanium Web Proxy, написаний на мові програмування C# [23]. За допомогою нескладних команд я зміг запустити програму проксі сервер, яка захоплювала увесь https трафік, розшифровувала його у мене на машині і передавала його далі на обробку на сервер керування. В підсумку мені вдалося перехоплювати лише https запити які містили html файли, адже інші файли такі, як картинки чи стилі, при статичному аналізі сторінки не потрібні.

Далі я продемонструю частину коду яку я використав для перехоплення цих запитів. Важливо також зазначити, що програма сама встановлювала свій кореневий TLS сертифікат для розшифрування веб запитів.

```
61      private async Task WebProxy_BeforeResponse(object sender, SessionEventArgs e)
62      {
63          try
64          {
65              if (e.HttpClient.Request.Method == "GET")
66              {
67                  var response = e.HttpClient.Response;
68                  if (response is { ContentType: { } } && response.ContentType.Contains("text/html") && response.HasBody)
69                  {
70                      string body = await e.GetResponseBodyAsString();
71                      try
72                      {
73  #if test
74  ...
83  #else
84                          await NetworkService.PostObject(CreateObjectToSend(e, body));
85  #endif
86                      }
87                      catch (Exception ex)
88                      {
89                          Logger.LogError(ex, "Error occured while intercepting request");
90                      }
91                  }
92              }
93          }
94          catch (Exception ex)
95          {
96              Logger.LogError(ex, "Error occured while intercepting request");
97          }
98      }
99
100     1 reference | 0 changes | 0 authors, 0 changes
101     private ObjectTextToHtml CreateObjectToSend(SessionEventArgs e, string body)
102     {
103         var request = e.HttpClient.Request;
104         return new ObjectTextToHtml { Url = request.Host, Text = body };
105     }
```

Малюнок 3.5. Практичне застосування проксі сервісу для аналізу html сторінки. Запит перехоплюється якраз перед відповіддю і перенаправляється на центральний сервер.

На самому центральному сервері ми спершу аналізуємо сторінку за допомогою AngleSharp бібліотеки для роботою з html документами. Тут я виділяю усі основні необхідні для роботи елементи. А саме: кнопки, поля для вводу (наприклад, поля для заповнення інформації та ін.), текстові поля, заголовки, картинки і посилання. Кожен з цих елементів має відповідну назву і буде використовуватися для створення активності по різному. Наприклад, на всі елементи від картинки до посилання можна натискати, а в поля можна вносити текст. Ця інформація збирається у базу даних і зберігається у ній для подальшого аналізу. Також при повторному запиті на цю сторінку система буде автоматично

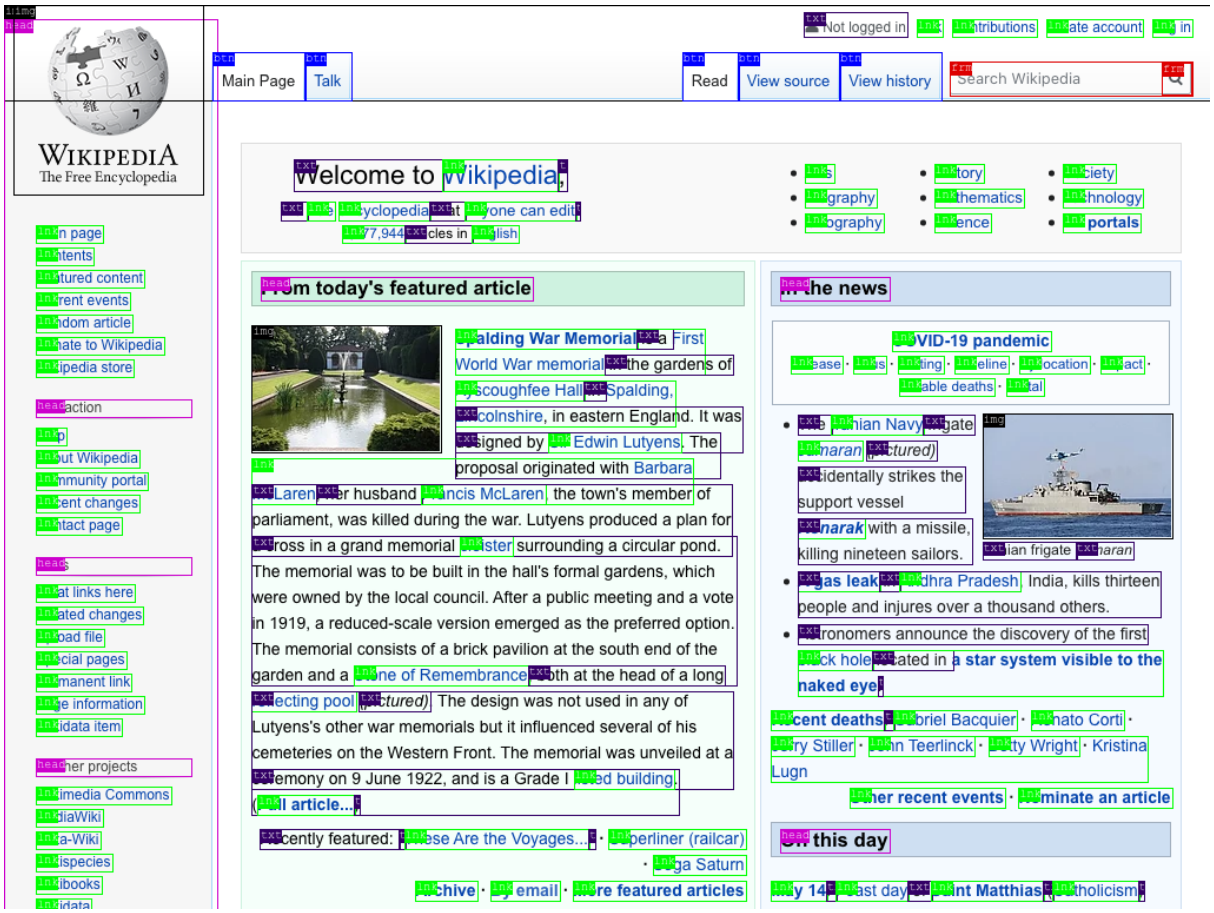
порівнювати скільки елементів змінилося з останнього переходу не неї і у випадку багатьох змін, вона перезапускає алгоритм аналізу сторінки (малюнок номер 3.6.).

Малюнок 3.6. Приклад коду для переведення елементів html сторінки в класи для подального використання програмою.

Отже, ми розглянули розробку статистичного аналізу сторінки. Тепер ми можемо розглянути роботу алгоритмів штучного інтелекту які можуть бути використані для аналізу веб сторінки за допомогою штучного зору.

Для того щоби зрозуміти яким чином здійснюється аналіз, нам треба зрозуміти як працюють ці алгоритми. Сучасні алгоритми штучного зору працюють за архітектурою CNN або конволюційною нейронною мережею. Ця технологія дозволяє зводити велике зображення з багатьма вимірами до меншого зображення з меншою вимірністю зі збереженням якостей великого. Це дозволяє значно швидше обробляти багатовимірні об’єми даних, ніж би це зробила звичайна модель. Для побудови алгоритму розпізнавання елементів на сторінці я використав алгоритм yolov5. Він перекладається як “Поглянь лише раз” і використовується для швидкого тренування на визначення елементів на фото та відео. Написаний він на мові програмування Python за допомогою фреймворку для написання алгоритмів машинного навчання pyTorch [26]. Я використав його

уже треновану модель YOLOv5m - це мережа середнього розміру яка відзначається одночасно і швидкістю і точністю. Також для того щоби почати роботу з визначенням елементів на сторінці, я використав уже готовий розмічений датасет на тисячу зображень [27]. В ньому уже анотовані всі готові елементи, які потрібні для тренування моделі.



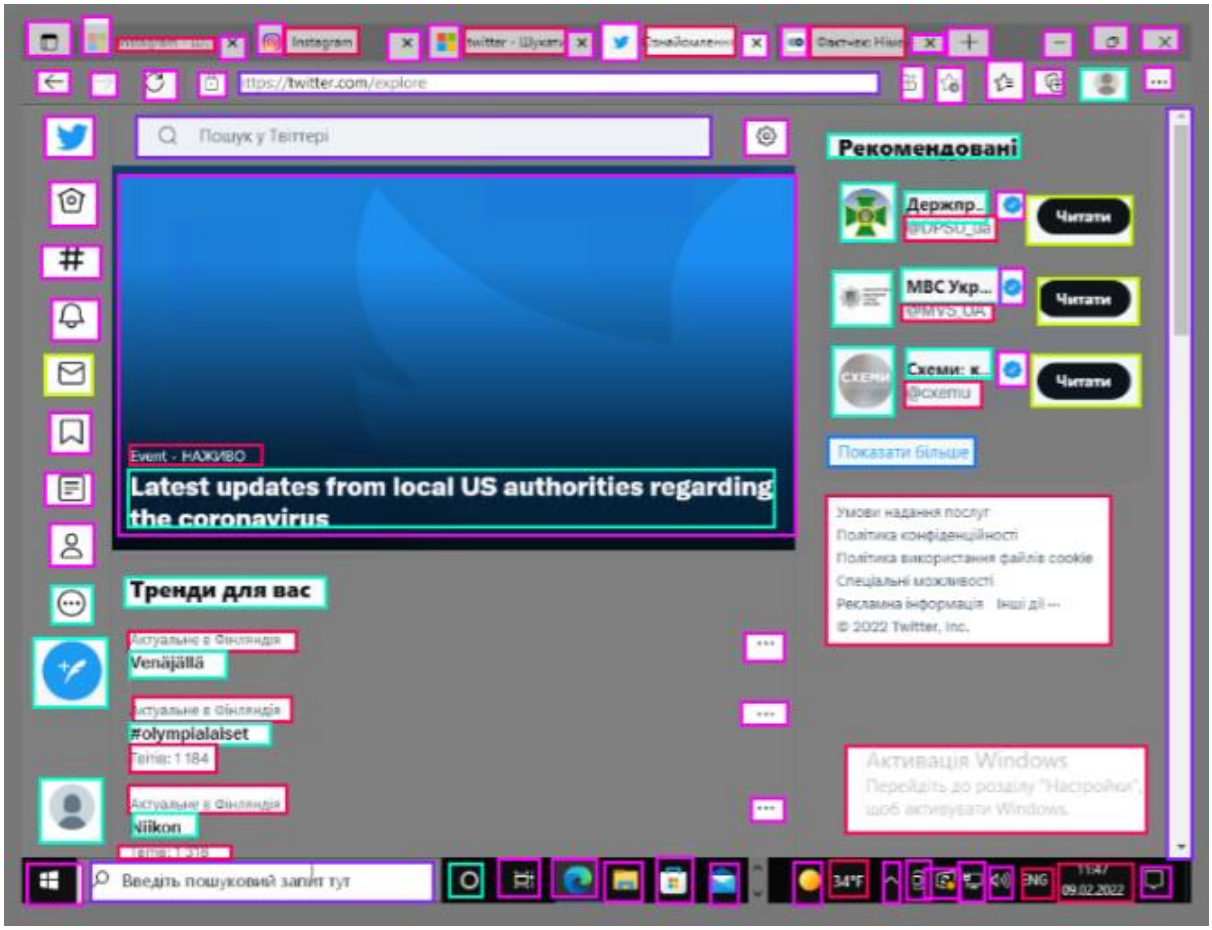
Малюнок 3.7. Приклад анотованого зображення. Кожний колір це інший клас. Всього класів - 7.

Також я створив свій власний додатковий датасет, який мав би визначати такі елементи на сторінці, як кнопки лайків, репостів ітд. Їх я знаходив у мережі інтернет. Також я зробив близько 50-ти скріншотів різних соціальних мереж і теж провів їх анотацію. Тобто усіх можливих елементів які можуть бути на сторінці соціальної платформи.

Разом я зібрав їх також близько 1000 і з кожної зробив анотацію. Використовував я для цього веб-сайт для анотації під назвою RoboFlow [28]. Крім цього я здійснив зміни на кожній з отриманих картинок і збільшив мій датасет в 6 разів.

- Горизонтальне віддзеркалення 50%
- Випадкове обтинання від 0 до 30 відсотків зображення
- Довільне обертання між -5 і +5 градусами
- Коригування випадкових бризок від -25 до +25 відсотків
- Коригування випадкового впливу від -25 до +25 відсотків
- На 2 відсотки пікселів застосовувався шум

На відкритому датасеті який я отримав онлайн я здійснив тренування моєї моделі та отримав 70% точності та рекол на рівні 80%. На власному датасеті який я створив згодом було 60% точності і 40% рекол. Відповідно другий датасет був недостатньо сформований і містив багато анотацій, які не змогли бути виділені моделлю.



Малюнок 3.8. Приклад анотації датасету

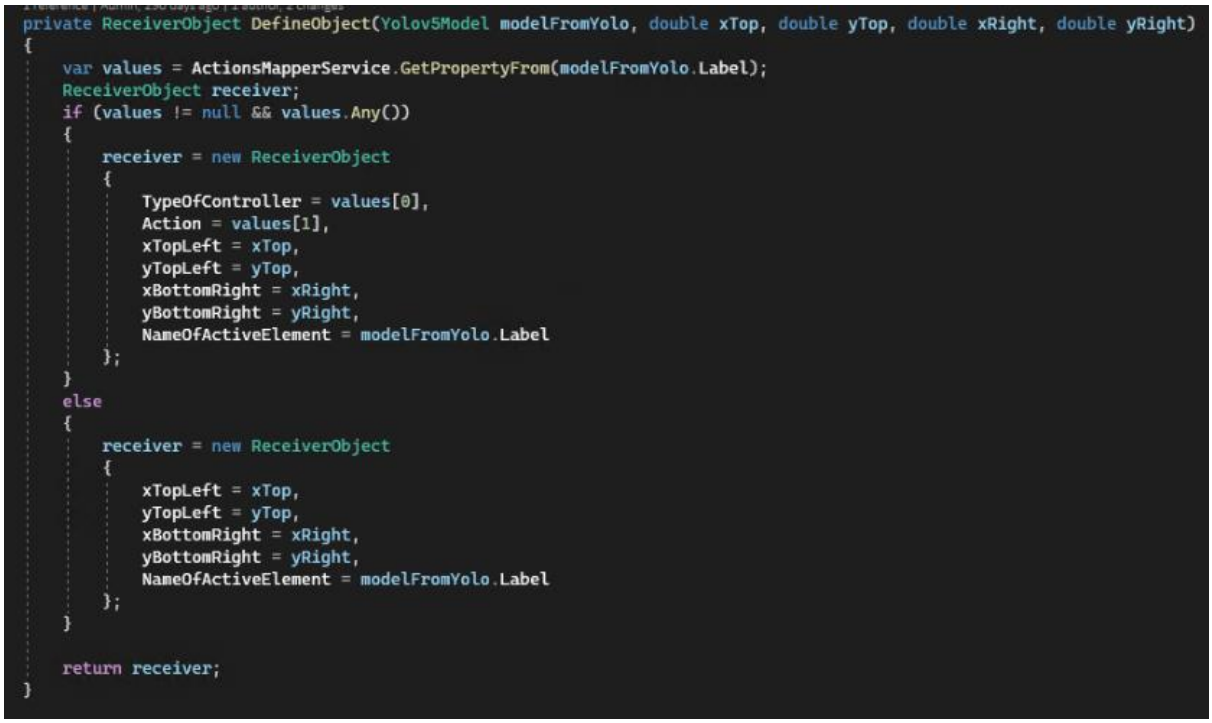
Адже до цього датасет містив 8 класів для визначення тепер в новому датасеті основних назв елементів для веб сторінки я також обрав і інші класи. Такі як: кнопка для посилання повідомлень (вона виділена картинкою поштового листа), кнопка для закривання (виділена косим хрестом) та кругла картинка. Цими анотаціями, я хотів отримати більш точну ідентифікацію елементів під час подальшої роботи програми, однак мені це не вдалося. Проте нова модель краще справляється з визначенням кольорових і заокруглених кнопок. Таких же, які є на сторінках реєстрації або логіну в соціальних мережах.
Тепер коли ми розібрали яким чином відбувається робота самого визначення інформації з зображень, ми розберемо те як система їх аналізу працює. Аналіз даних відбувається на сервері аналітики, який по суті є Flask застосунком написаним на Python 3. При своїй роботі він запускає код Yolov5 моделі

написаної за допомогою Pytorch. Для аналізу цей сервіс отримує лише зображення від центрального сервера. У відповідь надсилає усю зібрану інформацію про кількість знайдених елементів на сторінці та їх позиції відносно картинки. Саме зображення надсилається на сервер аналітики через центральний сервер з клієнтського додатку кожний вибраний часовий проміжок. Згодом для кожного класу, отриманого при оброблені фотографій, створюється новий об’єкт класу ReceiverObject, який здійснює перенесення інформації між різними частинами системи. В собі він зберігає інформацію про позицію об’єкту на сторінці, дію, яку має здійснити клієнт, дані, які має клієнт отримати і використати при своїй роботі. Далі ці об’єкти передаються в базу даних і зберігаються разом з інформацією про певну скановану сторінку. Тобто, як ми бачимо, обидва методи обробляють інформацію і зберігають ці дані до подальшого їх використання. При запиті на головний сервер про подальші дії ці дані будуть використані для створення нових команд для роботи з цим вебсайтом. Тож тепер, коли ми поговорили про роботу з аналізу вебресурсів, ми перейдемо до обговорення серверу контролю і клієнтської частини усієї системи.

3.3. Частина серверу контролю

Сервер контролю - це програма написана на C# з використанням фреймворку ASP.NET Core. Цей сервер має декілька класів для оброблення http запитів і найголовніші з них це: HttpController та HomeController. Ці два класи отримують запити напряму від клієнта. Через HttpController відбувається отримання запитів на інформацію про ту вебсторінку, яка використовується на даний момент. А через HomeController - йде робота метою якої є отримання команд від цільового сервера, отримання зображень з персонального екрану від клієнта і отримання команд від цільового користувача. Також через центральний сервер йде робота з аналізу отриманих даних, їх зберігання, відбувається відправка на сервер аналізу зображень та створення нових команд для клієнтів.

Про сервіс аналітики зображень ми уже говорили, тепер розглянемо як працює аналіз отриманих даних для створення нових команд.
Як було зазначено до цього, після отримання даних, створюються готові для роботи об’єкти в яких записана інформація про всі важливі елементи на сторінці (малюнок 3.9.).

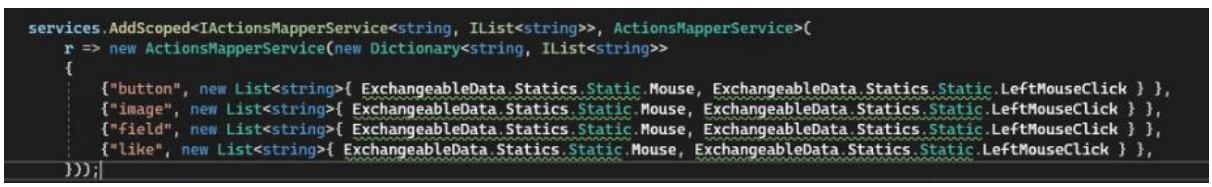


```
private ReceiverObject DefineObject(Yolov5Model modelFromYolo, double xTop, double yTop, double xRight, double yRight)
{
    var values = ActionsMapperService.GetPropertyFrom(modelFromYolo.Label);
    ReceiverObject receiver;
    if (values != null && values.Any())
    {
        receiver = new ReceiverObject
        {
            TypeOfController = values[0],
            Action = values[1],
            xTopLeft = xTop,
            yTopLeft = yTop,
            xBottomRight = xRight,
            yBottomRight = yRight,
            NameOfActiveElement = modelFromYolo.Label
        };
    }
    else
    {
        receiver = new ReceiverObject
        {
            xTopLeft = xTop,
            yTopLeft = yTop,
            xBottomRight = xRight,
            yBottomRight = yRight,
            NameOfActiveElement = modelFromYolo.Label
        };
    }

    return receiver;
}
```

Малюнок 3.9. Частина коду де відбувається перетворення даних отриманих з сервісу аналітики зображень в потрібний для роботи формат.

Для отримання інформації про дії, які необхідно виконати використовується просто звичайний маппінг властивостей за іменами класів, які виявила модель (Малюнок 3.10.). Маппінг - це статичне зіставлення вихідних даних з вхідними. Таким чином працюють, наприклад, словники, які є елементами різних мов програмування і видають значення відповідно до інших значень.



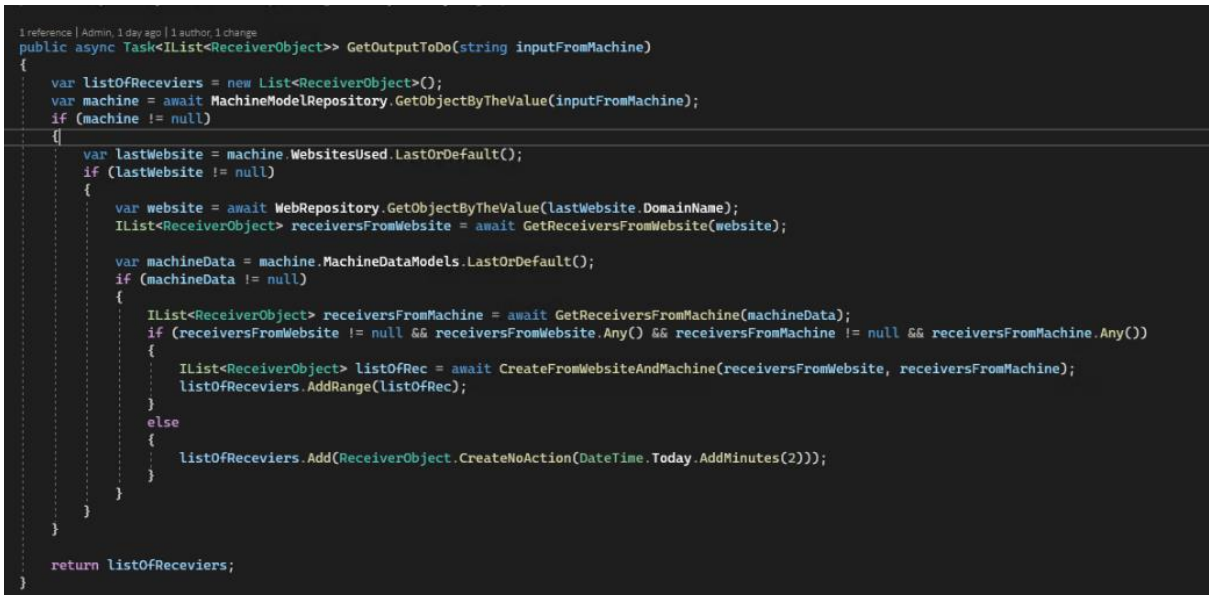
```
services.AddScoped<IActionsMapperService<string, IList<string>>, ActionsMapperService>(
    r => new ActionsMapperService(new Dictionary<string, IList<string>>
    {
        {"button", new List<string>{ ExchangeableData.Statics.Static.Mouse, ExchangeableData.Statics.Static.LeftMouseClick } },
        {"image", new List<string>{ ExchangeableData.Statics.Static.Mouse, ExchangeableData.Statics.Static.LeftMouseClick } },
        {"field", new List<string>{ ExchangeableData.Statics.Static.Mouse, ExchangeableData.Statics.Static.LeftMouseClick } },
        {"like", new List<string>{ ExchangeableData.Statics.Static.Mouse, ExchangeableData.Statics.Static.LeftMouseClick } },
    }));
```

Малюнок 3.10. Статичний мапінг для виявлених класів моделі до назви дій, які з ними можна виконувати. На даний момент використані лише дії з натисканням на клавішу комп’ютерної мишки.

Таке перетворення відбувається для того, щоби передати ці отримані дані в одному форматі на клієнта, який діє відповідно цим даним всередині своєї віртуальної машини. Адже, по суті, центральний сервер це програма, яка працює на господарській серверній машині та роздає накази іншим гостьовим віртуальним машинам через реверс http - запити. Реверс в даному випадку означає що клієнти самі кожний визначений проміжок часу запитують нові команди у центрального сервера. Така архітектура дуже подібна до роботи ботнетів з контролюючими та підлеглими машинами, про які було наведено приклад у першому розділі.

Отже, розглянемо що відбувається дали при надходженні запиту від клієнта на отримані команди. Сервер починає збирати інформацію про клієнта, який надіслав цей запит. Він отримує інформацію про вебсайт на якому зараз знаходиться клієнт, про зібрану інформацію щодо елементів які розташовані на вебсайті та на основі цього розробляє сценарій роботи для клієнта. Термін сценарій тут використаний в тому контексті, що кожна дія, яка має бути виконана клієнтом має бути виконана, таким чином щоби антифрод система не зрозуміла, що там присутня автоматизація. Наприклад, мишка користувача не може зразу потрапляти на визначену кнопку або він не може завжди вводити правильні дані в поля. Саме тому для імітації такої активності було вирішено створювати заготовлений список різних дій які може зробити користувач і під час запиту на нові команди цей список має бути заповнений випадковими помилковими діями, які згодом виправляються на правильні. Також це рішення порівнює інформацію отриману від обох джерел (проаналізованої вебсторінки і даних з зображення вебсайту) та отримує загальну найбільш точну картину того що відбувається на екрані у клієнта. Згодом отримується інформація про попередні дії які можуть

бути зроблені з вебсайтом, а також дії, які можуть бути вчинені з машиною клієнтом (наприклад вимкнути, запустити якусь програму) і, в підсумку, отримується список усіх дій, які необхідно виконати клієнту з урахуванням задачі, поставленої користувачем. А, якщо жодних задач на даний момент не може бути виконано, сервер передає запит на чекання в дві хвилини (Малюнок 3.11.).



```
1 reference | Admin, 1 day ago | 1 author, 1 change
public async Task<IList<ReceiverObject>> GetOutputToDo(string inputFromMachine)
{
    var listOfReceviers = new List<ReceiverObject>();
    var machine = await MachineModelRepository.GetObjectByTheValue(inputFromMachine);
    if (machine != null)
    {
        var lastWebsite = machine.WebsitesUsed.LastOrDefault();
        if (lastWebsite != null)
        {
            var website = await WebRepository.GetObjectByTheValue(lastWebsite.DomainName);
            IList<ReceiverObject> receiversFromWebsite = await GetReceiversFromWebsite(website);

            var machineData = machine.MachineDataModels.LastOrDefault();
            if (machineData != null)
            {
                IList<ReceiverObject> receiversFromMachine = await GetReceiversFromMachine(machineData);
                if (receiversFromWebsite != null && receiversFromWebsite.Any() && receiversFromMachine != null && receiversFromMachine.Any())
                {
                    IList<ReceiverObject> listOfRec = await CreateFromWebsiteAndMachine(receiversFromWebsite, receiversFromMachine);
                    listOfReceviers.AddRange(listOfRec);
                }
                else
                {
                    listOfReceviers.Add(ReceiverObject.CreateNoAction(DateTime.Today.AddMinutes(2)));
                }
            }
        }
    }

    return listOfReceviers;
}
```

Малюнок 3.11. Створення команд для клієнтів.

Отже, як ми можемо побачити, в загальному система складається з центрального сервера, сервера аналітики зображень та клієнтської частини. Проте, для того щоби перейти до клієнтської частини, нам потрібно також розглянути супутні елементи, які використовуються на клієнтській частині. Це засоби захисту від антифрод систем. Вони складаються з комплексної моделі, яка має на меті приховати всі можливі дані про своє використання та мімікрувати під звичайного користувача.

3.4. Антидетект частина системи мережі ботів

Створення антидетект частини нашої системи було виконане за допомогою платної програми Antidetect 3 Patreaon Edition. У своїй основі вона використовує віртуалізацію від опенсорс проекту VirtualBox. Більшість інформації про сам антидетект зберігають у таємниці, однак він допомагає своїм користувачам повністю підмінити всі ідентифікатори на цільовій операційній системі, перекрити можливість антифрод рішенням збирати інформацію про мережеве оточення і, в результаті, допомагає досягти найвищого рівня маскування. Також, крім цього, я використовував проксі сервіси від різних компаній провайдерів проксі такі як LuxSocks або AstroProxy. Для додаткової захищеності при роботі був використаний IpVanish приватний vpn сервіс. Vpn - це віртуальна приватна мережа, яка використовується для захисту свого вебтрафіку за допомогою його тунелювання через захищені сервери провайдера vpn.

Для початку роботи через антидетект програму створюється відповідна віртуальна машина на якій завантажена операційна система Windows 10. Для новоствореної віртуальної машини також можна вибрати базову конфігурацію. Наприклад, можна вибрати цільові платформи з якими буде працювати користувач. Це може бути Google, Facebook та інші інтернет гіганти з якими працюють підписники на цей програмний продукт. Після первинного налаштування на віртуальну машину встановлюється тунелювання трафіку через вибраний проксі сервіс. Також на головному сервері запускається vpn сервіс. Це робиться для закривання можливості віртуальній машині робити udp запити не через сервіси анонімізації, а також щоби захистити від витоків інформації про локальний днс сервер. На саму машину ставиться браузер та інше необхідне для роботи програмне забезпечення. Наприклад, це .Net Framework 4.7 на якому написаний клієнт. .Net Framework вибраний тому що він в основному знаходиться на машинах користувачів для запуску користувацьких програм. Якщо його використання виявить антифрод система вона не буде реагувати на це, як на

аномальну активність. Потім на саму віртуальну машину переносяться проксі сервіс, а також клієнтська програма. Потім клієнтська програма запускається і чекає на команди від сервера. Після виконання цих дій ця перша віртуальна машина зберігається і з неї в подальшому робляться копії. Копій може бути скільки завгодно, все залежить лише від потужності вибраного користувачем серверу.

3.5. Клієнтська частина

Клієнтська частина це .Net Framework 4.7 сервісний додаток написаний на C#. Який використовує перехоплення натискання мишки і клавіш через стандартні функції в операційній системі Windows 10. Програма здатна запускати деякі програми через виконання команд на сервері, а також вимикати операційну систему. Клієнт отримує всю інформацію, а також нові конфігурації напряму від сервера контролю, здійснюючи http запити на його веб адресу.

3.6. Додаткові сервіси

Крім роботи з клієнтом, варто також згадати що головний сервер здатний викликати команди для роботи з віртуальними машинами. Зроблено це було для того щоби була можливість керувати віртуальними машинами не лише зсередини, а і зовні. Команди, які підтримуються на даний момент мінімальні. Це увімкнути, ввімкнути і виділити віртуальну машину. Також було використано декілька старих емейл адрес з сервісу Gmail для створення облікових записів.

Висновок

Отже, ми розглянули усі частини нашої системи по автоматизації роботи. Тепер ми можемо побудувати повну модель того, яким чином відбувається наша робота з автоматизацією акаунтів і можемо зрозуміти, які ще кроки необхідно зробити власноруч або автоматизувати.

1. Генеративна частина

a. Всі дії відбуваються вручну, автоматизації немає

2. Автоматизація отримання смс або емейлів

a. Був використаний емейл для реєстрації, а також приватний номер телефону

b. Заповнення профілю

i. Також вручну, адже система не готова для роботи з аутентифікацією

ii. Проте вона може бути використана, уже після аутентифікації для автоматичного заповнення даних профілю

2. Аналітична частина

a. Вміє розрізняти різні елементи на сторінці

b. Проксі сервіс перехоплює всі запити і досить ефективно вдається статично аналізувати сторінку

2. Контролююча частина

a. На даний момент сервер вміє чекати, вводити інформацію в поля, натискати на кнопки, відкривати програми

b. Ще не реалізована функція створення сценарію через великий об’єм роботи з встановлення і тестування кожної окремої частини системи

c. Запускає віртуальні машини, розблоковує їх робить первинне налаштування

2. Інтерфейс для клієнтського відображення інформації

a. Готовий лише частково цей інтерфейс ще не підключений до контролюючого серверу

2. Клієнтський додаток

a. Працює ефективно, протестований на виконання всіх необхідних задач.

2. Робота з віртуальними машинами

a. Вона ведеться примітивно через використання автоматизації клавіатури і логіну в кожній машині через натискання клавіш на господарській машині

b. Це не є ефективною програмною реалізацію, однак іншого варіанту для автоматизації такої роботи, я не знайшов.

Проаналізувавши всі елементи системи, можна сказати що побудувати цілісну робочу систему мені не до кінця вдалося. Проте використовуючи названі до цього методи захисту від антифрод активності, я можу впевнено сказати, що вони дозволяють створити акаунти в будь-якій соціальній мережі, незважаючи антифрод захист.

А також мені вдалося виконати цільову дію на вебсайті соціальної мережі Twitter, а саме зайти в акаунт, та поставити лайк на першому пості, який відкрився. На жаль, через неспровоковану агресію з боку росії і окупацію великої частини української території у мене була втрачена можливість роботи з моїми серверними рішеннями на яких відбувалося тренування моїх моделей для машинного навчання. Через окупацію міста Херсон, в яких знаходилися мої орендовані серверні рішення і їх подальше розграбування - мені не вдалося відновити велику кількість моїх матеріалів для роботи. Однак, на щастя, був збережений програмний код, з яким я працював на даний момент та вдалося відновити деяку частину цієї інфраструктури.

ВИСНОВКИ

1. Боти можуть порушувати усі три компоненти системи кібербезпеки вебресурсів. Оскільки вони створені саме для того щоб обходити певні обмеження цих вебсайтів і успішно автентифікуватися в них, бо є програмами автоматичного доступу до закритої частини вебсайту, то найбільшу загрозу вони складають для збереження конфіденційності інформації.
2. Сканери є незамінними утилітами для роботи з тестуванням на вразливості. Але, коли мова йде про неправомірний доступ до цільового сервера за допомогою бота, боти використовують сканери на вразливості або зловмисні утиліти віддаленого доступу, які вони можуть залишати після успішної атаки. Після проходження аутентифікації бот може запустити свій вбудований сканер на вразливості, який буде шукати вразливості на вебсайті, а потім, при знаходженні потрібних, виконати корисне навантаження і отримати неправомірний доступ до цільового сервера.
3. Використання для атаки експлойтів 0-го дня або таких шкідливих програм, що були спеціально створені під конкретну задачу чи об'єкт враження значно ускладнюють захист “жертви” атаки. Проте, оскільки кібербезпека все більше розвивається, створення таких експлойтів стає все більш затратною справою в грошовому і часовому еквіваленті. Саме тому, більшість кіберзлочинців використовуватимуть доступні уже відомі експлойти і будуть направляти свої атаки на застарівше програмне забезпечення.
4. Зловмисна система з центром управління дає набагато більше гнучкості для її розробників, дозволяє використовувати її в довгостроковій перспективі, оновлювати своє програмне забезпечення, а також вести свою діяльність довгий час без того щоби бути поміченими правоохоронними органами. Тому на даний момент основний спосіб боротьби з таким програмним

забезпеченням - це активний моніторинг усіх вразливих девайсів, зміна паролів на більш захищені, створення навмисно вразливих систем-пасток (ханіпотів) для виявлення та ідентифікації злочинної діяльності.

5. Оскільки соціальні платформи поширюють інформацію серед величезної аудиторії та мають великий вплив на громадську думку майже третини населення в усьому світі, все більша частина зловмисників намагатиметься використовувати їх як зброю в інформаційних війнах. Для автоматизації взаємодії, зменшення витрат та максимального охоплення аудиторії, використовуються боти, які дозволяють автоматично взаємодіяти з акаунтами користувачів у соціальних мережах.
6. Бот – це ефективний інструмент, який робить процес просування ідей серед користувачів інтернету легким та доступним. Проте, щоб ідею успішно поширити, необхідно враховувати психологію впливу на користувача як на людину – ефективність донесення ідеї залежить від цього, як саме бот реалізує цей процес. Одночасне застосування кількох принципів переконання людей призводить до більш вираженого впливу прийняття рішень. Основна причина, через яку ці принципи ефективні, полягає в тому, що вони служать евристиками (розумовими спрощеннями) – способом обробки інформації, на який зазвичай покладаються.

   Використовуючи одночасно декілька принципів переконання (принцип авторитету, принцип соціального схвалення, принцип симпатії, принцип дефіциту), зацікавлені особи нав'язують інформацію користувачеві, спонукаючи його до дій фінансового характеру. Іноді це може дозволити іншим отримати доступ до конфіденційної інформації, яка може бути використана зловмисно.

7. Ілюзію соціального схвалення в соціальних мережах допомагають створити програми-боти, що генерують штучні позначки «мені подобається», коментарі під фото та записами на сторінці людини, що підвищують

кількість передплатників та друзів певного суб'єкта. В подальшому таке соціальне схвалення монітезується з використанням різних інструментів.

8. Мережі ботів є реальними загрозами для фінансового сектору. Особливо вразливими є біржі криптовалют, банківські установи та фондові ринки. Крім прямої крадіжки фінансових ресурсів та інформації, існують репутаційні ризики, що можуть штучно створюватись мережею ботів, та напряму впливають на фінансові рейтинги, вартість бізнесу, акцій, тощо.
9. Використання новітніх систем на основі штучного інтелекту для виявлення і знищення шкідливого програмного забезпечення є одним з найперспективніших рішень у сфері безпеки та використовується для виявлення та знищення шкідливого програмного забезпечення. Комбіннуючи SVM із отриманням інформації та BPSO, можна отримати два методи зменшення параметрів функцій та використати їх для створення системи виявлення на вторгнення, особливо у випадку атак «Відмова в обслуговуванні» (DoS), атаки U2R і R2L. Перспектиними є використання алгоритмів фільтрації на основі класифікатора SVM для ідентифікації кількох класів вторгнення. Вартим уваги є дослідження метою якого є створення системи виявлення вторгнень та запропоновано нові методи для покращення продуктивності класифікації вторгнень через kNN. Крім того для створення системи пошуку вторгнень досить часто використовують алгоритм машинного навчання – дерево рішень, зважаючи на його структуру та точність. Для виявлення загроз використовується також і логістична регресія.

10. Глибоке навчання є окремою групою машинного навчання. Алгоритми глибокого навчання так само широко використовується в кібербезпеці для боротьби з розподіленими атаками на відмову в обслуговуванні.
11. Загалом за своєю суттю системи протидії фроду це комплексні системи, які направлені на успішну ідентифікацію та унікалізацію своїх користувачів, з подальшим присвоєнням їм свого власного фрод рейтингу.

12. Головні етапи ідентифікації користувача при застосуванні системи протидії фроду - збір інформації про користувача, її подальшого аналізу, формування рейтингу користувача і згодом прийняття рішення про надання дозволу на використання, запит додаткової інформації або заборона на вчинення будь-яких дій у системі.
13. Аналіз рівнів ідентифікаторів користувача, які збирають антифрод системи (мережевий рівень, браузерний рівень, рівень операційної системи та апаратний рівень) дозволяють зрозуміти, яким чином нам потрібно будувати свою систему для подальшого їх обходу.
14. Антифрод системи не потрібно аналізувати, лише за одними аргументами. Кожний з аргументів, які отримуються фрод системою аналізуються окремо і формують загальний фрод рейтинг користувача. На основі якого і приймається згодом рішення про використання санкцій до цього користувача, його додаткової перевірки або дозволу на роботу в мережі. Наступним етапом, за статичним аналізом йде також і поведінковий аналіз. Соціальні мережі, а також фінансові організації, такі як банки, наприклад, розробляють свої власні системи поведінкового аналізу, які вирізняються досить високою здатністю до ідентифікації фроду.
15. Розгдянувши усі аспекти, можна побудувати повну модель того, яким чином відбувається робота з автоматизацією акаунтів і визначити дії, які необхідно зробити власноруч або автоматизувати.

1. Генеративна частина

a. Всі дії відбуваються вручну, автоматизації немає

2. Автоматизація отримання смс або емейлів

c. Був використаний емейл для реєстрації, а також приватний номер телефону

d. Заповнення профілю

i. Також вручну, адже система не готова для роботи з аутентифікацією

ii. Проте вона може бути використана, уже після аутентифікації для автоматичного заповнення даних профілю

3. Аналітична частина

a. Вміє розрізняти різні елементи на сторінці

b. Проксі сервіс перехоплює всі запити і досить ефективно вдається статично аналізувати сторінку

4. Контролююча частина

a. На даний момент сервер вміє чекати, вводити інформацію в поля, натискати на кнопки, відкривати програми

b. Ще не реалізована функція створення сценарію через великий об’єм роботи з встановлення і тестування кожної окремої частини системи

c. Запускає віртуальні машини, розблоковує їх робить первинне налаштування

5. Інтерфейс для клієнтського відображення інформації

a. Готовий лише частково цей інтерфейс ще не підключений до контролюючого серверу

6. Клієнтський додаток

a. Працює ефективно, протестований на виконання всіх необхідних задач.

7. Робота з віртуальними машинами

   a. Вона ведеться примітивно через використання автоматизації клавіатури і логіну в кожній машині через натискання клавіш на господарській машині

   b. Це не є ефективною програмною реалізацію, однак іншого варіанту для автоматизації такої роботи, я не знайшов.

Проаналізувавши всі елементи системи, можна сказати що побудувати цілісну робочу систему мені не до кінця вдалося. Проте використовуючи названі до цього методи захисту від антифрод активності, я можу впевнено сказати, що вони дозволяють створити акаунти в будь-якій соціальній мережі, незважаючи антифрод захист.

А також мені вдалося виконати цільову дію на вебсайті соціальної мережі Twitter, а саме зайти в акаунт, та поставити лайк на першому пості, який

відкрився. На жаль, через неспровоковану агресію з боку росії і окупацію великої частини української території у мене була втрачена можливість роботи з моїми серверними рішеннями на яких відбувалося тренування моїх моделей для машинного навчання. Через окупацію міста Херсон, в яких знаходилися мої орендовані серверні рішення і їх подальше розграбування - мені не вдалося відновити велику кількість моїх матеріалів для роботи. Однак, на щастя, був збережений програмний код, з яким я працював на даний момент та вдалося відновити деяку частину цієї інфраструктури.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Bach-Nutman, M. (2020) Understanding the top 10 Owasp vulnerabilities, arXiv.org. Available at: https://arxiv.org/abs/2012.09960 (Accessed: June 15, 2022).
2. OWASP. “OWASP Top 10 Vulnerabilities”. owasp.org., https://owasp.org/www-project-top-ten/ (Accessed Oct. 5, 2022).
3. OWASP, https://owasp.org/
4. Khalid, M.N. et al. (2020) “Web vulnerability finder (WVF): Automated black-box web vulnerability scanner,” International Journal of Information Technology and Computer Science, 12(4), pp. 38–46. Available at: https://doi.org/10.5815/ijitcs.2020.04.05.
5. Sharif, M.H.U. and Mohammed, M.A. (2022) A literature review of financial losses statistics for Cyber Security and future trend, World Journal of Advanced Research and Reviews. World Journal of Advanced Research and Reviews. Available at: https://wjarr.com/content/literature-review-financial-losses-statistics-cyber-security-and-future-trend (Accessed: November 17, 2022).
6. Samonas, S. (no date) The CIA strikes back: Redefining confidentiality, integrity and Availability. Available at: https://www.proso.com/dl/Samonas.pdf (Accessed: June 17, 2022).
7. Малюнок CIA - https://en.wikipedia.org/wiki/File:CIAJMK1209-en.svg
8. “NOTPETYA TECHNICAL ANALYSIS,” LogRhythm Labs. [Online]., Available: https://logrhythm.com/pdfs/threat-intelligence-reports/notpetya-technical-analysis-threat-intelligence-report.pdf
9. “NotPetya: World's First $10 Billion Malware”, By Rich Tehrani Group Editor-in-Chief, TMC, https://www.apextechservices.com/topics/articles/435235-notpetya-worlds-first-10-billion-malware.htm#

10. Antonakakis, M. et al. (2017) Understanding the Mirai botnet, USENIX. Available at: https://www.usenix.org/conference/usenixsecurity17/technical-sessions/presentation/antonakakis (Accessed: June 18, 2022).
11. Мал. номер, https://paalsh.com/wp-content/uploads/2021/03/mirai.png
12. “СБУ ліквідувала «ботоферму», яка фінансувалась з країни-агресора”, https://ssu.gov.ua/novyny/sbu-likviduvala-botofermu-yaka-finansuvalas-z-krainyahresora
13. Вебінар Дмитра Момота, “Что сайты могут узнать о вашей ОС, установленном ПО и как могут это использовать для вашей оценки”, Vektor T13 Youtube channel, 2020 рік, https://www.youtube.com/watch?v=puchBhU2Ivg
14. Вебінар Дмитра Момота, “Что сайты знают о вас и как антифрод системы оценивают ваше "железо"”, Vektor T13 Youtube channel, 2020 рік, https://www.youtube.com/watch?v=mN4IZP51teI
15. Вебінар Дмитра Момота, “Что сайты могут узнать о вашей ОС, установленном ПО и как могут это использовать для вашей оценки”, Vektor T13 Youtube channel, 2020 рік, https://www.youtube.com/watch?v=puchBhU2Ivg
16. Canvas, https://en.wikipedia.org/wiki/Canvas_element
17. Lindner, T. et al. (1970) Exploitation of public and private WIFI coverage for new business models, SpringerLink. Springer US. Available at: https://link.springer.com/chapter/10.1007/1-4020-8155-3_8 (Accessed: June 19, 2022).
18. Вебінар Дмитра Момота, “"Почему мой антидетект не работает" или как антифрод системы анализируют браузер посетителя”, Vektor T13 Youtube channel, 2020 рік, https://www.youtube.com/watch?v=EJwLhbgMkuM
19. Какую информацию сайты собирают о вас и как ее подменить. Вебинар 4. Сетевой уровень., Vektor T13 Youtube channel, 2020 рік, https://www.youtube.com/watch?v=EJwLhbgMkuM
20. https://en.wikipedia.org/wiki/IPv6

21. Karras, T. et al. (2020) Analyzing and improving the image quality of stylegan, arXiv.org. Available at: https://arxiv.org/abs/1912.04958 (Accessed: August 20, 2022)
22. https://www.simplilearn.com/tutorials/deep-learning-tutorial/rnn
23. Titanium web proxy, https://github.com/justcoding121/titanium-web-proxy
24. https://github.com/AngleSharp/AngleSharp
25. https://pptr.dev/
26. https://github.com/ultralytics/yolov5
27. https://public.roboflow.com/object-detection/website-screenshots
28. https://roboflow.com/
29. Robert B. Cialdini, Roselle L. Wissler and Nicholas J. Schweitzer. The Science of Influence. Using six principles of persuasion to negotiate and mediate more effectively. Dispute Resolution Magazine, Fall 2002.
30. Nicole L. Muscanell1, Rosanna E. Guadagno and Shannon Murphy. Weapons of Influence Misused: A Social Influence Analysis of Why People Fall Prey to Internet Scams. Social and Personality Psychology Compass 8/7 (2014): 388–396, 10.1111/spc3.12115.
31. Cialdini, RB (2009). Influence: Science and Practice (5th edn). Boston: Allyn & Bacon.
32. Fiske, SR, & Taylor, SE (2013). Social Cognition: From Brains to Culture (2nd edn). Los Angeles, London, New York, Singapore, Washington D.C.: SAGE.
33. Guadagno, RE, Okdie, BM, & Muscanell, NL (2013). Have we all just become 'Robo-sapiens'? Reflections on social influence processes in the Internet Age. Psychological Inquiry, 24, 1–9.
34. Regan, DT (1971). Effects of a favor and liking on compliance. Journal of Experimental Social Psychology, 7, 627–639.
35. Halper, D (February, 2014). FBI warns of 'Online Dating Scams' before Valentine's day. Retrieved on March 4, 2014 from http://www.weeklystandard.com/blogs/fbi-warns-online-dating-scams-valentines-day_781509.html

36. Ashley, C. and Tuten, T. (2015), “Creative strategies in social media marketing: an exploratory study of branded social content and consumer engagement”, Psychology and Marketing, Vol. 32 No. 1, pp. 15-27.
37. Gottfried, J. and Shearer, E. (2016), News Use across Social Medial Platforms 2016, Pew Research Center.
38. Gesser-Edelsburg, A. and Shir-Raz, Y. (2017), “Science vs. fear: the Ebola quarantine debate as a case study that reveals how the public perceives risk”, Journal of Risk Research, Vol. 20 No. 5, pp. 611-633.
39. Natalie Sauer. Algorithms, bots and elections in Africa: how social media influences political choices. The conversation. URL: https://theconversation.com/algorithms-bots-and-elections-in-africa-how-social-media-influences-political-choices-179121
40. Kai Kupferschmidt. Social media ‘bots' tried to influence the U.S. election. Germany may be next. Science, 2019. URL: https://www.science.org/content/article/social-media-bots-tried-influence-us-election-germany-may-be-next
41. Ross J. Schuchard, and Andrew T. Crooks. Insights into elections: An ensemble bot detection coverage framework applied to the 2018 U.S. midterm elections. PLoS One. 2021; 16(1): e0244309.
42. Efthimion, Phillip; Payne, Scott; Proferes, Nicholas (2018-07-20). "Supervised Machine Learning Bot Detection Techniques to Identify Social Twitter Bots". SMU Data Science Review. 1 (2).
43. Keller, Tobias R ; Klinger, Ulrike. Social bots in election campaigns: theoretical, empirical, and methodological implications. 2019, Zurich Open Repository and Archive.
44. Чароенвонг, Б. та Бернарді, М. (2021) *Десятиліття «зломів» криптовалюти: 2011 – 2021, Бен Чароенвонг, Маріо Бернарді :: SSRN*. Доступно за адресою: http://dx.doi.org/10.2139/ssrn.3944435.

45. Orabi, M. *та ін.* (2020) *Виявлення ботів у соціальних медіа: систематичний огляд, Обробка та управління інформацією*. Пергам. Доступно за адресою: https://www.sciencedirect.com/science/article/abs/pii/S0306457319313937?via%3Dihub.
46. "Why companies don't report incidents", https://www.csoonline.com/article/3398700/why-businesses-don-t-report-cybercrimes-to-law-enforcement.html
47. "The Rise Of Social Media Botnets", https://www.darkreading.com/attacks-breaches/the-rise-of-social-media-botnets
48. "Injection Flaws", https://owasp.org/www-community/Injection_Flaws
49. Revil with ransomware, https://www.bleepingcomputer.com/news/security/revil-ransomware-scans-victims-network-for-point-of-sale-systems/
50. J. Margolis, T. T. Oh, S. Jadhav, Y. H. Kim and J. N. Kim, "An In-Depth Analysis of the Mirai Botnet," *2017 International Conference on Software Security and Assurance (ICSSA)*, 2017, pp. 6-12, doi: 10.1109/ICSSA.2017.12.
51. "Zero days are rarer and more expensive than ever" https://www.cyberscoop.com/zero-day-vulns-are-rarer-and-more-expensive-than-ever/
52. G. Zhao, K. Xu, L. Xu and B. Wu, "Detecting APT Malware Infections Based on Malicious DNS and Traffic Analysis," in *IEEE Access*, vol. 3, pp. 1132-1142, 2015, doi: 10.1109/ACCESS.2015.2458581.
53. "How To Automate Scrolling on a Rooted Android Phone", https://www.guidingtech.com/36293/automate-scrolling-android/
54. "Get Started", https://developers.facebook.com/docs/graph-api/get-started
55. "Elon Musk buys Twitter", https://www.androidauthority.com/elon-musk-buys-twitter-3226401/
56. "Facebook Privaxy Policy", https://www.facebook.com/privacy/policy
57. "Most common resolutions", https://gs.statcounter.com/screen-resolution-stats/desktop/worldwide

58. Proxy server and tunneling, https://developer.mozilla.org/en-US/docs/Web/HTTP/Proxy_servers_and_tunneling
59. Project Fugu, https://www.chromium.org/teams/web-capabilities-fugu/
60. "Why is it so hard to get rid of deepfakes?", https://medium.com/predict/why-is-it-so-hard-to-get-rid-of-deepfakes-e7130b608069
61. "Marketplace for generated photos", https://generated.photos/faces
62. "RNNs part 5", https://towardsdatascience.com/recurrent-neural-networks-part-5-885fc3357792
63. https://1library.net/article/malware-traffic-content-agnostic-malware-detection-networks.yj5r6g2q
64. Tanner, O. (2022) *Multi-agent car parking using reinforcement learning*, *arXiv.org*. Available at: https://arxiv.org/abs/2206.13338 (Accessed: September 24, 2022).
65. "Deep Reinforcement Learning Doesn't Work Yet", https://www.alexirpan.com/2018/02/14/rl-hard.html
66. https://www.forbes.com/sites/zakdoffman/2019/08/03/warning-for-windows-users-as-new-malware-hides-malicious-traffic-to-infected-pcs/?sh=2edde64955c0