УДК 004.056

Куперштейн Л.М.
Примаков Б.С.

Про використання ChatGPT в кібербезпеці

Вінницький національний технічний університет

Анотація

Досліджено можливості використання сервісу ChatGPT для підтримки розв'язання задач кібербезпеки, а саме аналізу загроз та вразливостей систем. Дослідження показало, що ця технологія цікава як «білим капелюхам» та і «чорним капелюхам».

Ключові слова: *чат-бот, сервіс ChatGPT, кіберзагроза, кібербезпека*

Abstract

The possibilities of using the ChatGPT service to support cybersecurity tasks, specifically threat and vulnerability analysis, have been investigated. The research showed that this technology is of interest to both "white hat" and "black hat" actors..

Keywords: *chat-bot, ChatGPT service, cybersecurity.*

Вступ

З розвитком інформаційних технологій та поширенням Інтернету кібербезпека стає однією з найбільш важливих та актуальних тем в сучасному світі. Несанкціонований доступ до інформації, крадіжка особистих даних та зловмисні атаки на комп'ютерні системи можуть призвести до серйозних наслідків, включаючи втрату конфіденційної інформації та фінансових збитків [1]. У зв'язку з цим у сфері кібербезпеки актуальним є розвиток нових технологій та інструментів для захисту від цих загроз. Для побудови таких інструментів все частіше починають використовуватися технології штучного інтелекту та машинного навчання [2-4].

Результати дослідження

Сервіс ChatGPT є машинним інтелектом, який може генерувати текст на основі вхідних запитань або фраз. Завдяки своїм можливостям в генерації тексту на основі аналізу мови, він може бути використаний для виявлення кіберзагроз.

ChatGPT може бути корисним інструментом в кібербезпеці для виявлення, аналізу та реагування на потенційні кібератаки. Нижче наведено аналіз декількох кейсів використання ChatGPT для розвязання задач кібербезпеки.

1) Виявлення загроз: ChatGPT може бути використаний для аналізу текстової інформації, такої як повідомлення електронної пошти, соціальні медіа, чати та інші джерела, для виявлення потенційних загроз кібербезпеці. Штучний інтелект може сканувати великі обсяги даних та автоматично виявляти підозрілу активність, яка може вказувати на кібератаку. Для цього ChatGPT може бути навчений на прикладах текстових повідомлень, що містять елементи кібернебезпеки, такі як незвичайні запити, спроби шахрайства, шифрування даних, спроби несанкціонованого доступу до систем та інші ознаки, що вказують на потенційну кібератаку. Після навчання ChatGPT може використовуватися для автоматичного виявлення підозрілої активності та повідомлення про неї адміністратору системи. Наприклад, якщо відбувається спроба несанкціонованого доступу до системи або спроба крадіжки даних, то ChatGPT може автоматично виявити цю активність та повідомити адміністратора про неї для подальших заходів щодо захисту системи. Також, ChatGPT може використовуватися для аналізу текстових повідомлень, які не включають елементи кібернебезпеки. Наприклад, він може аналізувати повідомлення електронної пошти та інші текстові повідомлення для виявлення підозрілої активності або неправдивої інформації, яка може викликати кібернебезпеку. Враховуючи зростаючу кількість кібератак, з якими доводиться стикатися користувачам та адміністраторам систем, використання ChatGPT в кібербезпеці може бути важливим інструментом для забезпечення безпеки та захисту від кіберзагроз [5].

2) Аналіз вразливостей: ChatGPT може бути використаний для аналізу різних вразливостей системи та програмного забезпечення. Це допоможе виявити потенційні ризики та допоможе в попередженні можливих атак. Аналіз вразливостей - це процес виявлення потенційних слабких місць в інформаційних системах, що можуть бути використані для здійснення кібератак та їх зламу [6]. Цей процес зазвичай включає в себе сканування, тестування та оцінку вразливостей систем, виявлення потенційних загроз та розробку стратегій захисту. ChatGPT може допомогти в аналізі вразливостей, наприклад, шляхом збору інформації з багатьох джерел та використання інструментів машинного навчання для аналізу цієї інформації. Зокрема, ChatGPT може бути використаний для:

2.1) Аналізу журналів подій: ChatGPT може бути використаний для аналізу журналів подій, що

містять інформацію про події, що відбуваються в інформаційній системі. Можна використовувати ChatGPT для аналізу цих журналів та виявлення потенційних загроз.

2.2) Виявлення аномальної поведінки: ChatGPT може використовуватися для виявлення аномальної поведінки в мережі, що може свідчити про наявність хакера або бот-мережі. Можна навчити ChatGPT розпізнавати певні типи аномальної поведінки, наприклад, надмірний трафік чи кількість запитів до серверу.

2.3) Виявлення потенційних вразливостей: ChatGPT може використовуватися для аналізу коду програмного забезпечення та виявлення потенційних вразливостей в ньому. Наприклад, можна навчити ChatGPT розпізнавати певні типи вразливостей, такі як SQL-ін'єкції, міжсайтові скриптові атаки та інші.

3) Попередження фішингу: ChatGPT може бути використаний для попередження фішинг-атак, які можуть призвести до крадіжки особистої інформації або фінансових засобів. Штучний інтелект може аналізувати повідомлення електронної пошти та інші текстові джерела та виявляти підозру.
Фішинг - це вид атаки, який полягає у використанні соціальної інженерії з метою витягнути конфіденційну інформацію від користувачів, таку як імена користувачів та паролі. Чат-боти на базі моделі GPT можуть бути використані для попередження фішингу шляхом розпізнавання характерних ознак фішингових повідомлень та сповіщення користувачів про потенційні загрози. Зокрема, система може бути навчена розпізнавати типові характеристики фішингових повідомлень, такі як ланцюжки символів, які намагаються імітувати URL-адресу офіційного сайту, запити на конфіденційну інформацію або непрохані запити на персональні дані. Крім того, ChatGPT може бути навчений розпізнавати особливості відправника повідомлення, такі як незнайомий емейл або ім'я, що не співпадає зі звичним іменем відправника. Якщо система розпізнає фішингове повідомлення, вона може сповістити користувача та надати інформацію про потенційну загрозу. Зокрема, ChatGPT може надати поради користувачеві щодо того, які кроки варто виконати для захисту своїх даних. Це може включати рекомендації щодо зміни паролів, перевірки URL-адреси, перевірки ідентифікаційних даних відправника та інших кроків, які допоможуть запобігти втраті конфіденційної інформації. Система може на основі ChatGPT може бути навчена розпізнавати різні види загроз та попереджувати користувачів про їх наявність. Крім того, система може бути використана для навчання користувачів, щоб вони могли бути свідомими щодо ризиків та захисту своїх персональних даних в Інтернеті.

Крім корисних можливостей ChatGPT, темна сторона людини теж знайшла спосіб, як використовувати цей сервіс для хакінгу. В одному зі своїх звернень директор OpenAI, Сем Альтман сказав, що використання ШІ може стати «вимкненим світлом для всіх». І далеко не безпідставно, адже уже з'явилось багато аудиторії, зацікавленої у використанні штучного інтелекту у своїх темних цілях. Хоча самі розробники і вводять обмеження на використання ChatGPT, існують доволі прості способи перешкодити модерації вмісту. Так, наприклад, певні ресурси пропонують за менш ніж за $6 згенерувати у Chat GPT шкідливий код або фішингові листи. Для цього хакери використовують API OpenAI та за допомогою telegram-ботів інтегрують функціонал в свої канали, обходячи обмеження модерації. Клієнти таких ресурсів отримують 100 запитів всього за $5.5, а продавці демонструють переконливі приклади шкідливих речей, які можна згенерувати.

Інший спосіб передбачає обхід обмежень через спеціальні сценарії. Хакери засобами ChatGPT змогли згенерувати шаблони фішингової розсилки, маскуючись під адміністарацію банку, магазину тощо [7]. При цьому ChatGPT навіть радить в якому місці найкраще розмістити фішингове посилання. Лякає і те, що користувачі, які не мають навичок програмування, мають змогу генерувати в Chat GPT шкідливі скрипти, які може зашкодити іншим користувачам. Таким чином нас може накрити хвиля спам-листів, шкідливих файлів, замаскованих під архіви, відео та фото тощо.

Висновки

Під час дослідження було визначено потенційні напрями використання сервісу ChatGPT в сфері кібербезпеки. Дослідження показало, що сервіс ChatGPT може бути ефективним інструментом як для виявлення та запобігання кібератак, так і їх планування та виконання. При цьому було виділено лише загальні сценарії використання, але потенційні можливості цього інтелектуального інструменту значно більші і потребують більш ґрунтовного дослідження. Дослідження показало, що сервіс ChatGPT є дуже гнучким інструментом, який можна адаптувати до потреб конкретного користувача або компанії.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1.Лужецький В.А., Войтович О.П., Кожухівський В.Д. Основи інформаційної безпеки. Навчальний посібник [рекомендований МОН]. Вінниця: ВНТУ, 2013. 246 с.

2.Securityweek. Microsoft Puts ChatGPT to Work on Automating Cybersecurity. URL: https://www.securityweek.com/microsoft-puts-chatgpt-to-work-on-automating-cybersecurity/ (дата звернення 02.04.2023)

3. Kupershtein L., Martyniuk T., Voitovych O., Borusevych A. Remote Host Operation System Type Detection Based on Machine Learning Approach. *CEUR Workshop Proceedings*. 2021. Vol. 3106, pp. 65 – 81. URL: https://ceur-ws.org/Vol-3106/Paper_7.pdf (date of access: 03.04.2023).

4. Martyniuk T., Kupershtein L., Krukivskyi B., Lukichov V. Neural network model of heteroassociative memory for the classification task. *Radioelectronic and computer systems*. 2022. No. 2. P. 108–117. URL: https://doi.org/10.32620/reks.2022.2.09 (date of access: 03.04.2023).

5.ChatGPT. OpenAI. URL: https://chat.openai.com/chat. (дата звернення 03.04.2023)

6.Запорожець, О. Машинне навчання в кібербезпеці: проблеми та перспективи // Системні дослідження та інформаційні технології. 2022. 75-84 с.

7.Language models are unsupervised multitask learners. OpenAI Blog URL: https://d4mucfpksywv.cloudfront.net/better-language-models/language-models.pdf (дата звернення 05.04.2023)

Примаков Богдан Сергійович — студент групи 1БС-21МС, Факультет інформаційних технологій та комп'ютерної інженерії, Вінницький національний технічний університет, Вінниця, e-mail primakov.bogdan@gmail.com

Куперштейн Леонід Михайлович – к. т. н., доцент кафедри захисту інформації, Вінницький національний технічний університет, м. Вінниця, email: kupershtein@vntu.edu.ua

Primakov Bogdan S. — student 1BS-21MC, Faculty of Information Technologies and Computer Engineering, VinnytsiaNational Technical University, Vinnytsia, e-mail primakov.bogdan@gmail.com

Kupershtein Leonid— PhD (Eng), Associated Professor of Information Protection Department, Vinnytsia National Technical University, Vinnytsia, Ukraine, email: kupershtein@vntu.edu.ua