ШТУЧНИЙ ІНТЕЛЕКТ ЯК КЛЮЧ ДО БЕЗПЕКИ ЕЛЕКТРОННОЇ ПОШТИ

Шедін Д.А., Снігуров А.В.

Харківський національний університет радіоелектроніки, Україна

E-mail: dmytro.shedin@nure.ua,
arkadii.snihurov@nure.ua

Abstract

Email remains at the forefront of human-to-human communication. According to the statistics, the total daily number of letters is increasing. At the same time, threats spreading through this channel are also intensifying. Therefore, the purpose of the work is to improve the previously created system that is used to increase the level of electronic communication security by analyzing the sender's domain and email headers. For this purpose, the implementation of artificial intelligence (is proposed. This work also defines the analysis of the capabilities of natural language processing models and their advantages over manual research, the technical stack, and the architecture of the future system.

Електронна пошта наразі залишається популярним та важливим засобом комунікації. За підрахунками, у 2023 році кількість надісланих та отриманих листів у середньому на день сягає 347,3 мільярда. До 2026 року прогнозується, що ця цифра зросте до 392,5 мільярда [1]. Як наслідок, цей канал зв'язку найчастіше застосовується кіберзлочинцями для проведення атак. Статистика також показує, що у 2022 році 54% усіх загроз були поширені саме електронною поштою [2]. Отже, чинні засоби захисту не завжди є ефективними для протидії ним, а питання створення додаткових рішень для забезпечення безпеки є надзвичайно актуальним.

Більшість атак спрямовані безпосередньо на кінцевого користувача, який не завжди може визначити шкідливі листи. Одним із додаткових рішень для забезпечення безпеки цього каналу зв'язку є розроблене автором при написанні кваліфікаційної роботи освітнього ступеня «Бакалавр» програмне забезпечення «Adressant», яке шляхом криміналістичного аналізу домену відправника та заголовків повідомлення може вказувати на потенційні вразливості одразу ж у поштовому клієнті [3, 4]. Проте, застосунок не завжди є ефективним при великій кількості листів (бо кожне окремо треба перевіряти) та все ж потребує прийняття рішення від самого користувача. Як наслідок, – може спрацьовувати «людський фактор».

Штучний інтелект (ШІ) – це технологія, яка може допомогти з цим, бо дозволяє автоматизовано перевіряти безліч листів та визначати їх безпечність без участі користувачів. Обробка природної мови (NLP) належить до галузі ШІ та займається наданням комп'ютерам здатності розуміти текст і вимовлені слова майже так само, як це можуть робити люди. Ось кілька прикладів, як ця технологія може бути застосована при електронній комунікації:

- аналіз мови: системи ШІ можуть використовувати аналіз мови для виявлення ознак шкідливих листів, таких як використання певної лексики або граматики;
- аналіз поведінки: NLP-моделі можуть використовувати аналіз тексту для виявлення аномалій у поведінці відправників, що може бути ознакою шкідливих листів;
- аналіз контенту: ШІ може використовуватись для виявлення шкідливих файлів або посилань, що можуть бути приховані в листах;
- аналіз настрою: NLP-моделі можуть аналізувати настрої вхідних листів. Це є особливо корисним для визначення пріоритетів відправлень, чи загалом стану адресанта;

- категоризація електронної пошти: засоби ШІ можуть автоматично класифікувати електронні листи за папками або мітками, що полегшує їх керування. Наприклад, групувати на «особисті», «робочі», «інформаційні», «спам» тощо;
- пропозиції відповіді на лист: ШІ може пропонувати відповіді на електронні листи на основі вмісту та контексту повідомлення;
- виявлення аномалій: NLP-моделі можуть позначати незвичайні шаблони електронної пошти. Наприклад, у випадках раптового збільшення обсягу листів або незвичну поведінку раніше відомого відправника;
- автоматичне блокування: засоби ШІ можуть автоматично блокувати повідомлення, щоб воно не потрапило до користувача.

Під час написання кваліфікаційної роботи освітнього ступеня «Магістр» автором розробляються механізми розширення поточних можливостей системи «Adressant» та розробляється власна NLP-модель на базі Stanford CoreNLP [4] та з використанням мови програмування Java – «Adressant AI». Система дозволить отримувати лінгвістичні анотації для тексту, включаючи межі лексем і речень, частини мови, іменовані об'єкти, числові та часові значення, аналізи залежностей і конституентів, кореференцію, настрої, атрибуції цитат і зв'язки. Наразі модель підтримує 8 мов: арабську, китайську, англійську, французьку, німецьку, угорську, італійську та іспанську. Але в майбутньому планується їх розширення.

Загальна ж схема роботи системи представлена на рисунку 1.

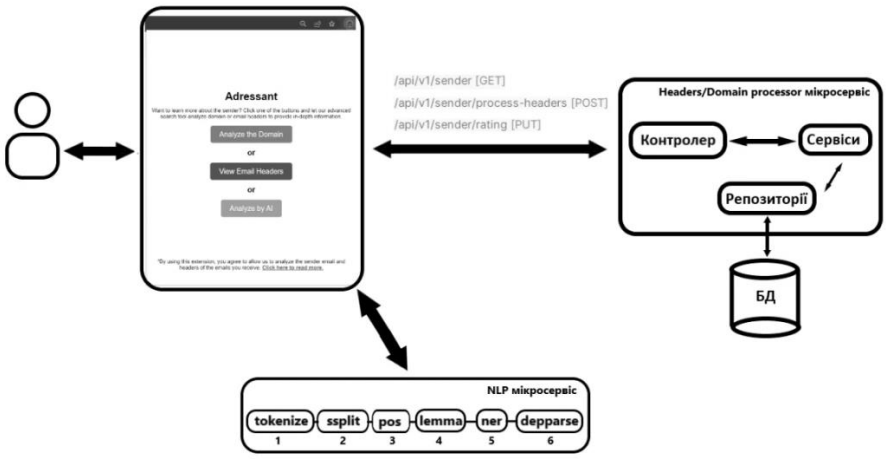

Рис. 1. Схема роботи системи «Adressant AI»

Отже, роль ШІ в захисті електронної комунікації є надзвичайно важливою, а майбутня розробка на базі NLP-моделі є перспективною та може застосовуватися в будь-якій сфері діяльності та користувачами з різним рівнем цифрової грамотності.

Література

1. Must-Know Email Statistics and Trends for 2023 (October 2023). *Mailbutler*. URL: https://www.mailbutler.io/blog/email/email-statistics-trends/ (дата звернення: 01.11.2023).

2. Worldwide 2022 Email Phishing Statistics and Examples. *Trend Micro*. URL: https://www.trendmicro.com/en_us/ciso/23/e/worldwide-email-phishing-stats-examples-2023.html (дата звернення: 01.11.2023).

3. Шедін Д.А. Підхід до цифрового криміналістичного аналізу повідомлень електронної пошти: кваліфікац. робота бакалавра / Шедін Дмитро Андрійович – Харків, ХНУРЕ, 2023. – 43 с.

4. Шедін Д.А. Програмне забезпечення для цифрового криміналістичного аналізу повідомлень електронної пошти. Радіоелектроніка та молодь у XXI столітті: каталог виставки технічної творчості молоді, м. Харків, 10–12 трав. 2023 р. / Харків. нац. ун-т радіоелектроніки. 1 с.

5. Overview. *CoreNLP*. URL: https://stanfordnlp.github.io/CoreNLP/ (дата звернення: 01.11.2023).