МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ

НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ
УКРАЇНИ "КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ
ІНСТИТУТ ІМЕНІ ІГОРЯ СІКОРСЬКОГО"

НАВЧАЛЬНО-НАУКОВИЙ
ФІЗИКО-ТЕХНІЧНИЙ ІНСТИТУТ

THEORETICAL AND APPLIED CYBERSECURITY

**Матеріали другої Всеукраїнської
науково-практичної конференції**

Випуск 2

Київ – 2024

Рекомендовано до друку Вченою радою
КПІ ім. Ігоря Сікорського
(протокол № 14 від 12 червня 2024 р.)

Theoretical and Applied Cybersecurity. Матеріали другої всеукраїнської науково-практичної конференції (TACS-2024). – Київ: Інжиніринг. – 190 с. ISBN 978-966-2344-98-1

До збірника увійшли матеріали доповідей, представлених на другій всеукраїнській науково-практичній конференції Theoretical and Applied Cybersecurity (TACS-2024, 30 травня 2024 року, м. Київ, Україна).

У збірнику представлені матеріали, присвячені питанням кібернетичної безпеки критичних інфраструктур, моделювання та протидії інформаційним операціям, технологій інформаційно-аналітичних досліджень на основі відкритих джерел інформації. Наведені матеріали з актуальних проблем інформаційної та кібернетичної безпеки, можливості застосування штучного інтелекту, системного аналізу при забезпеченні підтримки прийняття рішень, комп'ютерному моделюванні процесів і систем, актуальні завдання забезпечення інформаційної та кібербезпеки.

Для фахівців в області інформаційних технологій, інформаційної і кібернетичної безпеки, а також для аспірантів і студентів старших курсів вищої школи відповідних спеціальностей.

ISBN 978-966-2344-98-1

ФОРМУВАННЯ СЕМАНТИЧНОЇ МЕРЕЖІ КІБЕРЗАГРОЗ ШЛЯХОМ ОБРОБКИ ПРИРОДНОМОВНИХ ТЕКСТІВ

Анатолій Фегер, Дмитро Ланде

КПІ ім. Ігоря Сікорського, Київ, Україна

Ефективна аналітика кіберзагроз є важливою складовою сучасного кіберзахисту, зосередившись на використанні новітніх технологій штучного інтелекту, можна полегшити дослідження великих текстових масивів за допомогою контекстно-мережевого підходу. Це відкриває прискорені можливості для виділення загальних зв'язків і прихованих закономірностей між вилученими лінгвістичними сутностями. Запропоновано методологію забезпечення точності та повноти обробки NER з великого масиву новинних вирізок про кібератаки на Україну та висвітлено аналітичні можливості семантично-мережевого підходу.
Ключові слова: Cyber attacks, Semantic networks, Linguistics, GPT, NER

Вступ

У сфері кібербезпеки здатність швидко і точно ідентифікувати критичні елементи в неструктурованому тексті, такі як кіберзагрози і вразливості, має першорядне значення. Розпізнавання іменованих об'єктів, Named-Entity Recognition (NER) виділяється як важливий інструмент в обробці природної мови, Natural Language Processing (NLP), який полегшує вилучення структурованої інформації з текстів.

Це дослідження присвячене застосуванню методів NER, зокрема, використанню можливостей генеративного Штучного Інтелекту (ШІ) для обробки україномовних та російськомовних новинних статей про кібератаки проти України. Такий підхід дає змогу всебічно проаналізувати текстові корпуси, виокремити семантичні зв'язки та виявити релевантні об'єкти в контексті кожної новини.

На основі опрацьованих даних пропонується побудувати контекстну семантичну мережу, яка прояснює зв'язки та основні закономірності між об'єктами в текстовому просторі. Для підвищення точності та релевантності виокремлених об'єктів запропоновано додаткові процедурні рівні для підвищення точності та забезпечення повноти вихідних даних.

Методологія

Досліджуваний масив даних складається з новинних вирізок і статей, отриманих з відкритих джерел, що відображають кібератаки, пов'язані з Україною, за ключовими параметрами на україномовних та російськомовних ресурсах. Набір даних складається з 500 новинних вирізок, кожна з яких в середньому містить 976 слів, підготовлених як тестові документи для подальшого аналізу.

Для обробки було вилучено несуттєвий контекст, такий як дати, автори та посилання, а контент було розділено на 10 частин для зручності обробки. Кожна частина містила 50 вирізок, оброблених окремо ШІ з окремими запитами для забезпечення об'єктивності, без збереження попередньої історії запитів.

Дослідження використовує методи обробки природної мови для семантичного пошуку та вилучення зв'язків між сутностями [1]. Генеративний ШІ визначає еквіваленти в тексті, присвоює токенам мітки типів сутностей і визначає рівні зв'язків. Він також ітеративно заповнює контекстуальні прогалини. Для візуалізації семантичні пари сутностей нормалізуються та уніфікуються, щоб зменшити дискретність і фрагментарність даних.

Екстракція семантичних сутностей

Було використано описані техніки NER, як ключовий метод обробки природної мови - для ідентифікації та категоризації ключових елементів у просторі текстів та вилучення структурованої інформації з новинних вирізок. Обрано модель роду Generative Pre-trained Transformer

(GPT-4), оскільки вона продемонструвала більшу точність та ефективність при обробці україномовних новинних вирізок, в порівнянні з іншими моделями [2]. GPT-4 налаштовано на стандартній температурі 0,7 і використовує внутрішні алгоритми токенізації текстів, який вирішує проблему рідкісних слів, розбиваючи їх на підслова, які потім перетворюються на вектори високої розмірності, де фіксують як семантичні, так і синтаксичні деталі [3]. Архітектура GPT розширює його можливості, вбудовуючи лексеми контекстуально, враховуючи взаємозалежності між сусідніми лексемами.

Для підвищення ефективності NER було введено додатковий аналітичний рівень для вимірювання сили семантичних зв'язків між парами ідентифікованих сутностей. Цей рівень вимагає додаткового запиту в моделях GPT для оцінки якісних аспектів лінгвістичних зв'язків між сутностями в контексті новинної статті. Оцінка цих зв'язків включає низку внутрішніх методів GPT, від простих детермінованих підходів, таких як аналіз повторюваності слів, до більш складних методів, таких як оцінка косинусної схожості між векторами об'єктів для формування результату за запитом [4]. Ці методи інтегровані в подальшу мультимодальну стратегію яка відображає різноманітний характер запитів, промптів застосованих до визначеного набору даних.

Модель використовує ці методи для прогнозування рівня кореляції між вихідними сутностями, виводячи результати у форматі прикладу «Приватбанк; російські хакери; 55\%», де відсоток відображає оцінений лінгвістичний зв'язок за шкалою від 0 до 100. Цей підхід не лише ідентифікує сутності, а й аналізує їхні взаємозв'язки, тим самим підвищуючи точність даних, за впровадженням у подальшому відповідної фільтрації пар понять за цією ознакою, які демонструють мінімальну реляційну значущість.

Агрегація та нормалізація лексем

При обробці тексту великі мовні моделі, такі як GPT, генерують відповіді на основі розподілу ймовірностей, які

можуть здаватися правильними, але можуть містити шум та галюцинації. Таким чином, такі відповіді краще розглядати як рекомендацію віртуального експерта [5].

Для підвищення надійності кожен запит на розпізнавання об'єктів за NER дублюється, таке ітеративне злиття підказок і текстових просторів забезпечує кількісні варіації від декількох віртуальних експертів, створюючи більш повний набір даних. Описаний ітеративний метод покращує аналіз новинної інформації, інтерфейс дозволяє задавати значення кількості повторень запитів, що полегшує ітеративну генерацію відповідей у GPT-моделях.

Для забезпечення якісних та кількісних характеристик, стикаючись з проблемами синонімізації, та лінгвістичної невизначеності в лексемах виділяється висока варіативність сформованих взаємопов'язаних пар. Таке лінгвістичне розмаїття сильно ускладнює візуалізацію мережі, перешкоджаючи визначенню її ключових вузлів, центральності та категоризації, що має вирішальне значення для забезпечення аналітичних завдань.

Наприклад, «кібербезпека України» та «українська кібербезпека» які створюють додаткову сплутаність в побудованій мережі, тому методи лексичної нормалізації стандартизують термінологію для покращення структури мережі. Передові лінгвістичні моделі, такі як GPT, мають покращене контекстуальне розуміння простору текстів, точно ідентифікуючи та асоціюючи слова, вони підтримують автоматичну нормалізацію коренів слів, покращуючи оновлення наборів даних для використання.

Результати та обговорення

На основі вирізок новин про кібератаки на Україну було побудовано спрямовану семантичну мережу з використанням методологічно обґрунтованого процесу візуалізації та аналізу основних семантичних зв'язків у наборі даних за допомогою генеративного ШІ. Результатом цього процесу стало 50 семантичних пар на кожну стрічку новин загальною ємністю 25,000, які були відфільтровані за ступенем їх взаємозв'язку, залишивши по 20 релевантних пар на кожну новину. В результаті було оброблено

семантичних пар 10,000, з яких після нормалізації було отримано 1576, що сформували остаточний набір даних для аналізу.

На основі описаного процесу ми побудували репрезентативну візуалізацію, показану на Рис. 1, яка ефективно відображає взаємодію різних об'єктів у мережі. Розмір кожного вузла визначався довжиною імені об'єкта, а зв'язок між вузлами зображувався за допомогою кольорової схеми - від зеленого (низький зв'язок) до червоного (високий зв'язок), що ілюструє важливість і частоту взаємодії кожного об'єкта в мережі.

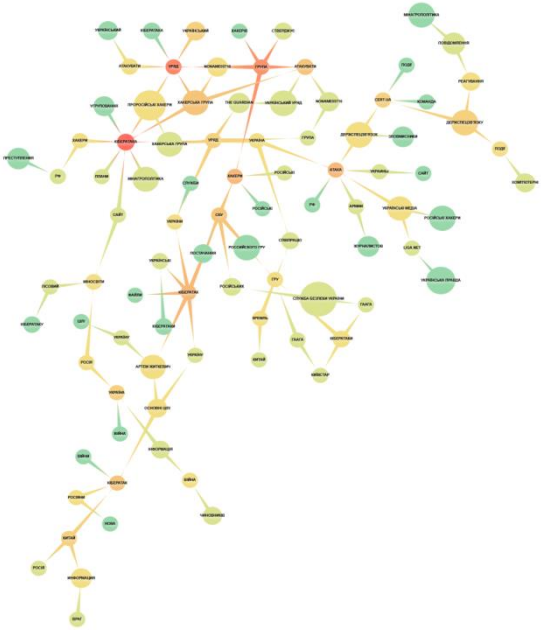

Рисунок 1 - Семантична мережа пар сутностей опрацьованих новинних видань

Подальший аналіз, включаючи оцінку центральності та ідентифікацію спільнот, дозволив глибше зрозуміти структуру мережі та ролі конкретних суб'єктів у ній. Аналіз

центральності визначив Україну, кібератаки та Міністерство освіти як центральні вузли, що мають вирішальне значення для підтримання узгодженості та потоків у мережі, як показано в Штаб. 1.

Близькість до центру продемонструвала, що такі організації, як Росія, Україна та Міністерство освіти, є не лише центральними, але й сильно впливають на інформацію в мережі, позиціонуючи їх як ключових акторів у реагуванні на інциденти, описані іншими вузлами. Крім того, використання методів виявлення спільнот ефективно ідентифікує групи в мережі, висвітлюючи різноманітні моделі взаємодії та зв'язки, починаючи від прямого впливу кібератак і закінчуючи ширшими геополітичними та освітніми впливами.

Таблиця 1. Значення типів центральності об’єктів в семантичній мережі

	Ступінь	Близькість
Україна	0.278	0.376
Росія	0.222	0.248
Кібератака	0.222	0.324
Мін-освіти	0.167	0.268
Китай	0.167	0.268

Висновки

Було використано передові методи обробки природної мови та генеративний ШІ для аналізу великого набору даних із вирізок новин про кібератаки, пов'язані з Україною. Використовуючи такі методи, як NER та побудова семантичних мереж, ми успішно виділили та проаналізували лінгвістичні зв'язки між текстовими об'єктами.

Описаний процес підвищує якість даних і розширює потенціал для виявлення ключових закономірностей і

зв'язків, які не є очевидними за допомогою традиційних методів.

Крім того, наша методологія демонструє ефективність мультимодальних підходів у дослідженнях кібербезпеки, де гнучкість і швидкість генеративного ШІ поглиблюють розуміння кіберзагроз та інформаційних війн.

Результати цієї роботи потенційно можуть бути використані для формування політики, посилення заходів безпеки та сприяти зміцненню глобальної кібербезпеки. Розширюючи межі аналізу тексту за допомогою штучного інтелекту, ця робота поглиблює наше розуміння кіберзагроз і створює основу для ефективної аналітики в галузі кібербезпеки та інших сферах.

Перелік використаних джерел

1. Dmytro Lande L. S. GPT Semantic Networking: A Dream of the Semantic Web - The Time is Now. — ISBN 978-966-2344-94-3. — 2023. — 168 p.
2. Koubaa A. GPT-4 vs. GPT-3.5: A Concise Showdown — 2023 — DOI: 10.20944/preprints2023 03.0422.v1.
3. Indurkhya N., Damerau F. J. Handbook of natural language processing. — CRC Press, 2010. — 2010. — 676 p.
4. Jianpeng Cheng L. D., Lapata M. Longshort-term memory-networks for machine reading // Conference on Empirical Methods in Natural LanguageProcess- ing. — 2016. — P. 551–561— DOI:10.48550/arXiv.1601.06733.
5. Dmitro Lande Anatolii Feher L. S. Cybersecurity in AI-Driven Casual Network Formation // Theoretical and Applied Cybersecurity. — 2023. — Vol. 5–Issue 2. — P. 105–113. — DOI:10.20535/tacs.2664-29132023.2.287139.