УДК 004.89
М. М. Рокош
(Тернопільський національний технічний університет імені Івана Пулюя, Україна)

ЗАСТОСУВАННЯ ГЕНЕРАТИВНОГО ШТУЧНОГО ІНТЕЛЕКТУ ДЛЯ ЗАХИСТУ КОРПОРАТИВНИХ КОМУНІКАЦІЙ

UDC 004.89
M. Rokosh

APPLICATION OF GENERATIVE ARTIFICIAL INTELLIGENCE FOR THE PROTECTION OF CORPORATE COMMUNICATIONS

Генеративний штучний інтелект (ГШІ) пропонує передові рішення для забезпечення безпеки корпоративних комунікацій, які є критично важливими для збереження інформаційної цілісності сучасних організацій. ГШІ дозволяє автоматизувати процеси виявлення та запобігання кіберзагрозам, таким як фішингові атаки, витоки даних та спроби соціальної інженерії. Це досягається завдяки інтеграції технологій обробки природної мови (NLP) та моделей глибокого навчання, здатних аналізувати зміст та контекст повідомлень. Armorblox є прикладом ефективного застосування ГШІ для захисту електронної пошти та інших каналів корпоративної комунікації. Використовуючи лексичний, синтаксичний та семантичний аналіз тексту, платформа автоматично виявляє потенційно небезпечні повідомлення, ідентифікує неправомірні запити доступу до даних та запобігає випадковим витокам інформації. Такий підхід мінімізує залежність від людського фактору, автоматизуючи виявлення загроз та значно знижуючи ризик помилок, спричинених недосвідченістю або недбалістю співробітників.

На технічному рівні Armorblox використовує моделі на основі трансформерів, такі як GPT (Generative Pre-trained Transformer) або їх модифіковані варіанти. Ці моделі проходять попереднє навчання на великих наборах даних, а потім донавчаються з урахуванням специфічних потреб у сфері кібербезпеки. У процесі розробки та впровадження систем захисту комунікацій на основі ШІ ці інструменти створюють багатовимірні представлення даних, що дозволяє ідентифікувати приховані зв’язки між контекстом повідомлення, його відправником і отримувачем. Наприклад, платформа може виявити підозрілий запит на зміну банківських реквізитів від імені фінансового директора, якщо аналіз історії комунікацій покаже аномалії у стилі письма або характері запиту.

Інтеграція генеративного ШІ у захист корпоративних комунікацій охоплює кілька технічних етапів. Спершу система здійснює попередню обробку даних, аналізуючи їх, видаляючи шум та нормалізуючи текст, наприклад, виправляючи помилки або стандартизуючи форматування. Далі система генерує контекстуальні вектори для кожного повідомлення, використовуючи архітектури трансформерів, такі як BERT чи GPT, що дозволяє враховувати як значення слів, так і їх контекст у загальному потоці комунікацій. Наступним етапом є виявлення аномалій за допомогою моделей машинного навчання, таких як LSTM (Long Short-Term Memory) чи GRU (Gated Recurrent Units), які відстежують поведінкові патерни, виявляючи аномалії у часі, форматі або змісті повідомлень, які можуть свідчити про загрозу. Нарешті, алгоритми ГШІ приймають рішення та здійснюють реагування, використовуючи механізми раннього попередження для автоматичного блокування підозрілих повідомлень або передачі їх на перевірку відповідальним працівникам. Наприклад, якщо система виявляє лист, схожий на фішингову атаку, вона ізолює його та сповіщає користувача, пояснюючи ризики.

Розгортання та оцінка систем безпеки на основі ШІ часто передбачають тестування у контрольованих умовах, таких як імітація фішингових кампаній або інших кіберзагроз, що дозволяє вдосконалити функціональні можливості системи та забезпечити її надійність. Такі симуляції допомагають оцінити здатність моделі виявляти нові загрози та адаптувати алгоритми до змін у патернах атак. Крім того, масштабованість і надійність цих систем забезпечуються за допомогою передових фреймворків розгортання, включаючи інструменти контейнеризації, такі як Docker і Kubernetes, які підтримують обробку в реальному часі та моніторинг роботи системи у хмарних або гібридних середовищах.

Реалізація таких рішень демонструє значні переваги у зниженні ризиків кіберзагроз, оптимізації витрат на безпеку та покращенні управління потоками комунікацій. Проте важливо вирішувати питання конфіденційності даних, етичного використання технологій та адаптації персоналу до роботи з новими інструментами. Збалансоване поєднання технічних та організаційних заходів дозволяє максимально ефективно використовувати потенціал генеративного ШІ у корпоративній кібербезпеці.

Література

1. Dhoni P., Ravinder K. Synergizing Generative Artificial Intelligence and Cybersecurity: Roles of Generative Artificial Intelligence Entities, Companies, Agencies and Government in Enhancing Cybersecurity, Authorea Preprints, Journal of Global Research in Computer Sciences, 2023. doi:10.4172/2229-371X.14.3.005.
2. Krishnamurthy O. Enhancing Cyber Security Enhancement Through Generative AI, International Journal of Universal Science and Engineering 9, pp. 35-50, 2023. URL: https://ijuse.org/admin1/upload/06%20Oku%20Krishnamurthy%2001155.pdf.
3. Devlin J., Chang M.-W., Lee K., Toutanova K. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. doi:10.48550/arXiv.1810.