УДК 004

ФОРМУВАННЯ СЕМАНТИЧНОЇ МАПИ КІБЕР ЗАГРОЗ ІЗ ЗАСТОСУВАННЯМ ШТУЧНОГО ІНТЕЛЕКТУ

О. О. Гуменюк[1], І. М. Свобода[1], А. В. Комар[1], Д. В. Ланде[1]

[1] Навчально-науковий Фізико-технічний інститут

Анотація

Формування семантичної мапи кібер загроз із застосуванням штучного інтелекту представляє новітній підхід для оперативного виявлення потенційних загроз в реальному часі. Розроблена система інтегрує алгоритми для збору, аналізу та класифікації даних, що забезпечує комплексний аналіз кіберзагроз і візуалізацію їх динаміки. Особливість методу полягає у виявленні зв'язків між кіберінцидентами та аналізі їх характеристик і трендів розвитку.

Ключові слова: кібербезпека, штучний інтелект, соціальні медіа, моніторинг, машинне навчання, обробка природної мови, telegram

Вступ

На сучасному етапі розвитку інформаційних технологій, кібербезпека набуває особливого значення у зв'язку зі зростанням числа кіберзагроз, що еволюціонують і стають все складнішими. Соціальні медіа, особливо платформа Telegram, відіграють ключову роль у поширенні інформації [1], яка може включати шкідливі повідомлення та кібератаки. Значний обсяг даних, генерований користувачами, вимагає нових методів та технологій, заснованих на штучному інтелекті, машинному навчанні та обробці природної мови, для швидкого та точного аналізу. Це дослідження зосереджене на розробці та застосуванні передових технік для аналізу потенційних загроз кібербезпеки через моніторинг соціальних медіа, з акцентом на платформі Telegram, яка, з огляду на її широке використання та великий обсяг обміну інформацією, стає цінним джерелом для ідентифікації кіберзагроз. Це дозволяє оперативно реагувати на них та забезпечувати вищий рівень захисту інформації.

1. Розробка методології

Аналіз кіберзагроз на основі соціальних медіа включає декілька етапів: збір даних, виявлення подій, ідентифікація оригінальних подій, встановлення причинно-наслідкових зв'язків, аналіз і візуалізація мережі подій, а також кластеризація [2]. Нижче наведено детальний опис кожного з цих етапів.

Збір даних. Збір новинних повідомлень здійснюється за допомогою існуючих систем пошуку новин, які можуть бути як безкоштовними, так і платними. Відібрані повідомлення фільтруються за темою дослідження та очищуються від шуму і незначущих даних. Для цього процесу використовують наступну формулу:

$$F_{\text{news}}(t,k,C) = \left\{ \text{articles} \mid \begin{array}{l} \text{articles.date} \in [t_{\text{start}}, t_{\text{end}}] \wedge \\ \text{keywords(articles)} \cap k \neq \varnothing \wedge \\ \text{articles.channel} \in C \end{array} \right\} \tag{1}$$

де:

- t – часові параметри (дата початку та кінця),
- k – набір ключових слів,
- C – набір ідентифікаторів каналів.

Виявлення подій. Для виявлення подій у текстах новинних повідомлень застосовуються генеративні мовні моделі, такі як GenAI [3]. Цей процес включає формування масиву коротких позначень виявлених подій. Використовується наступна формула:

$$E_{\text{detect}}(\text{text}) = \{e_1, e_2, \ldots, e_n \mid e_i \in \text{extract\_events(text)}\} \tag{2}$$

де функція *extract_events* використовує методи NLP для витягування подій із тексту.

Встановлення причинно-наслідкових зв'язків. Встановлення причинно-наслідкових зв'язків між подіями здійснюється за допомогою генеративних моделей, які аналізують події та визначають їхні взаємозв'язки. Для цього використовується формула:

$$C_{\text{link}}(E) = \{(e_i, e_j) \mid \text{cause}(e_i, e_j)\} \tag{3}$$

де функція *causecause* визначає пари подій, де подія e_i є причиною події e_j.

Аналіз та візуалізація мережі подій. Мережа подій аналізується для виявлення ключових подій, кластерів та ланцюгів, що їх зв'язують. Інтерактивна візуалізація мережі подій здійснюється за допомогою наступної формули:

$$S_{\text{map}}(E, L) = \{(e, \text{link}(e)) \mid e \in E \wedge \text{link}(e) \in L\}, \tag{4}$$

де кожна подія e пов'язана з URL або документальним посиланням $link(e)$.

Кластеризація подій. Для кластеризації подій застосовується метод максимізації модулярності. Модулярність (Q) визначається як міра сили поділу мережі на модулі (або кластери):

$$Q = \frac{1}{2m} \sum_{ij} \left(S_{ij} - \left[S_{ij} - \frac{k_i k_j}{2m} \right] \delta(C_i, C_j) \right), \quad (5)$$

де:

- $S_i j$— елемент матриці схожості,
- k_i і k_j – суми ваг ребер, прикріплених до вузлів i і j,
- m — сума всіх ваг у мережі,
- C_i ,C_j – спільноти вузлів i та j
- δ – дельта Кронекера, який дорівнює 1, якщо i і j, знаходяться в одній спільноті, і 0 в іншому випадку.

Для оцінки якості кластеризації можна використати коефіцієнт силуету, який розраховується за формулою:

$$s(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}} \quad (6)$$

де:

- $a(i)$ – середня відстань між i та всіма іншими точками у кластері, до якого належить i,
- $b(i)$ – мінімальна середня відстань від i до всіх точок у будь-якому іншому кластері, членом якого i не є.

Коефіцієнт силуету:

$$s(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}}. \quad (7)$$

Ця формула визначає коефіцієнт силуету для кожної точки i у кластерізованій мережі подій. Вона використовує середню відстань до точок у тому ж кластері $a(i)$ та найменшу середню відстань до точок у сусідніх кластерах $b(i)$ для оцінки якості кластеризації.

Візуалізація. Результати аналізу візуалізуються за допомогою інтерактивних дашбордів, які демонструють зв'язки між подіями, динаміку змін загроз у часі та географічний розподіл кіберінцидентів [4]. Це дозволяє користувачам системи швидко оцінювати рівень загрози та відповідно реагувати.

2. Реалізація системи та приклад застосування методології

Система призначена для забезпечення ефективної обробки даних у реальному часі з використанням масштабованої та надійної архітектури. Вона складається з кількох ключових компонентів, кожен з яких відіграє важливу роль у процесі збору, обробки та аналізу інформації.

Компоненти системи:

- *Модуль збору даних*: використовує API для постійного збору інформації з платформи Telegram. Цей модуль оснащений фільтрами, які налаштовуються відповідно до специфічних параметрів, таких як ключові слова, діапазони дат та частота подій, що дозволяє збирати лише релевантну інформацію для подальшого аналізу.
- *Модуль обробки даних*: використовує алгоритми машинного навчання та інструменти обробки природної мови для аналізу текстової інформації. Він видаляє шум та нерелевантні дані, класифікує вміст за рівнями загроз та видобуває ключові тематичні елементи.
- *Управління базами даних*: відповідає за зберігання оброблених даних у структурованому форматі у безпечній, масштабованій базі даних, що дозволяє швидко отримувати доступ до інформації та ефективно нею управляти.
- *Інтеграція з платформами кібербезпеки*: система розроблена з урахуванням безперервної інтеграції з існуючими платформами моніторингу кібербезпеки через добре визначені API та рішення проміжного програмного забезпечення, що підвищує ефективність існуючих систем реагування на інциденти [5].
- *Інтерфейс користувача*: реалізований з використанням Flutter, дозволяє ефективно взаємодіяти з системою як технічним, так і нетехнічним користувачам. Інтерактивні дашборди та налаштовувані перегляди надають можливість користувачам отримувати необхідну інформацію відповідно до їхніх потреб.

Інновації в аналізі даних. Система використовує генеративні моделі штучного інтелекту для поліпшення інтерпретації і розпізнавання даних, а також розширює можливості для ефективного виявлення та реагування на кіберзагрози [6]. Планується подальше оновлення та інтеграція новітніх технологій для забезпечення актуальності та конкурентоспроможності у боротьбі з сучасними кібервикликами.

2.1. Приклад застосування методології

Застосуємо нашу методологію для аналізу новин за ключовим словом «vulnerability».

1. Збір даних: Використовуючи системи пошуку новин, збираються новини, що відповідають заданим параметрам часу, ключовим словам та каналам.

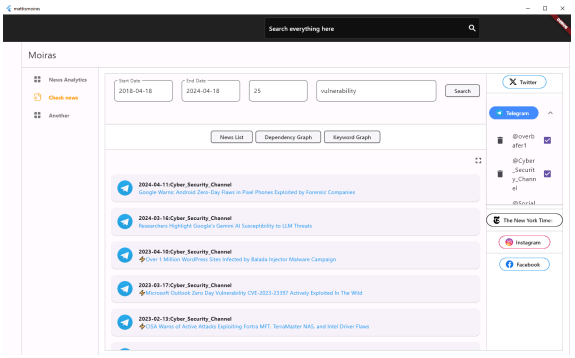

Рис. 1. Скріншот роботи клієнтського застосунку для введених параметрів

2. Виявлення подій: Зібрані новини аналізуються для виявлення подій.
3. Ідентифікація оригінальних подій: Ідентифікуються оригінальні події серед виявлених та вилучаються дублікати.
4. Встановлення причинно-наслідкових зв'язків: Встановлюються причинно-наслідкові зв'язки між подіями.

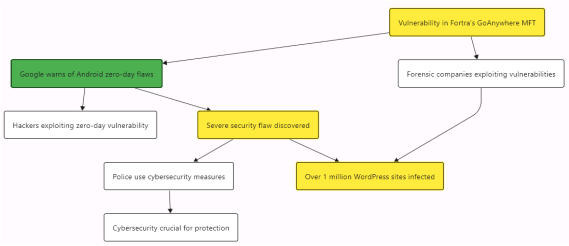

Рис. 2. Карта зв'язку між подіями у вигляді графу

5. Аналіз та візуалізація мережі подій: Формується мережа подій та створюється інтерактивна візуалізація, використовуючи формулу (5), після чого здійснюється побудова графів, що відображають зв'язки між цими подіями.
6. Кластеризація подій: Події кластеризуються для виявлення модулів у мережі за допомогою максимізації модулярності, використовуючи формулу (6). Кластеризація перевіряється за допомогою коефіцієнта силуету, використовуючи (7).

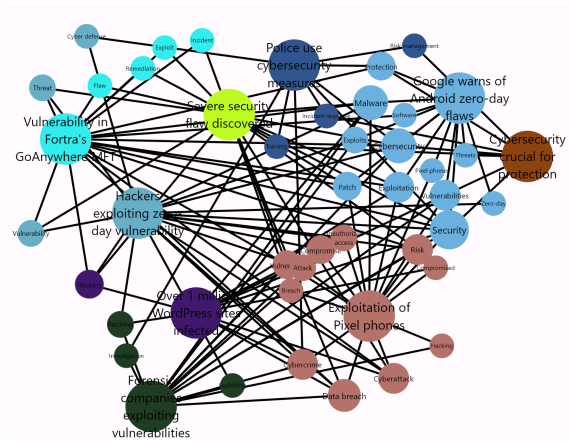

Рис. 3. Граф ключових слів до запиту «vulnerability»

Висновки

У результаті проведеного дослідження було розроблено нові методи аналізу та моніторингу кіберінцидентів через соціальні медіа, що дозволяють автоматизовано виявляти, ідентифікувати та аналізувати кіберзагрози, а також візуалізувати їх у реальному часі. Автоматизація цього процесу значно зменшує час та ресурси, необхідні для аналізу великих обсягів даних, дозволяючи оперативно досліджувати динаміку загроз та ефективно відреагувати на них. Запропонована методологія є гнучкою і адаптується до різних платформ соціальних медіа, що робить її універсальним інструментом для дослідження кіберзагроз. Використання штучного інтелекту сприяє виявленню нових тенденцій та причинно-наслідкових зв'язків між кіберінцидентами, що відкриває нові можливості для прогнозування майбутніх загроз та розробки ефективних стратегій кіберзахисту. Перспективи розвитку методології включають розширення джерел даних, розробку складніших інструментів візуалізації, використання машинного навчання для прогнозування атак, а також інтеграцію навчальних модулів для підвищення обізнаності у кібербезпеці.

Перелік використаних джерел

1. Detection and resolution of rumours in social media: A survey / A. Zubiaga, A. Aker, K. Bontcheva, M. Liakata, P. R // Natural Hazards and Earth System Sciences. — 2018. — Feb. — Vol. 21. — P. 10. — URL: https://doi.org/10.1145/3161603.
2. *Kruspe A.*, *Kersten J.*, *Klan F.* Review article: Detection of actionable tweets in crisis events // ACM Computing Surveys. — 2021. — June. — Vol. 51. — P. 20. — URL: https://doi.org/10.5194/nhess-21-1825-2021.
3. *Lande D.*, *Strashnoy L.* GPT Semantic Networking: A Dream of the Semantic Web - The Time is Now // Engineering. — 2023. — Vol. 2131. — P. 168. — URL: https://ssrn.com/abstract=4541673.
4. *Cherven K.* Mastering Gephi Network Visualization // Packt Publishing. — 2015. — P. 378. — URL: http://gephi.michalnovak.eu/Mastering%20Gephi%20Network%20Visualization.pdf.
5. Cyber-Attack Features for Detecting Cyber Threat Incidents from Online News / M. Abdullah, A. Zainal, M. Maarof, M. Nizam // Proceedings of the 2018 Cyber Resilience Conference. — 2018. — URL: https://doi.org/10.1109/CR.2018.8626866.
6. *Munkhdorj B.*, *Yuji S.* Cyber attack prediction using social data analysis // Journal of High Speed Networks. — 2017. — Vol. 23. — P. 26. — URL: https://doi.org/10.3233/JHS-170560.