

УДК 004.56

Чорний П.Р., Скарга-Бандурова І. С., д.т.н., проф.

Тернопільський національний технічний університет ім. Івана Пулюя

**МЕТОДИ ІДЕНТИФІКАЦІЇ ТА ВИЛУЧЕННЯ ГЕОГРАФІЧНО ПОВ'ЯЗАНИХ
ОБ'ЄКТІВ З ДАНИХ ПРО АТАКИ НА ОБ'ЄКТИ КРИТИЧНОЇ
ІНФРАСТРУКТУРИ УКРАЇНИ**

P.R. Chorny, I.S. Skarga-Bandurova, DSc, Prof

**METHODS FOR IDENTIFYING AND EXTRACTING GEOGRAPHICALLY
RELATED OBJECTS FROM TEXTUAL DATA ON ATTACKS ON CRITICAL
INFRASTRUCTURE FACILITIES IN UKRAINE**

Кібератаки на об'єкти критичної інфраструктури в Україні стали однією з найактуальніших проблем національної безпеки. Для забезпечення ефективної протидії таким атакам необхідно точно та оперативно ідентифікувати та ліквідувати пов'язані з ними загрози. Виявлення географічно пов'язаних об'єктів з текстових та історичних даних спрощує аналіз та розуміння впливу кібератак у географічному контексті, що сприяє розробці ефективних стратегій реагування на інциденти та забезпечує захист кіберпростору.

Метою даного дослідження є розробка та застосування методів ідентифікації географічно пов'язаних об'єктів та їх взаємозв'язків. Це дозволить виявити та аналізувати закономірності між об'єктами.

Основними етапами роботи є:

1. Збір даних і попередня обробка. На цьому етапі здійснюється збір необхідних даних про об'єкти, їх географічні координати та атрибути, також попередня обробка;
2. Етап використання методів ідентифікації передбачає застосування різних методів для ідентифікації географічно пов'язаних об'єктів;
3. Геокодування. Процес полягає у перетворенні об'єктів, пов'язаних із розташуванням, у географічні координати за допомогою бібліотек геокодування;
4. Геопросторовий аналіз допомагає виявити розміщення об'єктів у просторі та знаходити закономірності їх розподілу;
5. На етапі візуалізації та аналізу даних географічно пов'язані об'єкти представляються в графічній формі.

На основі текстової інформації з відкритих джерел в роботі планується картографування щоденних подій із використанням технологій для обробки природньої мови та геопросторового аналізу. Для візуалізації обробленої інформації пропонується використання сучасних інструментів для створення графіків та картографічних візуалізацій.

Один із найважливіших технологічних компонентів методології полягає у використанні методів обробки природньої мови для ідентифікації іменованих сутностей, що є ключовим елементом аналізу текстових даних. За допомогою таких методів, іменовані сутності, зокрема імена, локації, дати, організації та інші, можуть бути виділені та класифіковані у текстових даних для подальшого використання в аналізі та визначенні ключових елементів тексту. Для роботи з іменованими сутностями використовуються завдання з бібліотек spaCy, NLTK та Stanford NER, в яких реалізовані відповідні алгоритми та класифікатори для виявлення іменованих сутностей у текстових даних.

В результаті дослідження очікується набуття всебічного розуміння складної природи атак на критичну інфраструктуру України (ККІ). Засновуючись на аналізі історичних даних та закономірностей фізичних та кібератак, проводиться ідентифікація критичних аспектів у захисті об'єктів ККІ і розробка ефективних стратегій захисту.