Література:

1. Operating System Market Share Worldwide. URL: https://gs.statcounter.com/os-market-share
2. Android version history. URL: https://en.wikipedia.org/wiki/Android_version_history
3. iOS version history. URL: https://en.wikipedia.org/wiki/IOS_version_history
4. The Six Most Popular Cross-Platform App Development Frameworks. URL: https://www.jetbrains.com/help/kotlin-multiplatform-dev/cross-platform-frameworks.html
5. React Native Guide. URL: https://reactnative.dev
6. Flutter Documentation. URL: https://flutter.dev
7. Kotlin Multiplatform. URL: https://kotlinlang.org/docs/multiplatform.html

DOI https://doi.org/10.36059/978-966-397-479-8-129

РОЛЬ ШТУЧНОГО ІНТЕЛЕКТУ В РОЗПІЗНАВАННІ І ПРОТИДІЇ КІБЕРАТАКАМ З ВИКОРИСТАННЯМ СОЦІАЛЬНОЇ ІНЖЕНЕРІЇ

Ковальчук Д. А.
здобувач вищої освіти третього (освітньо-наукового) рівня
за спеціальністю 125 – Кібербезпека
Міжнародний гуманітарний університет
Науковий керівник: ***Йона Л. Г.***
кандидат технічних наук, доцент,
доцент кафедри комп'ютерної інженерії та інноваційних технологій
Міжнародний гуманітарний університет
м. Одеса, Україна

Соціальна інженерія стала потужним інструментом у руках зловмисників, які використовують психологічні слабкості людей для отримання доступу до конфіденційної інформації. Зі збільшенням складності та масштабів таких атак зростає потреба в інноваційних засобах захисту. Штучний інтелект має значний потенціал для виявлення та запобігання соціально-інженерним атакам завдяки здатності аналізувати великі обсяги даних, автоматично виявляти аномалії та адаптуватися до нових загроз.

Методи виявлення соціально-інженерних атак за допомогою ШІ

1. Машинне навчання для аналізу поведінкових патернів

Машинне навчання (ML) забезпечує аналіз та ідентифікацію шаблонів поведінки користувачів, які можуть вказувати на спроби соціальної інженерії. Наприклад, раптова зміна часу доступу до корпоративної системи, незвичний обсяг завантажених даних або поведінка, яка відхиляється від типового профілю користувача, може сигналізувати про потенційну загрозу. У таких випадках алгоритми ML відстежують ці аномалії, попереджаючи безпекові команди про можливі ризики. Особливо ефективним є поєднання ML із системами управління доступом, які автоматично блокують потенційно шкідливі дії.

2. Обробка природної мови (NLP)

NLP дозволяє виявляти ознаки шахрайства в текстових повідомленнях, електронних листах та інших формах комунікації. Наприклад, системи на основі NLP можуть аналізувати емоційний тон тексту, частоту вживання імперативних форм або маніпулятивних фраз. Аналізуючи фішингові повідомлення, ШІ визначає, чи містять вони характерні маркери, як-от погрози, терміновість або надмірну емоційність. Такі моделі можуть бути інтегровані у фільтри електронної пошти для автоматичного блокування підозрілих повідомлень.

3. Генеративний ШІ для моделювання атак

Використання генеративних моделей, таких як GPT, дозволяє створювати сценарії потенційних атак на основі даних про попередні інциденти. Це корисно для тренування систем кіберзахисту та персоналу, забезпечуючи підготовку до реальних загроз. Наприклад, згенеровані ШІ сценарії фішингових атак допомагають працівникам розпізнавати ознаки маніпуляцій і вдосконалювати свої навички кіберграмотності.

4. Адаптивні моделі виявлення загроз

ШІ дозволяє створювати системи, які постійно вдосконалюються завдяки зворотному зв'язку. Наприклад, якщо новий тип атаки виявлено в одній організації, адаптивна система вивчає її характеристики й автоматично оновлює захисні алгоритми для інших клієнтів. Це забезпечує глобальний проактивний підхід до боротьби з соціальною інженерією.

Переваги використання ШІ у протидії соціальній інженерії

1. Автоматизація та швидкість реагування

Традиційні підходи до захисту часто залежать від ручної перевірки підозрілих дій, що може займати багато часу і ресурсів. Завдяки ШІ автоматизуються такі завдання, як моніторинг систем, виявлення загроз і реагування на інциденти. Наприклад, у випадку фішингової атаки система на основі пояснювального ШІ миттєво ідентифікує підозрілий лист і блокує/ізолює його ще до того, як користувач встигне його

прочитати. Це значно зменшує час реагування та обмежує можливості зловмисників.

2. Мінімізація людського фактору

Люди залишаються найуразливішим елементом будь-якої системи безпеки, оскільки вони можуть піддаватись емоційному чи психологічному впливу. ШІ не тільки автоматизує процеси, але й допомагає навчати працівників через інтерактивні симуляції атак. Наприклад, за допомогою генеративного ШІ можна створювати персоналізовані фішингові сценарії, щоб працівники краще розуміли, як уникати маніпуляцій.

3. Проактивне виявлення загроз

На відміну від традиційних систем безпеки, які часто реагують вже після факту атаки, ШІ дозволяє прогнозувати потенційні загрози. Завдяки аналізу великих обсягів даних і поведінкових патернів ШІ моделі можуть передбачати вірогідність соціальних атак на основі історичних даних та контексту поточних взаємодій. Це дозволяє організаціям діяти на випередження, вживаючи заходів до того, як загроза стане реальною.

4. Покращення взаємодії з користувачами

Системи ШІ дозволяють забезпечити інтуїтивно зрозумілу взаємодію для кінцевих користувачів. Наприклад, ШІ може автоматично проводити освітні модулі після виявлення помилок у поведінці працівника, підвищуючи загальну кіберграмотність у компанії.

5. Економія ресурсів

Інтеграція ШІ в кіберзахист дозволяє організаціям знижувати витрати на ручний моніторинг і розслідування. Завдяки автоматизації організації можуть ефективніше розподіляти свої ресурси, концентруючись на стратегічних завданнях.

Таким чином, ШІ не лише забезпечує багаторівневий захист від соціально-інженерних атак, але й трансформує підходи до кібербезпеки, роблячи їх більш гнучкими, адаптивними та ефективними.

Штучний інтелект відкриває нові горизонти у сфері кібербезпеки, зокрема у боротьбі з атаками, що використовують методи соціальної інженерії. Завдяки машинному навчанню, обробці природної мови та генеративним моделям, ШІ здатний аналізувати та виявляти аномалії в поведінці користувачів, автоматизувати захист і прогнозувати нові загрози. Це значно підвищує швидкість реагування на атаки та знижує ризики, пов’язані з людським фактором.

Проте, незважаючи на значний потенціал ШІ, його впровадження супроводжується викликами, такими як етичні аспекти, прозорість прийняття рішень та захист даних. Для забезпечення максимальної ефективності таких систем важливо дотримуватися балансу між

автоматизацією та контролем з боку людини, розвивати кіберграмотність користувачів і впроваджувати інтерактивні навчальні програми.

Компромісним варіантом може бути розробка програмного забезпечення, яке використовує пояснювальний ШІ для аналізу вхідних і вихідних електронних листів, дзвінків, повідомлень тощо, з подальшим ранжуванням цих даних за рівнем ризику надання несанкціонованого доступу кіберзлочинцям або витоку чутливої інформації. Якщо відсоток ризику перевищує визначений поріг, лист, повідомлення або запис дзвінка автоматично переміщується в карантин для подальшого аналізу відповідним ІТ-відділом.

Такий підхід дозволяє автоматизувати пошук вразливостей у системі, знижуючи ймовірність успішного здійснення соціально-інженерних кібератак. При цьому працівникам не потрібно постійно проходити нові тренінги з кібербезпеки, що зменшує витрати часу і ресурсів, зберігаючи високий рівень захисту.

Використання ШІ в кібербезпеці є перспективним напрямом, який дозволяє створювати проактивні та адаптивні стратегії захисту. Це робить його невід'ємним елементом сучасних цифрових екосистем і фундаментом для подальшого розвитку систем протидії соціальній інженерії.

Література:

1. Gray J. Practical Social Engineering: A Primer for the Ethical Hacker. 2022.

2. Gururaj H. L., Janhavi V., Ambika V. Social Engineering in Cyber-security: Threats and Defenses. 2024.

3. Hadnagy C. Human Hacking: Win Friends, Influence People, and Leave Them Better Off for Having Met You. 2021.

4. Sipola T., Kokkonen T., Karjalainen M. Artificial Intelligence and Cybersecurity: Theory and Applications. 1st ed. 2023.

5. Sarker H. I. AI-Driven Cybersecurity and Threat Intelligence: Cyber Automation, Intelligent Decision-Making and Explainability. 2024.