

УДК 004.7

*Адамчук Р.С., здобувач
Дячук О.Ю., ст. викладач
Колощук М.С., ст. викладач
Державний університет «Житомирська політехніка»*

МЕТОДИ АВТОМАТИЗОВАНОГО ЗБОРУ ТА АНАЛІЗУ ДАНИХ З TELEGRAM-СПІЛЬНОТ У ЗАДАЧАХ OSINT

У роботі досліджено методи автоматизованого збору та аналізу інформації з відкритих Telegram-спільнот у контексті OSINT. Систематизовано підходи до отримання даних із використанням клієнтського доступу до Telegram через звичайний акаунт користувача, а також допоміжних методів web-scraping. Проаналізовано методи обробки текстового контенту та запропоновано узагальнену функціональну модель системи моніторингу для задач приватного бізнесу.

Зростання обсягів відкритих даних у цифровому середовищі зумовлює необхідність їх ефективного аналізу. Концепція OSINT передбачає використання відкритих джерел для отримання аналітичної інформації [7]. Telegram як інформаційна платформа характеризується високою швидкістю поширення повідомлень і значною кількістю тематичних спільнот.

Автоматизований збір даних реалізується із використанням клієнтського доступу до Telegram через звичайний обліковий запис користувача, що дозволяє отримувати повідомлення з відкритих спільнот без необхідності використання офіційного API. Такий підхід забезпечує більшу гнучкість доступу до чатів, зокрема у випадках, коли використання ботів є обмеженим. Методи web-scraping можуть застосовуватися як допоміжний інструмент за умови дотримання політик платформи.

Офіційні інструменти Telegram API можуть використовуватися як альтернативний підхід для організації збору даних [1], [2].

Обробка даних здійснюється із використанням методів обробки природної мови [3], що включають нормалізацію, токенизацію, лематизацію та векторизацію тексту. Для підвищення точності аналізу можуть застосовуватися сучасні трансформерні моделі [4].

Важливим аспектом функціонування системи є попередня фільтрація інформації та визначення її релевантності. Для цього можуть застосовуватися методи ранжування повідомлень за ключовими словами, частотними характеристиками та семантичною близькістю. Це

дозволяє зменшити обсяг нерелевантних даних і підвищити ефективність подальшого аналізу.

Крім того, у системах OSINT-моніторингу доцільним є використання механізмів автоматичного виявлення аномалій, що дозволяють ідентифікувати нетипові інформаційні патерни або різкі зміни активності у спільнотах. Такі підходи можуть базуватися на статистичних методах або моделях машинного навчання та сприяють своєчасному виявленню потенційних загроз.

Практичні підходи до використання Telegram як джерела даних підтверджуються сучасними дослідженнями у сфері кіберрозвідки, зокрема при формуванні датасетів та виявленні загроз [5], [6].

Функціональна модель системи включає модулі збору, обробки, зберігання та аналізу інформації. Взаємодія між модулями забезпечує обробку інформації у квазі-реальному часі та масштабованість системи. Обмеженням підходу є залежність від доступності даних і політик Telegram.

У роботі систематизовано методи збору та аналізу даних із Telegram-спільнот. Встановлено, що ефективним є комбінований підхід із використанням API та методів NLP. Наукова новизна полягає у формалізації такого підходу для задач OSINT. Оцінювання ефективності може здійснюватися за метриками precision, recall та F1-score. Практична значущість полягає у застосуванні результатів для підтримки прийняття рішень у бізнесі.

Список використаних джерел

1. Telegram. Telegram APIs / TDLib. URL: <https://core.telegram.org>
2. Telegram. Telegram Bot API. URL: <https://core.telegram.org/bots/api>
3. Jurafsky D., Martin J. *Speech and Language Processing*. URL: <https://web.stanford.edu/~jurafsky/slp3/>
4. Devlin J. et al. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. URL: <https://arxiv.org/abs/1810.04805>
5. Arikkat D. R. et al. CTI Dataset Construction from Telegram. URL: <https://arxiv.org/pdf/2509.20943>
6. Saeed M. H. et al. SENTINEL: A Multi-Modal Early Detection Framework for Cyber Attacks. URL: <https://arxiv.org/pdf/2512.21380>
7. Van Puyvelde D., Tabárez F. The Rise of Open-Source Intelligence // *European Journal of International Security*. URL: <https://www.cambridge.org/core/journals/european-journal-of-international-security/article/rise-of-opensource-intelligence/21122432399ECB8078BF0D89A76D0586>