зосереджувалося на досвіді працівників експертних установ. Зазначене обумовлює актуальність дослідження даної проблеми у майбутньому.

1. Edwards, D., & Burnard, P. (2003). A systematic review of stress and stress management interventions for mental health nurses. *Journal of Advanced Nursing*, *42*, 169–200.

2. Sutherland, V., & Cooper, C. (1990). *Understanding stress: A psychological perspective for health care professionals.* London: Chapman & Hall.

3. Maslach, C., Jackson, S. E., & Leiter, M. P. (1996). Maslach Burnout Inventory manual (3rd ed.). Palo Alto: Consulting Psychologists Press, Inc.

4. Duquette, A., Kerovac, S., Sandhu, B. K., & Beauatt, L. (1994). Factors related to nursing burnout: A review of empirical knowledge. *Issues in Mental Health Nursing*, *15*, 337–358.

5. Schaufeli, W., & Peeters, M. (2000). Job stress and burnout amongst correctional officers: A review of the literature. International Journal of Stress Management, 7, 19–48.

Олена ГОРУН,
головний науковий співробітник
Українського науково-дослідного інституту
спеціальної техніки та судових експертиз
Служби безпеки України

АКТУАЛЬНІ ПРОБЛЕМИ ЗАХИСТУ ПРАВ ЛЮДИНИ У КІБЕРПРОСТОРІ В УМОВАХ ТЕХНОЛОГІЧНОГО РОЗВИТКУ СИСТЕМ ШТУЧНОГО ІНТЕЛЕКТУ

Кібербезпека напряму пов’язана із стрімким розвитком інтернет технологій, сервісів та додатків. Чат-боти зі штучним інтелектом, які колись вважалися просто автоматизованими розмовними програмами від тепер можуть навчатися та вести розмови, які майже не відрізняються від людських. Швидкий розвиток штучного інтелекту за останні роки призвів до появи вражаючих технологій чат-ботів. Ці віртуальні помічники на основі штучного інтелекту стають все більш популярними завдяки своїй здатності навчатися і надавати персоналізовану допомогу в різних доменах. Під час використання чат-боту важливим є його інтеграція з існуючими системами та процесами. Це включає підключення чат-бота до системи управління та наявність відповідного програмного забезпечення, яке надає йому здатність отримувати доступ до необхідної інформації. Окрім інтеграції чат–ботів з системами, важливо відстежувати ключові показники його ефективності (KPI). Однак разом з очевидними перевагами використання новітніх технологій виникають актуальні проблеми, що створює та провокує суттєві ризики, які доцільно розкрити.

Ризики дискримінації. Проблеми, пов’язані з дискримінацією можуть виникати по-різному, коли використовуються системи штучного інтелекту.

Однією з найбільших небезпек чат-ботів (ШІ) є їх схильність до шкідливих упереджень, тобто можуть виникнути упередженість даних, на яких навчаються інструменти ШІ. Оскільки моделі штучного інтелекту створюються людьми та навчаються, поглинаючи дані, цілком логічно, що людські упередження можуть бути вбудовані в дизайнерську модель, розробку, впровадження та використання ШІ. Як наслідок, чат-боти можуть швидко навчитися та поширювати протиправний або дискримінаційний контент, навіть якщо нічого такого не було в його початкових даних. Зачасту зловмисники можуть цілеспрямовано маніпулювати системами штучного інтелекту та чат-ботами для отримання упереджених результатів.

Ризики для кібербезпеки. У цьому сегменті чат-боти створюють ризики за двома основними напрямками. По-перше, зловмисники без складних навичок та вмінь програмування можуть використовувати чат-боти для створення шкідливих програм з метою кіберзломів. По-друге, оскільки чат-боти можуть переконливо імітувати вільну розмовну англійську, їх можна використовувати для фіктивного створення людських розмов, які можуть використовуватися для соціальної інженерії, фішингу та зловмисних рекламних схем. Це вимагає зусиль з метою забезпечення кібербезпеки. При цьому, наслідки впливу штучного інтелекту на кібербезпеку можуть бути катастрофічними. Небезпека технології чат-ботів штучного інтелекту також може становити більш пряму загрозу кібербезпеці, оскільки рівень фішингових атак постійно зростає. В контексті викладеного хакерами та зловмисниками активно практикується новітня методика під назвою «отруєння даних» (DNS cache poisoning) – це випадки, коли хакери успішно передають дані ШІ для створення уразливостей. Причинами, чому отруєння даних є ефективним, полягає в тому, що воно використовує недостатню обізнаність штучного інтелекту. Тобто отруєння даних – це прототип кібератаки, спрямованої безпосередньо на технології штучного інтелекту. Для здійснення атаки використовується вразливість у конфігурації DNS [1]. Хоча компанії, які спеціалізуються на технологіях штучного інтелекту, зазвичай, зберігають у таємниці джерела своїх даних, кіберзловмисники можуть визначити, які з них вони використовують, і маніпулювати даними. Хакери знаходять способи підробки наборів даних, які використовуються для навчання штучного інтелекту, дозволяючи їм маніпулювати своїми рішеннями та відповідями. Вони можуть маніпулювати програмами чат-ботів, щоб надавати неправдиву та фальсифіковану інформацію клієнтам, яка може змусити їх натиснути посилання, що містить зловмисне програмне забезпечення або шахрайський веб-сайт. У випадку, коли штучний інтелект починає витягувати зіпсовані дані, його важко виявити, і це може призвести до значного порушення стану кібербезпеки, що залишається непоміченим протягом тривалого часу. У разі здійснення кібератаки штучний інтелект має вирішальне значення, оскільки він може допомогти швидше реагувати на інциденти, аналізуючи дані в реальному часі та надаючи рекомендації щодо

запобіганню протиправних дій. Адже попри позитивні аспекти існує реальна загроза через здатність чат-бота, керованого штучним інтелектом необережно допомагати кіберзлочинцям писати шкідливий код. Саме тому найважливішим захистом від отруєння даних є надійно побудовані архітектура та інфраструктура кібербезпеки.

Ризики для конфіденційних даних. Також існує реальний ризик для забезпечення безпеки конфіденційних даних, оскільки чат-боти можуть на постійній основі збирати особисту та публічну інформацію, тривалий час зберігати її. Конфіденційність даних може бути порушена, коли йдеться про додатки, керовані саме штучним інтелектом. Програми, керовані ШІ призначені для збору та аналізу великих масивів даних, та існує ризик того, що ці дані можуть бути використані неналежним чином, або передані третім особам. Крім того, програми, керовані ШІ не можуть гарантувати повний захист даних, оскільки алгоритми ШІ можуть бути вразливими до маніпуляцій. Тобто однією з важливих спроможностей чат-ботів є можливість збирати інформацію про користувачів, відслідковувати їхні дії, а потім, за потреби, проаналізувати їхні звички, при цьому зібрані дані про користувачів дозволяють персоналізувати пропозиції і розсилку [2, с.69].

Наприклад, чат-бот «ChatGPT» збирає інформацію про IP-адресу користувача, тип браузеру та налаштування, дані про взаємодію користувача із певним сайтом та алгоритм дій користувача у веб-переглядачі протягом певного часу та на різних веб-сайтах, усіма даними, якими він може ділитися «третім особам». Особливістю є те, що у випадку, якщо користувач не надасть таку особисту інформацію, це може вірогідно призвести до непрацездатності послуг чат-бота. Також цілком вірогідно, що чат-бот «ChatGPT» може розкривати особисту інформацію реальних користувачів зі своїх навчальних даних. Тому, щоб зменшити ризики конфіденційності даних, компанії, які використовують чат-боти та генеративні інструменти штучного інтелекту, повинні постійно переглядати політику конфіденційності та розкриття інформації, дотримуватися законодавства про захист персональних даних щодо обробки особистої інформації, вживати заходів для захисту даних, контролювати, щоб конфіденційна інформація не передається через сторонні застосунки. Оскільки використання генеративного ШІ продовжує швидко зростати у світових масштабах, важливо визначити пріоритетність заходів кібербезпеки для захисту конфіденційних даних. Таким чином, алгоритми штучного інтелекту здатні отримувати персональну інформацію про людей шляхом аналізу великих даних, вилучати її з метаданих. Збираючи інформацію за допомогою штучного інтелекту про певну людину, власник алгоритму – компанія, державна організація або правоохоронні органи можуть з високим ступенем точності виявити уподобання, пріоритети та характерні властивості тієї чи іншої особи.

Ризики дезінформації. Чат-боти можуть допомагати зловмисникам швидко створювати неправдиву або фальсифіковану інформацію, що лунає

досить авторитетно. Тобто чат-бот може поширювати дезінформацію і відображати упередженість. Це пояснюється тим, що бот «вчиться» на інформації з реального світу, в якій існують такі упередження. Саме з цієї причини у відповідях на запитання користувачів можуть з'явитися стереотипи і хибна інформація (припущення). Чат-боти можуть писати новинні статті, есе та сценарії, які поширюють теорії змови, згладжуючи людські помилки, як-от поганий синтаксис і неправильний переклад, і просуваючись за межі легко виявлених завдань копіювання та вставки. Неправдиві наративи, що поширюються в мережі Інтернет регулярно шкодять не тільки бізнесу, але й державним інтересам. Більше того, зловмисники можуть навчити моделі штучного інтелекту фальшивою інформацією, вводячи в їхні моделі брехню, яку потім поширюють відповідні моделі. Управління ризиками дезінформації є досить складним процесом.

З метою запобігання та боротьби з дезінформацією, Європейська комісія ініціювала підписання за участі 44 компаній та організацій, серед яких «Google», «Facebook», «Microsoft» Кодекс практики щодо онлайн-дезінформації з пропозицією маркувати контент, створений за допомогою штучного інтелекту [3]. Підписанти, які інтегрують генеративний штучний інтелект у свої сервіси, такі як «Bing», «Chat» в Microsoft і «Bard» у Google, повинні створити необхідні гарантії, щоб ці сервіси не могли використовуватися зловмисниками для створення дезінформації. Законодавці в ЄС прагнуть маркувати діпфейки та інший контент, створений за допомогою ШІ, щоб звичайні користувачі відразу могли зрозуміти, що це створено саме автоматизованою машиною, а не людиною.

На жаль, маніпулятивні повідомлення, згенеровані за допомогою ШІ, удосконалилися настільки, що їх усе важче відрізнити від реальних. Зловмисники, зацікавлені в дезінформації за допомогою ШІ можуть повністю автоматизувати як її генерування, так і її поширення, що є головною небезпекою, яку несе ця технологія, і є основним викликом технічного прогресу. Узагальнюючи вищевикладене доцільно вказати, що масштабне використання чат-ботів на основі технологій штучного інтелекту тісно пов`язано із настанням потенційних ризиків, що вимагає прискорення розробки відповідних заходів їхньому запобіганню. Чат-боти зі штучним інтелектом відтепер можуть навчатися та вести розмови, які майже не відрізняються від людських, що відкриває нову еру технологічної революції. Однак небезпеки чат-ботів штучного інтелекту настільки ж різноманітні, що вимагають, у першу чергу: певної обережності під час використання чат-ботів і генеративного штучного інтелекту; чіткого розуміння того, що чат-боти можуть також робити помилки; постійно підвищувати захист кібербезпеки ІКТ систем від загроз, пов`язаних із штучним інтелектом; мати переконання, що штучний інтелект використовується відповідно до етичних норм та відповідних професійних стандартів; перевіряти результати штучного інтелекту на предмет їхнього упередженого та дискримінаційного впливу;

визначати алгоритми протидії дезінформації, яка може бути поширена за допомогою систем штучного інтелекту тощо.

1. DNS cache poisoning URL: https://owasp.org/www-pdf-archive/DNS_Cache_Poisoning(OWASP_GHANA).pdf.

2. Трофименко О.Г. Сфери застосування чат-ботів // Інформаційне суспільство: проблеми та перспективи: Матеріали VII Всеукраїнської науково-практичної конференції (м. Одеса, 20 травня 2022 р.). Одеса, 2022. С. 68-71. URL: http://dspace.onua.edu.ua/bitstream/handle/11300/18203.

3. The 2022 Code of Practice on Disinformation EU URL: https://digital-strategy.ec.europa.eu/en/policies/code-practice-disinformation.

Ігор ЧОБОТЬКО,
старший викладач кафедри фізичного виховання та тактико-спеціальної підготовки Дніпропетровського державного університету внутрішніх справ

ВПЛИВ УМОВ ВВЕДЕННЯ ПРАВОВОГО РЕЖИМУ ВОЄННОГО СТАНУ НА ДІЯЛЬНІСТЬ НАЦІОНАЛЬНОЇ ПОЛІЦІЇ В УКРАЇНІ

Умови введення правового режиму воєнного стану відображаються на діяльності органів Національної поліції відповідно до чинного законодавства і мають свої особливості. Основними нормативними актами, які регулюють функціонування поліції в цей період, є Закон України «Про Національну поліцію» і Закон України «Про правовий режим воєнного стану». Успішність діяльності поліції в цей період суттєво впливає на забезпечення громадського порядку [1].

Згідно зі статтею 1 Закону України «Про правовий режим воєнного стану», воєнний стан встановлюється в разі збройної агресії, загрози нападу або інших ситуацій, які становлять загрозу державній незалежності, територіальній цілісності України [2]. В цей період відповідним органам і органам місцевого самоврядування надаються додаткові повноваження для забезпечення національної безпеки та обмеження прав і свобод громадян.

Організація та матеріально-технічне забезпечення поліції враховує оперативну обстановку в кожному регіоні. Ефективність діяльності поліції в умовах воєнного стану досягається завдяки чіткому плануванню заходів, який регулярно відпрацьовується особовим складом поліції на практичних заняттях. Ці плани дозволяють ефективно використовувати час керівництва поліції для управлінських рішень в умовах кризових ситуацій [3].

Керівництво поліції в період воєнного стану повинно вживати заходів, визначених планом підготовки, щоб забезпечити громадський порядок і