УДК 004.056/.57+339.986:004.91+316.472.4
DOI: https://doi.org/10.17721/ISTS.2024.8.34-41

Олександр ПУЧКОВ, канд. філос. наук, проф.
ORCID ID: 0000-0002-8585-1044
e-mail: iszzi@iszzi.kpi.ua
Національний технічний університет України
"Київський політехнічний інститут імені Ігоря Сікорського", Київ, Україна

Дмитро ЛАНДЕ, д-р техн. наук, проф.
ORCID ID: 0000-0003-3945-1178
e-mail: dwlande@gmail.com
Національний технічний університет України
"Київський політехнічний інститут імені Ігоря Сікорського", Київ, Україна

Ігор СУБАЧ, д-р техн. наук, проф.
ORCID ID: 0000-0002-9344-713X
e-mail: igor_subach@ukr.net
Національний технічний університет України
"Київський політехнічний інститут імені Ігоря Сікорського", Київ, Україна

МЕТОДИКИ ЕКСТРАГУВАННЯ ОБ'ЄКТІВ КІБЕРБЕЗПЕКИ З ЕЛЕКТРОННИХ ДЖЕРЕЛ ІЗ ЗАСТОСУВАННЯМ ШТУЧНОГО ІНТЕЛЕКТУ

Вступ. ***Стрімкий розвиток інформаційних технологій призвів до виникнення нових загроз і викликів у сфері кібербезпеки. Кібервійна стала реальністю та справжньою проблемою для держав, організацій та окремих користувачів кіберпростору. В Україні вживають заходи з розроблення системи кібердій у кіберпросторі, які включають сукупність взаємопов'язаних підсистем кіберрозвідки, кіберзахисту та кібервпливу. Однією з форм кіберрозвідки є розвідка з відкритих джерел – комп'ютерна розвідка (OSINT), яка здійснюється для пошуку й добування розвідувальної інформації, зокрема і для виявлення й аналізу об'єктів кібербезпеки щодо прогнозування можливих проявів кіберзагроз та їхніх наслідків. Це вимагає розроблення ефективних методів виявлення й аналізу об'єктів кібербезпеки за допомогою екстрагування фактографічних даних про об'єкти кібербезпеки з великих масивів неструктурованої текстової інформації.***

Методи. ***Досліджено технології штучного інтелекту, зокрема й великі мовні моделі (ВММ), та генеративного штучного інтелекту (ГШІ) в контексті застосування їх для розв'язання задач комп'ютерної розвідки об'єктів кібербезпеки з відкритих електронних джерел і соціальних мереж.***

Результати. ***У результаті проведеного дослідження, для здійснення дієвої аналітики результатів добування інформації, запропоновано методику екстрагування іменованих сутностей – назв хакерських угруповань та їхніх контекстуальних зв'язків із текстів повідомлень електронних мережних джерел, що стосуються предметної області кібербезпеки, а також формування мереж їхніх взаємозв'язків і змістовного аналізу цих мереж. Для визначення áкторів, які мають відношення до кібервійни, запропоновано методику аналізу відібраних документів, доступних в електронних джерелах інтернету та соціальних мережах. Обидві методики ґрунтуються на застосуванні ГШІ.***

Висновки. ***Результати дослідження демонструють ефективність запропонованих підходів і можливість їхнього застосування на практиці під час розв'язання завдань забезпечення кібербезпеки. Запропоновані методики можуть стати важливим інструментом для фахівців сфери кібербезпеки щодо розроблення ними ефективних стратегій захисту від кіберзагроз.***

Ключові слова: ***кібервійна, кібербезпека; генеративний штучний інтелект; великі мовні моделі; інтернет; відкриті електронні джерела; соціальні мережі, аналіз тексту; об'єкти кібервійни.***

Вступ

Стрімкий розвиток інформаційних технологій (ІТ) призвів до виникнення нових загроз і викликів у сфері кібербезпеки. Кібервійна стала реальністю та справжньою проблемою для держав, організацій та окремих користувачів інтернету. Кібервпливи все частіше стають ефективним інструментом для досягнення мети щодо контролю й управління як об'єктами критичної інфраструктури держави, так і окремо взятими громадянами та їхніми об'єднаннями (Даник, Воробієнко, & Чернега, 2019). Нині практично всі провідні держави світу зіткнулися з кіберзагрозами й необхідністю формувати системи кібербезпеки та кібероборони.

Залежність безпеки й економіки держави від стану важливих об'єктів та інформаційної інфраструктури об'єктів критичної інфраструктури, таких як: енергозабезпечення, водопостачання, транспорт, електронні комунікації тощо, обумовлює визнання кібербезпеки провідним елементом державної безпеки. Кіберзагрози у сучасному суспільстві набувають значного масштабу та являють собою наявні та потенційно можливі явища і чинники, що створюють небезпеку життєво важливим національним інтересам України у кіберпросторі та негативно впливають на стан кібербезпеки держави.

З огляду на це в Україні вживають заходів із розроблення системи кібердій у кіберпросторі, під якими розуміють сукупність взаємопов'язаних підсистем кіберрозвідки, кіберзахисту, кібервпливу та кіберконтррозвідки, які утворюють цілісну єдність, на яку покладаються функції із забезпечення кібербезпеки (Даник, Воробієнко, & Чернега, 2019).

Зауважимо, що нині розвідка перемістилася у новий вимір бойових дій **–** кіберпростір – і стала важливою складовою системи кібердій – кіберрозвідкою, під якою розуміють процес добування усіма наявними технічними засобами розвідки (космічної, повітряної, радіоелектронної, мережної, програмно-комп'ютерної, розвідки систем управління тощо) й засобами розвідки з відкритих джерел (OSINT) інформації, наявної в кіберпросторі про протилежну сторону, а також подальше її оброблення, що здійснюється за єдиним задумом і планом із метою викриття процесів управління, які протікають у кібернетичних системах під час їхнього функціонування та формування вихідних даних для здійснення заходів кіберзахисту та кібервпливу на фізичні, соціальні, інформаційні й інші кібернетичні системи (Даник, Воробієнко, & Чернега, 2019).

ISSN 2707-1758

Однією з форм кіберрозвідки є розвідка з відкритих джерел – комп'ютерна розвідка, яка здійснюється для пошуку й добування розвідувальної інформації, у тому числі, виявлення й аналіз об'єктів кібербезпеки для прогнозування можливих проявів кіберзагроз та їхніх наслідків. Основними джерелами отримання інформації для OSINT є відкриті електронні джерела в інтернеті, соціальних мережах і месенджерах.

Відповідно до цього, актуальним є наукове завдання щодо розроблення ефективних методів виявлення й аналізу об'єктів кібербезпеки екстрагуванням фактографічних даних про них із великих масивів текстової інформації.

Традиційні методи аналізу тексту, які використовують нині, мають свої обмеження, особливо коли йдеться про оброблення й аналіз великих за обсягом даних, що мають складну структуру.

Розв'язати цю проблему можна, застосувавши сучасні ІТ, такі як ВММ і ГШІ, які дозволяють ефективно обробляти й аналізувати текстові дані великого обсягу та складної структури.

Метою статті є розроблення нових методик екстрагування об'єктів кібербезпеки з великих за обсягом і складних за структурою текстових документів, розміщених у відкритих електронних джерелах інтернету та соціальних мережах із використанням ВММ та ГШІ для ефективного розв'язання задач комп'ютерної розвідки фахівцем із кібербезпеки для відпрацювання ним ефективних стратегій захисту від кіберзагроз.

Для досягнення мети дослідження розв'язувались такі часткові завдання:

1. Розроблення методики виявлення суб'єктів кібербезпеки засобами ГШІ.
2. Розроблення методики екстрагування а́кторів кібервійни з використанням ГШІ.

Суть першого завдання полягає у використанні системи ГШІ, наприклад, ChatGPT, для екстрагування понять і зв'язків між ними, шляхом звернення до неї зі змістовними запитами (промптами), та розробці алгоритму для формування мереж взаємозв'язків між суб'єктами кібербезпеки на основі екстрагованих даних.

Друга задача пов'язана з визначенням об'єктів кібербезпеки – а́кторів кібервійни на основі даних, отриманих у результаті інформаційно-пошукових запитів (ІПЗ) до агрегаторів інформації та формування промптів до ГШІ, та, відповідно до цього, розробити алгоритм для побудови мережі а́кторів у формі графа, який враховує взаємозв'язки між ними.

Огляд літератури. У науковій літературі широко обговорюють методи аналізу текстових даних. Наприклад, у статті (Yi et al., 2020) розглянуто модель RDF-CRF розпізнавання іменованих об'єктів безпеки на основі регулярних виразів і словника відомих об'єктів, а також умовних випадкових полів у поєднанні з чотирма шаблонами ознак. Ця модель, на основі правил, дозволяє здійснювати зіставлення об'єктів безпеки з необхідною точністю в простих ситуаціях. Словник відомих об'єктів застосовують для вилучення загальних і специфічних об'єктів безпеки, а екстрактор на основі CRF використовує ідентифіковані об'єкти за допомогою екстракторів на основі правил і словника для подальшого покращення продуктивності розпізнавання.

У роботі (Halbouni et al., 2022) наведено аналіз ефективності традиційних алгоритмів машинного навчання для розв'язування задач кібербезпеки. Проте наведені у статтях методи мають свої обмеження, особливо під час роботі з великими за обсягом та складною структурою даними. З іншого боку, у теперішній час спостерігається стрімкий розвиток ІТ, які ґрунтуються на досягненнях у сфері штучного інтелекту, зокрема і ВММ, і ГШІ, які призвані для подолання недоліків та обмежень, що властиві для традиційних методів оброблення текстових даних.

У статті (Bayer et al., 2024) розглянуто застосування ВММ для аналізу текстових даних у сфері кібербезпеки. Зокрема, запропоновано використання спеціально адаптованої до сфери кібербезпеки мовної моделі, яка може слугувати основою для розроблення систем кібербезпеки.

Наведено аналіз ефективності застосування ГШІ для розв'язування задач у сфері національної безпеки у праці Hassanin і Moustafa. У цій роботі досліджуються наслідки інтеграції ВММ, аналізується їхній потенціал щодо оброблення інформації, прийняття рішень і оперативності рішень, що приймаються. Проаналізовано переваги застосування ВММ, такі як автоматизація завдань і покращення аналізу даних, ризики і проблеми, пов'язані з конфіденційністю даних і їхньою вразливістю до атак з боку супротивника. Зроблено висновок про доцільність поєднання застосування ВММ із методами теорії прийняття рішень, що може значно полегшити перехід від даних до обґрунтованих рішень, дозволяючи особам, які приймають рішення, швидко отримувати й аналізувати наявну інформацію з меншими витратами людських ресурсів.

У статті (Gao, Zhang, & Han, 2021) обговорено застосування методу розпізнавання іменованих об'єктів для виявлення іменованих сутностей у текстах, що є важливим аспектом у виявленні об'єктів кібербезпеки. Описано різні підходи і методи для розпізнавання іменованих сутностей у сфері кібербезпеки, включаючи підхід на основі правил, підхід на основі словників і підхід на основі машинного навчання, а також проведено аналіз проблем, з якими стикаються дослідження в цій області. Запропоновано майбутні напрями розпізнавання іменованих сутностей у сфері кібербезпеки серед яких: застосування неконтрольованої або напівконтрольованої технології навчання; розробка більш повної онтології кібербезпеки; розробка більш повної моделі глибокого навчання.

Автори роботи Hanks et al. (2022) розглядають використання бібліотеки spaCy для аналізу текстових даних у кібербезпеці. Тут наведено початковий неструктурований корпус інформації, що описує вектори загроз, вразливості та кібератаки з різних відкритих джерел, який використовується для навчання та тестування моделей об'єктів кібербезпеки за допомогою фреймворку spaCy та вивчення методів самонавчання для автоматичного розпізнавання об'єктів кібербезпеки.

У статті (Alam et al., 2022) проаналізовано ефективність бібліотеки Flair у розпізнаванні іменованих сутностей. Описано бібліотеку CyNER, з відкритим вихідним кодом на мові python для розпізнавання іменованих об'єктів у сфері кібербезпеки, яка поєднує в собі трансформаційні моделі для вилучення об'єктів, пов'язаних із кібербезпекою, евристики для вилучення різних індикаторів компрометації, а також загальнодоступні моделі для загальних типів об'єктів.

Застосування методу тематичного моделювання для аналізу текстових даних, що належать до сфери кібербезпеки, обговорено в роботі (Piyush, & Okamura,

2021). Запропонований метод дозволяє виявляти тематичні зв'язки між різними об'єктами кібербезпеки,

У статті (Lande, Puchkov, & Subach, 2022) запропоновано інформаційну технологію вилучення концептів із текстів повідомлень мережних джерел, що належать до предметної області кібербезпеки. Ці концепти фільтруються за статистичними характеристиками та ранжуються. Розглянуто питання створення, кластеризації та візуалізації мережі концептів та їхніх взаємозв'язків.

У дослідженні (Lande, Puchkov, & Subach, 2020) запропоновано й обґрунтовано підходи до побудови системи моніторингу й аналізу соціальних медіа з питань кібербезпеки, які базуються на концепції оброблення великих обсягів даних, складних мереж, добування знань із текстових масивів. Основна ідея створення цієї системи – одночасне застосування методів і засобів інформаційного пошуку, аналізу даних та агрегування інформаційних потоків.

Питанням формування в реальному часі моделей предметних областей і дайджестів на основі автоматичного аналізу великої кількості повідомлень із соціальних мереж, присвячено публікацію (Lande et al., 2020]. Запропоновано метод кластерного аналізу, який базується на оцінюванні дискримінантного значення термів, новизна якого полягає у використанні найбільш значущих дискримінантних значень як центроїдів для визначення кластерів.

Проте проведений аналіз вказує на відсутність ефективних методів, які повною мірою задовольняли б потреби користувачів – фахівців сфери кібербезпеки щодо екстрагування об'єктів кібербезпеки з великих масивів текстових документів, що мають складну структуру.

Отже, зважаючи на те, що існуючі методи мають суттєві обмеження для роботи з великими за обсягом масивами текстових даних з одного боку, та сучасні досягнення у цій сфері та сфері штучного інтелекту, з іншого, стає доцільним і перспективним розроблення та застосування новітніх ІТ, які ґрунтуються на нових методах екстрагування об'єктів з електронних джерел і соціальних мереж, зокрема й у сфері кібербезпеки, на основі застосуванні ВММ і ГШІ.

Методи

Виділення іменних сутностей – об'єктів кібербезпеки є важливим для ідентифікації можливих кіберзагроз і розкриття кіберзлочинів, виявлення і розслідування діяльності злочинних хакерських угруповань тощо.

Результати виділення іменних сутностей можуть бути використані для подальшого їхнього аналізу у контексті кібербезпеки. Є кілька методів та інструментів, які можна використовувати для виокремлення іменних сутностей:

1. Аналіз вебсторінок:
   - використання інструментів для вилучення текстової інформації з вебсторінок відповідно до тег веброзмітки;
   - отримання метаданих (наприклад, EXIF-даних із фотографій), які можуть містити іменні сутності.
2. Аналіз соціальних мереж:
   - застосування інструментів для виявлення згадок конкретних імен, компаній або інших сутностей на публічних сторінках;
   - вивчення зв'язків між різними сутностями через аналіз друзів, фоловерів тощо.
3. Аналіз електронної пошти:
   - використання інструментів для аналізу публічної інформації, яка може містити імена, адреси, телефонні номери тощо.
4. Методи роботи з неструктурованим текстом:
   - використання шаблонів (регулярних виразів – regular expressions, RegExp);
   - застосування алгоритмів оброблення текстів природною мовою (Natural Language Processing, NLP) для виділення імен, назв організацій тощо з текстової інформації;
   - автоматична ідентифікація іменованих сутностей – Named Entity Recognizer (NER): spaCy, Natural Language Toolkit (NLTK), Stanford Named Entity Recognizer;
   - використання інструментів для розпізнавання іменованих сутностей (осіб, місць, організацій, а́кторів тощо) в тексті, що може вказувати на сталі словосполучення.
5. Аналіз коду:
   - аналіз вихідного коду програм (зокрема, коментарів) для виявлення імен та інших інформаційних сутностей, пов'язаних із розробниками чи компаніями сфери кібербезпеки.
6. Застосування ВММ і ГШІ:
   - формування спеціальних запитів (промптів) до систем ГШІ, таких як ChatGPT, Gemini, Groq тощо для виявлення відповідних іменованих сутностей сфери кібербезпеки.
   - Для виявлення іменованих сутностей можуть використовуватися такі інструментальні засоби OSINT:
   - бази даних кіберзлочинців, такі як SANS Internet Storm Center, The Honeynet Project тощо;
   - інструменти соціальних мереж, такі як Social Mention , BuzzSumo тощо;
   - системи аналізу журналів роботи систем, такі як Splunk, LogRhythm тощо;
   - інструменти OSINT, такі як Maltego, OSINT Framework тощо.

Методика виявлення суб'єктів кібербезпеки засобами генеративного штучного інтелекту. Для здійснення дієвої аналітики результатів добування інформації, запропоновано методику екстрагування іменованих сутностей – назв хакерських угруповань та їхніх контекстуальних зв'язків із текстів повідомлень електронних мережних джерел, що стосуються предметної сфери кібербезпеки, а також формування мереж їхніх взаємозв'язків і змістовного аналізу цих мереж засобами ГШІ, зокрема і системи ChatGPT, яка дозволяє отримувати результати змістовних запитів (промптів) через програмний інтерфейс (Application Programming Interface, API).

Екстрагування понять і зв'язків між ними здійснюється за допомогою звернення до системи ГШІ змістовними промптами, що належать до заздалегідь відібраних текстів повідомлень із вебпростору і соціальних мереж.

Суть методики така.

На першому кроці для отримання інформаційного масиву публікацій щодо кібербезпеки визначають необхідний період (наприклад, місяць до і місяць після початку події у кіберпросторі) й опрацьовують тематичні запити до системи OSINT, як-от:

– кібервійна в Україні:

- (кібератак~/3/украї)|(хакер~/3/атак~/2/украї)|(кибератак~/3/украин)| (хакер~/3/атак~/2/украин)
- (hack~/3/ukrain)|(cyber~attack~/3/ukrain)|(cyberattack~/3/ukrain)

На другому кроці для кожного з отриманих документів застосовують промпт до системи ChatGPT, результати якого надходять до програм через API й агрегуються для подальшого формування мереж (рис. 1).

ISSN 2707-1758

Надай список назв хакерських угруповань з тексту без пояснень у Вигляді переліку. Текст:...

Рис. 1. Приклад промпта до системи ChatCPT
для отримання списку назв хакерських угруповань

На третьому та четвертому кроці здійснюють аналіз відібраної мережі та її візуалізацію (рис. 2).

Формальна постановка задачі й алгоритм її розв'язання виглядає так.

Дано:

множина документів $D = \{d_1, d_2, ..., d_N\}$ – набір документів, отриманих за допомогою OSINT-систем на основі тематичних запитів; H – множина назв хакерських угруповань, які потрібно виявити з текстів документів; C – множина контекстуальних зв'язків між хакерськими угрупованнями, що екстрагуються з текстів документів.

Необхідно: відібрати з множини текстових документів D суб'єкти кібербезпеки та сформувати і візуалізувати мережу їхніх взаємозв'язків.

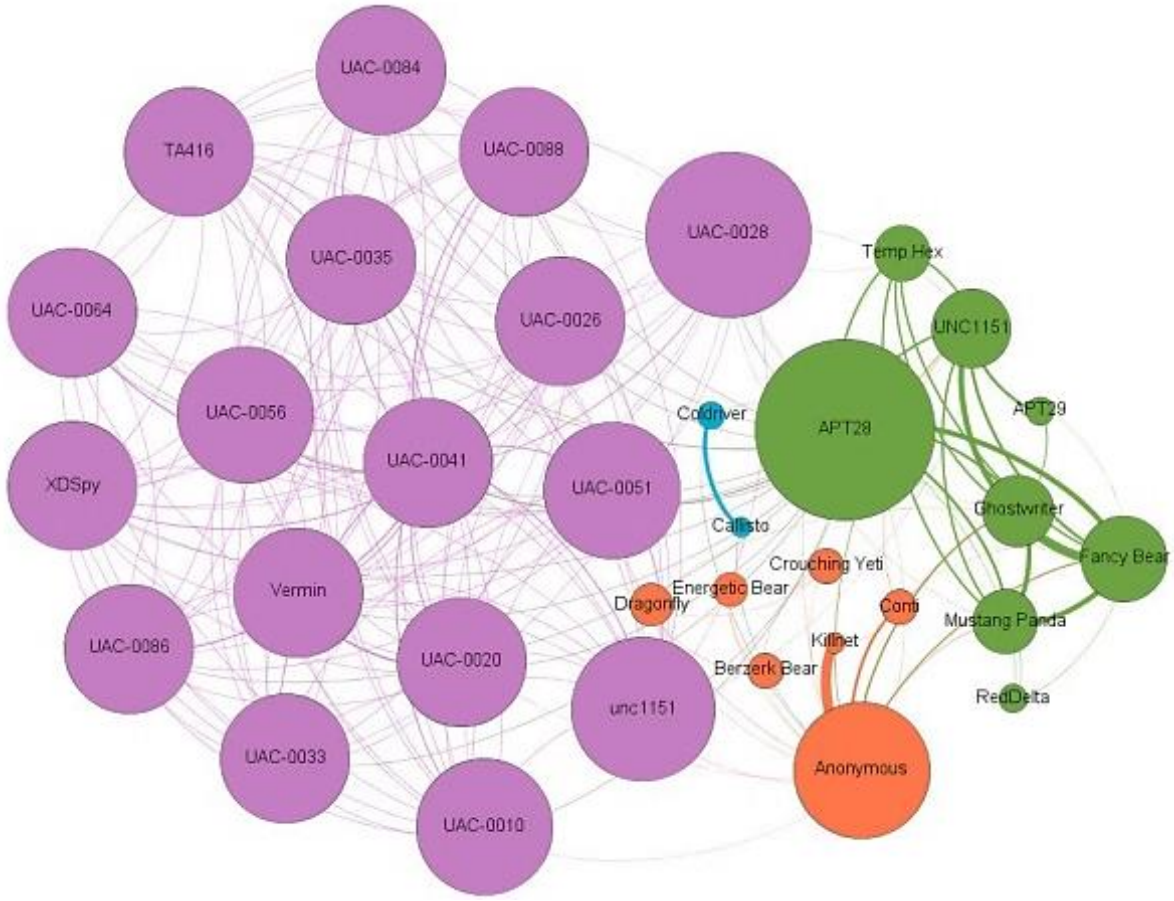

Рис.2. Фрагмент мережі злочинних хакерських угруповань,
пов'язаних із російсько-українською кібервійною

Алгоритм розв'язку

Крок 1. Для кожного набору тематичних запитів $q \in Q$ (напр., запитів щодо кібератак в Україні), отримати множину документів $D = \{d_1, d_2, ..., d_N\}$, що відповідають цим запитам:

$$D = \bigcup_{q \in Q} OSINT(q), \quad (1)$$

де $OSINT(q)$ – функція, що повертає множину документів за тематичним запитом q.

Крок 2. Екстрагування назв хакерських угруповань. Для кожного документа $d \in D$ формується відповідний промпт до системи ГШІ (напр., ChatGPT) для екстракції назв хакерських угруповань:

$$H(d) = GenAI(prompt, d), \quad (2)$$

де $H(d)$ – множина хакерських угруповань, екстрагованих із документа d; prompt – змістовний запит до системи GenAI.

Крок 3. Побудова мережі зв'язків. На основі екстрагованих назв хакерських угруповань, для кож-

ISSN 2707-1758

ного документа сформувати множину контекстуальних зв'язків:

$$C(d) = \{(h_i, h_j) \mid h_i, h_j \in H(d)\}, \quad (3)$$

де $C(d)$ – множина парних зв'язків між хакерськими угрупованнями з документа d.

Загальна множина зв'язків для всіх документів визначається як

$$C = \bigcup_{d \in D} C(d). \quad (4)$$

Крок 4. Візуалізація та аналіз мережі. Мережа зв'язків хакерських угруповань, побудована на основі множини H і множини зв'язків C, може бути представлена у вигляді графа $G = (H, C)$, де: H – множина вершин (хакерських угруповань); C – множина ребр (контекстуальних зв'язків між угрупованнями).

Відповідно до наведено вище, обчислювальну складність алгоритму можна оцінити, оцінивши обчислювальну складність його кроків.

Обчислювальна складність кроку формування інформаційного масиву публікацій залежить від кількості запитів Q та кількості документів N. Її можна оцінити як $O(|Q| \times N)$.

Обчислювальна складність кроку екстрагування назв хакерських угруповань оцінюється з таких міркувань: для кожного документа d здійснюється звернення до системи GenAI. Якщо t_{AI} – середній час оброблення одного запиту до системи ГШІ, то загальна обчислювальна складність цього етапу складатиме: $O(N \times t_{AI})$.

Обчислювальна складність побудови мережі зв'язків розраховується таким способом: для кожного документа d екстрагуються зв'язки між угрупованнями. Якщо в документі d знайдено $|H(d)|$ хакерських угруповань, то кількість зв'язків між ними оцінюється як $O(|H(d)|^2)$. Тоді загальна складність процесу побудови мережі буде така: $O(\sum_{d \in D} |H(d)|^2)$.

У свою чергу, обчислювальна складність кроку візуалізації та аналізу мережі залежить від кількості вершин $|H|$ та ребер $|C|$ у графі $G = (H, C)$. У найгіршому випадку, складність візуалізації та аналізу можна оцінити як $O(|H| + |C|)$.

З урахуванням наведеного, загальна складність алгоритму дорівнюватиме

$$O(|Q| \times N + N \times t_{GPT} + \sum_{d \in D} |H(d)|^2 + |H| + |C|). \quad (5)$$

Отже, можна зробити висновок про те, що запропонована методика й алгоритм її реалізації дозволяють ефективно екстрагувати й аналізувати взаємозв'язки між хакерськими угрупованнями на основі даних із текстових джерел, використовуючи засоби ГШІ.

Методика екстрагування а́кторів кібервійни засобами генеративного штучного інтелекту. Для визначення а́кторів, які мають відношення до кібервійни, запропоновано методику аналізу відібраних документів, доступних в електронних джерелах інтернету та соціальних мережах, шляхом застосування системи ГШІ.

Формально цю задачу можна сформулювати так.

Дано:

множина документів $D = \{d_1, d_2, ..., d_N\}$ – набір текстових документів, відібраних з електронних джерел інтернету та соціальних мереж;

$K = \{k_1, k_2, ..., k_m\}$ – набір ключових слів, обов'язкових для наявності їх у документі для подальшого аналізу; $N = \{(n_1, s_1), (n_2, s_2), ..., (n_p, s_p)\}$ – множина пар імен і прізвищ, екстрагованих із текстів електронних документів; $f_{\min}$ – мінімальна кількість поява імені та прізвища для збереження під час аналізу (мінімальний поріг частотності).

Необхідно: побудувати мережу а́кторів кібервійни та взаємозв'язків між ними.

Суть запропонованої методики розв'язання сформульованої задачі така.

На першому кроці методики формується запит до пошукової системи-агрегатора, наприклад, системи "КіберАгрегатор" (Ланде, Субач, & Соболєв, 2019) з ключовими словами, які мають міститися в документі для подальшого аналізу. Після знаходження достатньої кількості текстових повідомлень, вони фільтруються за допомогою програмного коду, згенерованого ГШІ (напр., ChatGPT), для пошуку пар понять, які мають формат "Ім'я Прізвище".

На наступному кроці розв'язується завдання фільтрації наданих словосполучень. Інформація конвертується у файл формату PDF і формується запит (промпт) до системи ГШІ, наприклад, ChatGPT з наступним формулюванням (рис. 3).

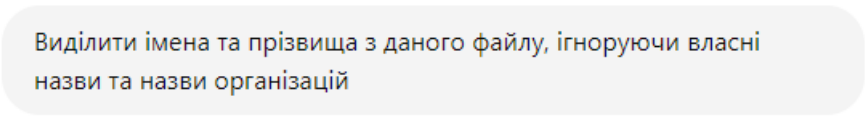

Рис. 3. Приклад промпта до системи ГШІ для фільтрації наданих словосполучень

Під час проведення експерименту з понад 30 000 словосполучень виділено близько 700 імен. Для оптимізації побудови мережі був розроблений програмний код мовою програмування Python за допомогою якого здійснювався підрахунок кількості повторень і вилучення всіх появ сутностей, окрім першої, а також вилучення слів, які згадуються кількість разів, що менша заданого порогового значення (у проведеному експерименті – 3), оскільки вони не мають статистичної важливості й лише перевантажують мережу зайвою інформацією.

За допомогою системи ГШІ (у нашому дослідженні – ChatGPT) створюються зв'язки між а́кторами кібервійни (рис. 4).

ISSN 2707-1758

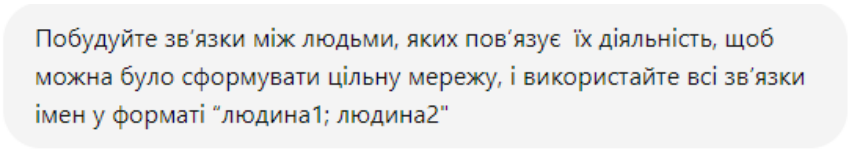

Рис. 4. Приклад промпта до системи ГШІ для створення зв'язків між а́кторами кібервійни

На третьому кроці, після встановлення зв'язків між учасниками в заданому форматі, отримана інформація записується у CSV-файл.

На четвертому, заключному кроці, використовуть спеціальний програмний застосунок і створюють графічне представлення мережі а́кторів кібервійни та їхніх зв'язків (рис. 5).

Побудована мережа а́кторів у формі графа надає важливий інструмент для аналізу та візуалізації взаємодії між ними. Проте варто зауважити, що під час застосування систем ГШІ потрібна експертна перевірка для уточнення складу а́кторів та їхніх зв'язків.

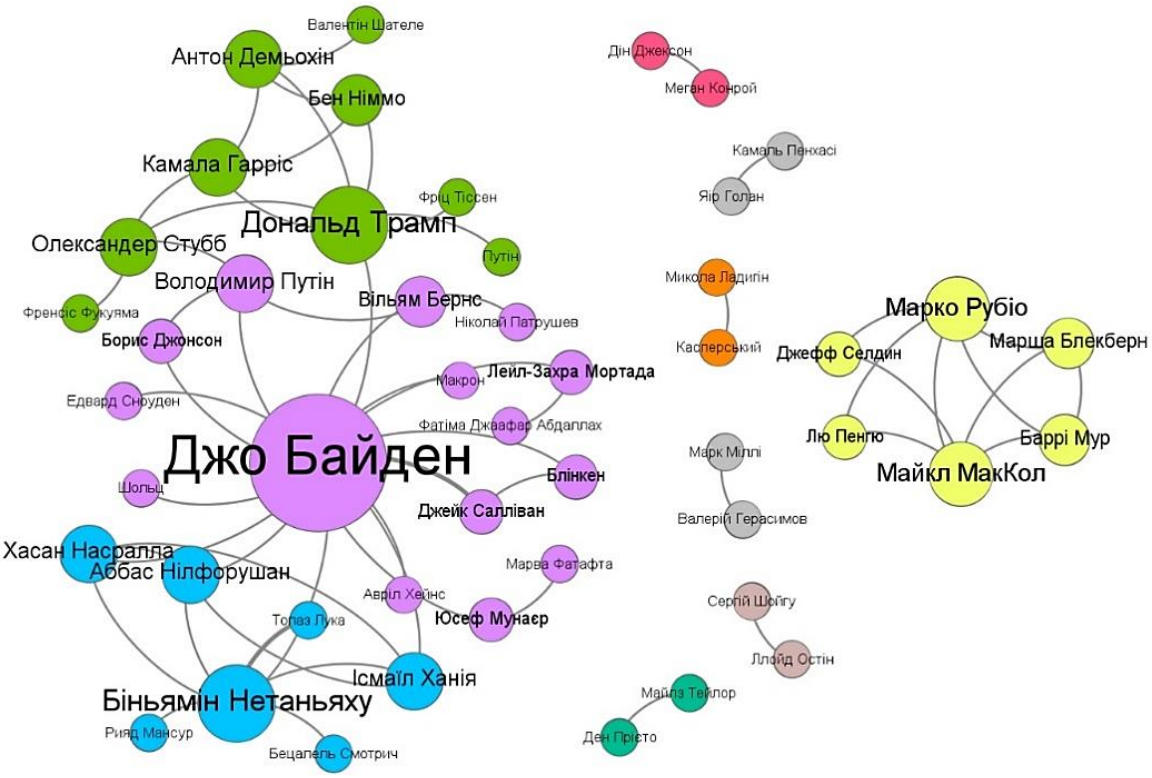

Рис. 5. Фрагмент мережі а́кторів кібервійни ВИПРАВИТИ НА РИС. – БЕНЬЯМІН НетаньяГу

Алгоритм розв'язання задачі

Крок 1. Сформувати запит до пошукової системи-агрегатора з використанням ключових слів K:

$$Query(K) = aggregate\_search(K), \quad (6)$$

де $aggregate\_search(K)$ – функція, що виконує пошук документів *D*, які містять усі ключові слова з множини *K*.

Крок 2. Екстрагувати імена та прізвища. З відібраних документів D екстрагувати пари імен і прізвищ, що мають формат "Ім'я Прізвище", за допомогою системи ГШІ:

$$N = \{(n,s) \mid (n,s) \in extract\_names(d), d \in D\}, \quad (7)$$

де $extract\_names(d)$ – функція, яка використовує систему ГШІ для вилучення імен і прізвищ із документа d.

Крок 3. Фільтрувати й оптимізувати список імен. На цьому кроці застосовується фільтрація отриманих імен для відкидання власних назв і назв організацій, а також для вилучення менш частотних слів. Фільтрація власних назв виконується системою ГШІ для відокремлення імен від власних назв і назв організацій.

Фільтрація отриманих пар імен за частотністю здійснюється так. Для кожної пари $(n,s) \in N$ підраховують кількість її появ:

$$f(n,s) = \sum_{d \in D} count(n,s,d), \quad (8)$$

де $count(n,s,d)$ – кількість появ пари імен (n,s) у документі d.

Після підрахунку вилучають пари, для яких $f(n,s) < f_{\min}$:

$$N' = \{(n,s) \mid (n,s) \in N, f(n,s) \geq f_{\min}\}. \quad (9)$$

Крок 4. Побудувати мережу зв'язків між а́кторами кібервійни. На цьому кроці визначають зв'язки між

ISSN 2707-1758

а́кторами кібервійни на основі їхньої наявності в документах або контекстних взаємодіях:

$$C = \{(a_1, a_2) \mid a_1, a_2 \in N', connected(a_1, a_2, D)\}, \quad (10)$$

де $connected(a_1, a_2, D)$ – функція, що визначає зв'язок між акторами a_1 та a_2 на основі їхньої наявності в одному або кількох документах, або на основі контексту, визначеного системою ГШІ.

Крок 5. Візуалізувати мережу а́кторів. Зібрана інформація записується у файл формату CSV, після чого будується графічне представлення мережі а́кторів кібервійни. Мережу а́кторів G представлено у вигляді графа $G = (V, E)$, де $V = N'$ – множина вузлів, що відповідають а́кторам; $E = C$ – множина зв'язків між а́кторами.

Граф G візуалізують із використанням програмного забезпечення для аналізу соціальних мереж.

Обчислювальна складність наведеного алгоритму буде така.

- обчислювальна складність кроку формування запиту до пошукової системи залежить від кількості ключових слів $|K|$ і кількості знайдених документів N;
- обчислювальна складність кроку екстрагування імен залежить від кількості документів N та обсягу тексту в кожному документі M_d. Отже, загальна складність цих кроків буде $O(N \times M_d)$;
- обчислювальна складність кроку фільтрації та оптимізації складається зі складності підрахунку частотності: $O(N \times M_d)$ та складності відбору значущих пар: $O(p)$, де p – кількість пар (n, s);
- обчислювальна складність кроку побудови мережі зв'язків визначається кількістю а́кторів $|N'|$ та кількістю зв'язків між ними $|C|$. Тоді загальна складність цього кроку буде $O(|N'|^2)$.

Відповідно до наведеного, загальна складність алгоритму з урахуванням усіх кроків дорівнюватиме

$$O(N \times M_d + p + |N'|^2). \quad (11)$$

Реалізована на практиці запропонована методика дозволяє ефективно визначати а́кторів кібервійни з використанням системи ГШІ, а також формувати та візуалізувати мережу їхніх взаємодій.

Дискусія і висновки

У статті ми представили методики екстрагування об'єктів кібербезпеки з використанням ВММ і ГШІ. Основну увагу приділяли розробленню підходів та алгоритмів до виявлення фактографічних даних про об'єкти кібербезпеки з текстових документів. Результати дослідження демонструють ефективність запропонованих підходів і можливість їхнього застосування на практиці для розв'язання задач сфери кібербезпеки.

Використання системи ГШІ, такої як ChatGPT, дозволяє екстрагувати поняття та зв'язки між ними за допомогою звернення до системи із змістовними промптами. Проведене дослідження показує, що вказаний підхід може бути застосованим для ефективного оброблення й аналізу електронних текстів із вебпросторе та соціальних мереж, дозволяючи формувати мережі взаємозв'язків між суб'єктами кібербезпеки.

Запропонована методика екстрагування а́кторів кібервійни дозволяє визначати їх на основі даних, отриманих унаслідок інформаційно-пошукових запитів до агрегаторів інформації та формування промптів до системи ГШІ. Результатом є побудована мережа а́кторів у формі графа, що враховує взаємозв'язки між ними.

Проведені експерименти показують, що отримані результати дозволяють досягти високої ефективності у виявленні й аналізі об'єктів кібербезпеки, а використання ВММ і ГШІ уможливлює подолання обмежень традиційних методів і забезпечує високу точність і швидкість оброблення текстових даних великого обсягу та складної структури.

Реалізація на практиці отриманих результатів може стати важливим інструментом для фахівців сфери кібербезпеки, допомагаючи їм у розробленні ефективних стратегій захисту від кіберзагроз.

Перспективи подальшого розвитку цієї сфери включають удосконалення алгоритмів екстрагування, а також інтеграцію їх з іншими системами кібербезпеки для отримання більш повної картини кіберзагроз.

Внесок авторів: Олександр Пучков – формальний аналіз, методологія; Дмитро Ланде – програмне забезпечення, валідація даних; Ігор Субач – концептуалізація, написання (перегляд і редагування).

Список використаних джерел

Даник, Ю. Г, Воробієнко, П. П., & Чернега, В. М. (2019). *Основи кібербезпеки та кібероборони.* Одеська національна академія зв'язку імені О. С. Попова.

Ланде, Д., Субач, І., & Соболєв А. (2019). Комп'ютерна програма контент-моніторингу соціальних мереж з питань кібербезпеки – КіберАгрегатор ("КіберАгрегатор"). *Свідоцтво про реєстрацію авторського права на твір № 92744.* Міністерство економіки.

Alam, T., Bhusal, D., Park, Y., & Rastogi, N. (2022). CyNER: A Python Library for Cybersecurity Named Entity Recognition. https://doi.org/10.48550/arXiv.2204.05754

Bayer, M., Kuehn, P., Shanehsaz, R., & Reuter, C. (2024). CySecBERT: A Domain-Adapted Language Model for the Cybersecurity Domain. *ACM Transactions on Privacy and Security*, 27(2), 1–20. https://doi.org/10.1145/3652594

Lande, D., Subach, I., Puchkov, O., & Soboliev, A. (2020). A Clustering Method for Information Summarization and Modelling a Subject Domain. *Information & Security*, 50(1), 79–86. https://doi.org/10.11610/isij.5013

Gao, C., Zhang, X., & Han, M. (2021). A review on cyber security named entity recognition. *Front Inform Technol Electron Eng*, 1153–1168. https://doi.org:/10.1631/FITEE.2000286

Halbouni, A., Gunawan, T. S., Habaebi, M. H., Halbouni, M., Kartiwi, M., & Ahmad, R. (2022). Machine Learning and Deep Learning Approaches for CyberSecurity. *IEEE Access*, 10, 19572–19585. https://doi.org/10.1109/ACCESS.2022.3151248

Hanks, C., Maiden, M., Ranade, P., Finin, T., & Joshi, A. (2022). Recognizing and extracting cybersecurity entities from text. *In Workshop on Machine Learning for Cybersecurity. International Conference on Machine Learning.* https://www.researchgate.net/publication/361618891_Recognizing/_and_Extracting_Cybersecurity_Entities_from_Text

Lande, D., Puchkov, O., & Subach, I. (2020). Система аналізу великих обсягів даних з питань кібербезпеки із соціальних медіа. *Collection "Information Technology and Security*, 8(1), 4–18. https://doi.org/10.20535/2411-1031.2020.8.1.217993

Lande, D., Puchkov, O., & Subach, I. (2022). Method of Detecting Cybersecurity Objects Based on OSINT Technology. *In XXII International Scientific and Practical Conference "Information Technologies and Security" (ITS 2022), Vol. 3503* (pp. 115–124). State University of Information and Communication Technologies. https://ceur-ws.org/Vol-3503/paper11.pdf

Piyush, G., & Okamura, K. (2021). Investigating Cybersecurity News Articles by Applying Topic Modeling Method. *In International Conference on Information Networking (ICOIN)* (pp. 432–438). IEEE.

Yi, F., Jiang, B., Wang, L., & Wu, J. (2020). Cybersecurity Named Entity Recognition Using Multi-Modal Ensemble Learning. *IEEE Access*, 8, 63214–63224. https://doi.org/10.1109/ACCESS.2020.2984582

References

Alam, T., Bhusal, D., Park, Y., & Rastogi, N. (2022). CyNER: A Python Library for Cybersecurity Named Entity Recognition. https://doi.org/10.48550/arXiv.2204.05754

ISSN 2707-1758

Bayer, M., Kuehn, P., Shanehsaz, R., & Reuter, C. (2024). CySecBERT: A Domain-Adapted Language Model for the Cybersecurity Domain. *ACM Transactions on Privacy and Security*, 27(2), 1–20. https://doi.org/10.1145/3652594

Danyk, Y. G., Vorobienko, P. P., & Chernega, V. M. (2019). *Fundamentals of Cybersecurity and Cyberdefense.* O. S. Popov Odessa National Academy of Telecommunications [in Ukranian].

Lande, D., Subach, I., & Soboliev, A. (2019). Computer Program for Content Monitoring of Social Networks on Cybersecurity Issues – CyberAggregator ("CyberAggregator"). *Copyright registration certificate No 92744.* Ministry of Economy [in Ukranian].

Lande, D., Subach, I., Puchkov, O., & Soboliev, A. (2020). A Clustering Method for Information Summarization and Modelling a Subject Domain. *Information & Security*, 50(1), 79–86. https://doi.org/10.11610/isij.5013

Gao, C., Zhang, X., & Han, M. (2021). A review on cyber security named entity recognition. *Front Inform Technol Electron Eng*, 1153–1168. https://doi.org:/10.1631/FITEE.2000286

Halbouni, A., Gunawan, T. S., Habaebi, M. H., Halbouni, M., Kartiwi, M., & Ahmad, R. (2022). Machine Learning and Deep Learning Approaches for CyberSecurity. *IEEE Access*, 10, 19572–19585. https://doi.org/10.1109/ACCESS.2022.3151248

Hanks, C., Maiden, M., Ranade, P., Finin, T., & Joshi, A. (2022). Recognizing and extracting cybersecurity entities from text. *In Workshop on Machine Learning for Cybersecurity. International Conference on Machine Learning.* https://www.researchgate.net/publication/361618891_Recognizing/_and_Extracting_Cybersecurity_Entities_from_Text

Lande, D., Puchkov, O., & Subach, I. (2020). Система аналізу великих обсягів даних з питань кібербезпеки із соціальних медіа. *Collection "Information Technology and Security*, 8(1), 4–18. https://doi.org/10.20535/2411-1031.2020.8.1.217993

Lande, D., Puchkov, O., & Subach, I. (2022). Method of Detecting Cybersecurity Objects Based on OSINT Technology. *In XXII International Scientific and Practical Conference "Information Technologies and Security" (ITS 2022), Vol. 3503* (pp. 115–124). State University of Information and Communication Technologies.https://ceur-ws.org/Vol-3503/paper11.pdf

iyush, G., & Okamura, K. (2021). Investigating Cybersecurity News Articles by Applying Topic Modeling Method. *In International Conference on Information Networking (ICOIN)* (pp. 432–438). IEEE.

Yi, F., Jiang, B., Wang, L., & Wu, J. (2020). Cybersecurity Named Entity Recognition Using Multi-Modal Ensemble Learning. *IEEE Access*, 8, 63214–63224. https://doi.org/10.1109/ACCESS.2020.2984582

Отримано редакцією журналу / Received: 17.10.24
Прорецензовано / Revised: 27.10.24
Схвалено до друку / Accepted: 05.11.24

Olexandr PUCHKOV, PhD (Philos.), Prof.
ORCID ID: 0000-0002-8585-1044
e-mail: iszzi@iszzi.kpi.ua
National Technical University of Ukraine
"Igor Sikorsky Kyiv Polytechnic Institute", Kyiv, Ukraine

Dmytro LANDE, DSc (Engin.), Prof.
ORCID ID: 0000-0003-3945-1178
e-mail: dwlande@gmail.com
National Technical University of Ukraine
"Igor Sikorsky Kyiv Polytechnic Institute", Kyiv, Ukraine

Ihor SUBACH, DSc (Engin.), Prof.
ORCID ID: 0000-0002-9344-713X
e-mail: igor_subach@ukr.net
National Technical University of Ukraine
"Igor Sikorsky Kyiv Polytechnic Institute", Kyiv, Ukraine

METHODS OF EXTRACTING CYBERSECURITY OBJECTS FROM ELECTRONIC SOURCES USING ARTIFICIAL INTELLIGENCE

Background. ***The rapid development of information technology (IT) has led to new threats and challenges in the field of cybersecurity. Cyber warfare has become a reality and a real problem for states, organizations and individual users of cyberspace. Ukraine is taking a number of measures to develop a system of cyber actions in cyberspace, which include a set of interconnected subsystems of cyber intelligence, cyber defense, cyber influence and cyber counterintelligence. One of the forms of cyber intelligence is open-source computer intelligence (OSINT), which is used to search for and obtain intelligence information, including the identification and analysis of cybersecurity objects to predict possible manifestations of cyber threats and their consequences. This requires the development of effective methods for detecting and analyzing cybersecurity objects by extracting factual data on cybersecurity objects from large amounts of unstructured textual information.***

Methods. ***The paper investigates artificial intelligence technologies, in particular, large language models (LLM) and generative artificial intelligence (GenAI) in the context of their application to solve the problems of computer intelligence of cybersecurity objects from open electronic sources and social networks.***

Results. ***As a result of the study, in order to carry out an effective analysis of the results of information extraction, a methodology for extracting named entities - the names of hacker groups and their contextual connections from the texts of messages of electronic network sources related to the subject area of cybersecurity, as well as the formation of networks of their interconnections and a substantive analysis of these networks is proposed. To identify the actors involved in cyber warfare, the author proposes a methodology for analyzing selected documents available in electronic sources on the Internet and social networks. Both methods are based on the use of artificial intelligence.***

Conclusions. ***The results of the study demonstrate the effectiveness of the proposed approaches and the possibility of their practical application in solving cybersecurity problems. The proposed methods can be an important tool for cybersecurity professionals to develop effective strategies to protect against cyber threats.***

Keywords: ***cyber warfare, cybersecurity; generative artificial intelligence; large language models; Internet; open electronic sources; social networks, text analysis; objects of cyber warfare.***

Автори заявляють про відсутність конфлікту інтересів. Спонсори не брали участі в розробленні дослідження; у зборі, аналізі чи інтерпретації даних; у написанні рукопису; в рішенні про публікацію результатів.

The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses or interpretation of data; in the writing of the manuscript; in the decision to publish the results.