УДК 004.8:004.056.53+004.942 DOI: 10.31673/2412-4338.2024.040560

Ланде Дмитро Володимирович
Національний технічний університет України "Київський політехнічний інститут імені Ігоря Сікорського", Київ
ORCID 0000-0003-3945-1178

Пучков Олександр Олександрович
Національний технічний університет України "Київський політехнічний інститут імені Ігоря Сікорського", Київ
ORCID 0000-0002-8585-1044

Субач Ігор Юрійович
Національний технічний університет України "Київський політехнічний інститут імені Ігоря Сікорського", Київ
ORCID 0000-0002-9344-713X

МЕТОДИКА ФОРМУВАННЯ ПРИЧИННО-НАСЛІДКОВИХ МЕРЕЖ У СФЕРІ КІБЕРБЕЗПЕКИ ЗАСОБАМИ ГЕНЕРАТИВНОГО ШТУЧНОГО ІНТЕЛЕКТУ

Анотація. *У цій статті запропоновано методику формування причинно-наслідкових мереж (ПНМ) у сфері кібербезпеки за допомогою генеративного штучного інтелекту (ГШІ). Методика базується на ієрархічному зверненні до систем ГШІ, таких як ChatGPT, для визначення центрального вузла та рівнів ієрархії, а також для подальшого уточнення причинно-наслідкових зв'язків. Суть запропонованої методики полягає у визначенні центрального вузла та рівнів ієрархії, формуванні множини пов'язаних понять, візуалізації первинної казуальної мережі, взаємодії з роєм віртуальних експертів (РВЕ) для покращення точності й повноти мережі та формуванні кінцевої ПНМ. Розглянуто можливість використання програми Gephi для візуалізації графу, що представляє казуальну мережу. Приведено методику вибору та застосування порогу значущості для фільтрації незначних зв'язків з метою формування більш точної та повної кінцевої ПНМ для подальшого сценарного аналізу в сфері кібербезпеки. Розглянуто різні варіанти застосування порогу значущості відповідно до характеристик мережі, попередніх знань або аналізу тренувальних даних, а також на основі таких статистичних показників, як середнє значення ваги та стандартне відхилення. Проаналізовано можливість динамічного коригування порогу значущості на основі оцінки якості кінцевої мережі, з врахуванням таких її показників, як кількість кластерів, згуртованість мережі та значущість зв'язків. Наведено приклади запитів до систем ГШІ та результати їхнього виконання, які дозволяють краще зрозуміти процес формування мережі. Результати експериментів показують, що запропонована методика дозволяє ефективно формувати ПНМ, які можуть бути використані для подальшого сценарного аналізу в сфері кібербезпеки.*

Ключові слова: *кібербезпека, генеративний штучний інтелект, ChatGPT, ієрархічне звернення, віртуальні експерти, сценарний аналіз, каузальні мережі, Gephi, текстова аналітика, мережевий аналіз.*

Lande Dmytro
National technical university of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Kyiv
ORCID 0000-0003-3945-1178

Puchkov Oleksandr
National technical university of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Kyiv
ORCID 0000-0002-8585-1044

Subach Ihor
National technical university of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Kyiv
ORCID 0000-0002-9344-713X

METHODOLOGY FOR CONSTRUCTING CAUSAL NETWORKS IN CYBERSECURITY USING GENERATIVE ARTIFICIAL INTELLIGENCE

Abstract. *This article proposes a methodology for the formation of causal networks in the field of cybersecurity using generative artificial intelligence (GAI). The methodology is based on a hierarchical approach to AI systems, such as ChatGPT, to determine the central node and levels of the hierarchy, as well as to further clarify the causal relationships. The essence of the proposed methodology is to determine the central node and hierarchy levels, form a set of related concepts, visualize the primary casual network, interact with a swarm of virtual experts to improve the accuracy and completeness of the network, and form the final causal network. The possibility of using the Gephi program to visualize a graph representing a casual network is considered. The article presents a methodology for selecting and applying a significance threshold for filtering insignificant connections in order to form a more accurate and complete final causal network for further scenario analysis in the field of cybersecurity. Various options for applying the significance threshold are considered, depending on the characteristics of the network, prior knowledge or analysis of training data, as well as on the basis of statistical indicators such as the average weight and standard deviation. The possibility of dynamically adjusting the significance threshold based on an assessment of the quality of the final network, taking into account such indicators as the number of clusters, network cohesion, and the significance of links, is analyzed. Examples of queries to the GCI systems and the results of their execution are presented, which allow us to better understand the process of network formation. Experimental results show that the proposed methodology allows to effectively form causal networks that can be used for further scenario analysis in the field of cybersecurity.*

Keywords: *Cybersecurity, generative artificial intelligence, ChatGPT, hierarchical approach, virtual experts, scenario analysis, causal networks, Gephi, text analytics, network analysis.*

1. Вступ.

У сучасному світі, де забезпечення кібербезпеки стає все більш критичною проблемою, ефективне управління ризиками та прогнозування потенційних загроз є надзвичайно важливим. Одним з ефективних підходів до вирішення цієї проблеми є використання причинно-наслідкових (каузальних) мереж, які дозволяють моделювати та аналізувати взаємозв'язки між різними факторами, що впливають на безпеку інформаційних систем (ІС). Однак, традиційні методи формування таких мереж часто вимагають значних ресурсів та залучення експертів, що ускладнює їхнє застосування в реальних умовах.

У зв'язку з цим, актуальним стає пошук нових підходів, які дозволяють автоматизувати процес формування ПНМ. Одним з таких підходів є використання генеративного штучного інтелекту (ГШІ), який вже довів свою ефективність у різних сферах, включаючи обробку природної мови та аналіз даних. Системи ГШІ, такі як ChatGPT, здатні генерувати тексти, що містять важливу інформацію про взаємозв'язки між різними поняттями, що є основою для побудови ПНМ. Слід зауважити, що саме причинно-наслідкові зв'язки необхідні, коли моделі впроваджуються в критично важливих сферах, таких як кібербезпека.

У цій статті запропоновано методику формування ПНМ у сфері кібербезпеки за допомогою ГШІ. Основна ідея методики полягає в ієрархічному зверненні до систем ГШІ для визначення центрального вузла та рівнів ієрархії, а також для подальшого уточнення причинно-наслідкових зв'язків. Цей підхід дозволяє формувати ПНМ, які можуть бути використані для подальшого сценарного аналізу в сфері кібербезпеки.

Запропонована методика, включає етапи визначення центрального вузла, формування множини пов'язаних понять, візуалізацію первинної каузальної мережі, а також взаємодію з РВЕ. Крім того, наводяться приклади запитів до системи ГШІ та результати їхнього виконання, які дозволяють краще зрозуміти процес формування мережі.

2. Аналіз останніх досліджень і публікацій.

ГШІ є однією з найбільш перспективних галузей інформаційних технологій (ІТ), яка зосереджена на створенні нових даних або вмісту на основі наявних даних. Системи ГШІ, такі

як GPT (Generative Pre-trained Transformer 3) і Gemini, вже довели свою ефективність у різних сферах, включаючи обробку природної мови, генерацію тексту, синтез мови та інші.

У статті [1] наведено огляд великих мовних моделей (LLMs), зокрема сімейства моделей GPT-3 (GLLMs), таких як ChatGPT і GPT-4. У ній обговорюються базові концепції, як то трансформери, самонавчання, передтреновані мовні моделі, а також досягнення GLLMs у різних завданнях обробки природної мови.

Статтю [2] присвячено огляду напрямку контрольованої генерації тексту, який є важливою частиною в галузі генерації природної мови. Основна увага приділяється використанню великих попередньо натренованих мовних моделей (PLM) на основі трансформерів, що стали основою для більш різноманітної та плавної генерації тексту.

ПНМ є важливим інструментом для моделювання та аналізу взаємозв'язків між різними факторами, що впливають на безпеку ІС. Вони дозволяють виявляти причини та наслідки різних подій, що є критично важливим для ефективного управління ризиками в сфері кібербезпеки.

У роботі [3] досліджується застосування концепції причинності у сфері кібербезпеки, яка до цього мало розглядалася в цьому контексті. Автори пропонують фреймворк під назвою Cybersecurity Granger Causality, який визначає наявність причинних зв'язків у часових рядах атак і використовує їх для прогнозування рівнів кібернападів.

Стаття [4] пропонує новий підхід до виявлення прихованих шкідливих програм, які важко виявити традиційними методами безпеки, що базуються на відомих сигнатурах коду або поведінки. Автори визначають причинні зв'язки між подіями у мережевих запитах і розробляють методи для їх виявлення, використовуючи правила та навчання на основі даних.

Автоматизоване формування ПНМ є важливою проблемою, яка вимагає розробки ефективних методів та інструментів для їхнього створення без необхідності залучення експертів. Існують різні підходи до автоматизованого формування таких мереж, включаючи використання обробки природної мови та машинного навчання.

У статті [5] досліджується, як LLMs формують мережі в контексті соціальних взаємодій, порівнюючи їх поведінку з людською соціальною динамікою. Автори аналізують принципи формування мереж, такі як переважне приєднання, тріадна закритість, структура спільнот і малий світ, та доводять, що LLMs демонструють подібні характеристики. Використання ГШІ для формування ПНМ є відносно новим підходом, який поєднує в собі здатність ГШІ генерувати тексти з можливостями аналізу причинно-наслідкових зв'язків. Цей підхід дозволяє автоматизувати процес формування мереж, зменшуючи залежність від експертів та знижуючи витрати на ресурси.

Стаття [6] присвячена рішенню проблеми виявлення причинно-наслідкових відносин між парами подій з тексту. Автори визначають дві основні проблеми: відхилення меж подій та невідповідність їхніх пар. Для їх вирішення вони використовують LLM для оптимізації визначення завдання та адаптованої рамки витягу причинності подій.

У дослідженні [7] пропонується використовувати можливості LLMs, зокрема генеративних попередньо навчених трансформерів, для покращення методів верифікації безпеки.

Візуалізація та аналіз ПНМ є важливим етапом для їхнього ефективного використання. Існують різні інструменти та методи для візуалізації та аналізу мереж, включаючи програмне забезпечення Gephi та інші. Стаття [8] представляє нову структуру, названу "Автономна структура відкриття причин", яка поєднує алгоритми відкриття причин і LLMs для автоматизації, візуалізації і створення точніших і зрозуміліших каузальних графів. У статті [9] представлено пакет візуалізації Causalvis. Автори створили інтерактивні модулі візуалізації, які підтримують завдання причинно-наслідкового аналізу.

У роботі [10] розглядається можливість інтеграції традиційних систем розвідки у відкритих інформаційних джерелах з передовими технологіями ГШІ, які стають ключовим фактором для розвитку аналітичних систем. Головна увага дослідження спрямована на

вдосконалення функціонування системи контентмоніторингу соціальних медіа з питань кібербезпеки. Детально описано встановлення залежностей та відпрацювання запитів, які трансформуються в промпти для системи ГШІ.

3. Мета і задачі дослідження.

Метою цієї статті є розробка методики формування ПНМ у сфері кібербезпеки за допомогою ГШІ.

Завдання статті:

- проаналізувати досвід застосування систем ГШІ щодо вирішення задач, пов'язаних з формуванням ПНМ;
- описати розроблену методику ієрархічної взаємодії з системами ГШІ, такими як ChatGPT, для автоматизованого визначення центрального вузла, рівнів ієрархії та причинно-наслідкових зв'язків;
- навести результати експериментальної перевірки запропонованої методики на прикладі, що стосується сфери кібербезпеки.

Даний підхід є певним кроком у напрямку автоматизації задач аналізу кібербезпеки та управління ризиками.

4. Результати дослідження.

ПНМ, що будуються з причино-наслідкових зв'язків забезпечують можливість подальшого переходу до сценарного аналізу. Основною проблемою, яка виникала до цього при проведенні сценарного аналізу на основі ПНМ, є саме створення таких мереж, що в традиційних випадках вимагає великих ресурсних витрат та залучення експертів. Підхід до формування РВЕ [11, 12] значно спрощує і пришвидшує процес формування ПНМ.

Методика формування ПНМ ґрунтується на ієрархічному зверненні до систем ГШІ і подальшому уточненні. Основна ідея полягає в тому, щоб використовувати здатність ГШІ генерувати тексти, що містять інформацію про взаємозв'язки між різними поняттями для автоматизованого формування ПНМ. Отже, цей підхід дозволяє зменшити залежність від експертів та знизити витрати на ресурси, необхідні для формування мереж.

Суть етапів методики, яка реалізує запропонований підхід полягає у наступному:

1. Визначення центрального вузла та рівнів ієрархії: вибір центрального вузла, який визначає головну тему або проблему сфери кібербезпеки, для якої необхідно побудувати ПНМ шляхом ієрархічного звернення до системи ГШІ для визначення причин, що впливають на центральний вузол та формування першого рівня ієрархії.

2. Формування множини пов'язаних понять:

- для кожного елемента з множини причин, визначеної на першому рівні, виконуються подальші запити до ГШІ для виявлення нових причин, що впливають на них;
- процес повторюється для кожного нового рівня ієрархії, поки не буде досягнута необхідна глибина мережі.

3. Візуалізація первинної каузальної мережі:

- формування зв'язків між вузлами на різних рівнях ієрархії у вигляді спрямованого графу;
- використання інструментів візуалізації, таких як Gephi, для відображення та аналізу отриманої мережі.

4. Взаємодія з роєм віртуальних експертів:

- визначення РВЕ, які взаємодіють з системою ГШІ для отримання додаткових концептів та зв'язків;
- узагальнення результатів РВЕ для покращення точності та повноти мережі.

5. Формування кінцевої ПНМ:

- вибір порогу значущості для фільтрації незначних зв'язків;
- формування кінцевої мережі на основі відфільтрованих зв'язків та вузлів.

Таким чином, задача формування ПНМ сфери кібербезпеки полягає у наступному.

Дано: C_0 – центральний вузол – поняття, що визначає головну тему або проблему сфери кібербезпеки, для якої необхідно побудувати ПНМ;

P_i – причина або концепт, що безпосередньо впливає на інше поняття (вузол), утворюючи зв'язок виду "причина-наслідок";

L_k — множина рівнів ієрархії (множина вузлів на певному рівні ієрархії), де k є номером рівня.

Необхідно: сформувати кінцеву семантичну мережу $G^\theta = \left(V^\theta, E^\theta\right)$, де V^θ – множина вузлів (понять), а E^θ – множина ребер, що зв'язують ці поняття.

Обмеження та припущення: інтерпретація результатів вимагає досвіду в досліджуваній області людини-експерта для контролювання процесу побудови мережи з метою забезпечення необхідної точності отриманих результатів.

Алгоритм рішення задачі.

Крок 1: визначення центрального вузла та першого рівня ієрархії.

Задається центральний вузол C_0 та виконується запит до системи ГШІ для отримання причин, що впливають на C_0:

$$L_1 = \left\{P_{1,1}, P_{1,2}, \ldots, P_{1,|L_1|}\right\},$$

де L_1 – множина причин $P_{1,i}$ – вузлів, що впливають на C_0.

Нехай нас, наприклад, цікавить проблематика фішингових атак [12]. Тоді до системи ГШІ направляється запит (промпт), наведений на рис. 1. Успішне відпрацювання такого запиту дозволяє визначити другий рівень ієрархії – поняття пов'язані з фішингом, його причини.

Крок 2: формування множини пов'язаних понять.

Для кожного такого поняття $P_{1,i}$ (елемента з множини L_1), визначається множина причин, що вплинули на нього, шляхом виконання подальших запитів до системи ГШІ:

$$L_2^j = \left\{P_{2,j,1}, P_{2,j,2}, \ldots, P_{2,j,|L_2^j|}\right\},$$

де $L_2 = \bigcup_j L_2^j$ – множина причин, які впливають на кожний елемент з L_1.

List the causes of phishing in cyber security. The reason is to use no more than three words. The results should be presented in the format "cause; phishing". Each such entry - from a new line

Рис. 1. Приклад промпту до системи ГШІ

Даний процес може тривати велику кількість ітерацій, проте результати дослідження вказують на доцільність формування трьох рівнів. Зауважимо, що незважаючи на ієрархічне формування такої каузальної мережі, отримана мережа загалом не буде строго ієрархічною структурою.

Направивши до системи ГШІ типу ChatGPT для відпрацювання нею деякий запит (промпт), отримаємо множину причин первинного поняття. Система ГШІ може допомогти отримати зміст CSV-файлу (поля, відповідні поняттям, розділені точкою з комою). Для цього можна застосувати, наприклад, наступний промпт до системи ГШІ (див. рис.1):

Відповідно до наведеного промпту, система ГШІ видає відповідь наступного вигляду (див. рис.2).

poor awareness; phishing
technical flaws; phishing
social engineering; phishing
weak passwords; phishing
untrained staff; phishing
lack of policies; phishing
email spoofing; phishing
vulnerable systems; phishing
fake websites; phishing
link manipulation; phishing
malicious attachments; phishing
trust exploitation; phishing

Рис.2. Приклад відповіді системи ГШІ на промпт користувача

Промпти наступного рівня формуються відповідно до наведених у відповіді концептів і мають вигляд, що повністю відповідає первинному запиту, наприклад (див. рис.3):

List the causes of poor awareness in cyber security. The reason is to use no more than three words. The results should be presented in the format "cause; poor awareness". Each such entry – from a new line

Рис.3. Приклад промпту наступного рівня, відповідно до розглядаємої задачі

Крок 3: Формування та візуалізація первинної каузальної мережі.

Об’єднані в одному CSV-файлі відповіді системи ГШІ завантажуються для аналізу та візуалізації, наприклад, програмою Gephi. Завантаживши отримані дані до системи Gephi, вибирається розмір вузлів, пропорційний ступеню (кількості суміжних зв’язків), і, розділивши мережу на кластери за критерієм модулярності, отримуємо наочний граф (див. рис.4).

Зв’язки формуються між вузлами C_0, L_1, та L_2, а також наступними рівнями, якщо це необхідно і оформлюються у вигляді спрямованого графу:

$$G_1 = (V, E),$$

де V – множина вузлів, що включає C_0, L_1, L_2 та інші рівні, а E – множина зв’язків виду: $\left(P_{k,i}, P_{k+1,j}\right)$.

З рис. 4 видно, що найбільш впливовими вузлами цієї мережі (*Out-Degree*), є: *social engineering* (4), *poor awareness* (3), *insufficient training* (3), *lack of policies* (3) та *weak authentication* (3).

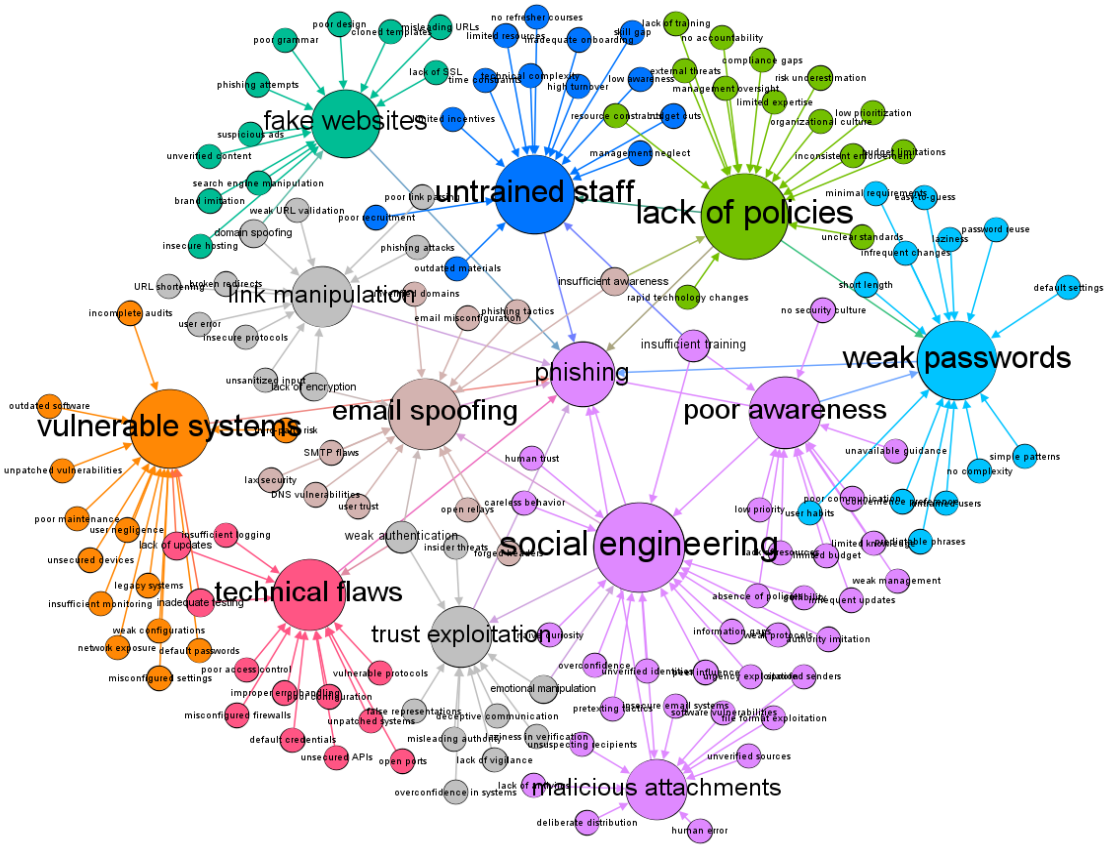

Рис. 4. Спрямована первинна каузальна мережа, отримана шляхом найпростішого ієрархічного звернення до системи ГШІ

Вочевидь, сформована мережа є слабопов'язаною, неповною, а представлені в ній концепти можуть не точно відображати причини та наслідки. Вважатимемо, що це мережа, отримана в результаті опитування лише одного штучного експерта.

Крок 4: Взаємодія з PBE.

Слід зауважити, що у різний час система ГШІ типу ChatGPT може надавати різні варіанти відповідей під час обробки промпту, причому більшість з них є точними та логічно обґрунтованими з людської точки зору. Кожну таку відповідь можна сприймати як відповідь якогось одного віртуального експерта. Можна припустити, що шляхом узагальнення відповідей множини подібних експертів (рою експертів) можна отримати більш повну та точну відповідь.

PBE – це набір незалежних запитів до LLM, з яких отримується відповідь щодо пари найтісніше пов'язаних понять у тексті [12]. Кожен такий запит можна вважати "голосом" окремого віртуального експерта.

Віртуальні експерти можуть бути розділені на групи за функціоналом: аналітики, які відшукують найтісніші зв'язки між поняттями; синтезатори, які аналізують агреговані зв'язки та виявляють нові взаємодії або підтверджують існуючі; рецензенти, що оцінюють значущість отриманих зв'язків і відфільтровують слабкі або нерелевантні.

Отже, впроваджуючи PBE, можна декілька разів виконати однакові промпти, пов'язані як з ієрархіями першого, так і другого рівнів. Таким чином, *для кожного рівня* L_k *виконується багаторазове звернення до системи ГШІ для отримання додаткових концептів, які утворюють нові вузли та зв'язки в мережі:*

$$L_k^{(m)} = \left\{ P_{k,1}^{(m)}, P_{k,2}^{(m)}, \ldots, P_{k,n_k}^{(m)} \right\},$$

де m *– кількість звернень до ГШІ для отримання множини причин* L_k .

Нехай $Q = \{q_1, q_2, ..., q_n\}$ – множина промптів до *ГШІ*, де кожен промпт q_i відповідає одному віртуальному експерту E_i .

Кожний віртуальний експерт E_i відповідає на промпт q_i , генеруючи множину пар понять $P_i = \{(c_{ij}, c_{ij}')\}$, де c_{ij}, c_{ij}' – це два найтісніше повязані поняття в тексті згідно з думкою експерта E_i .

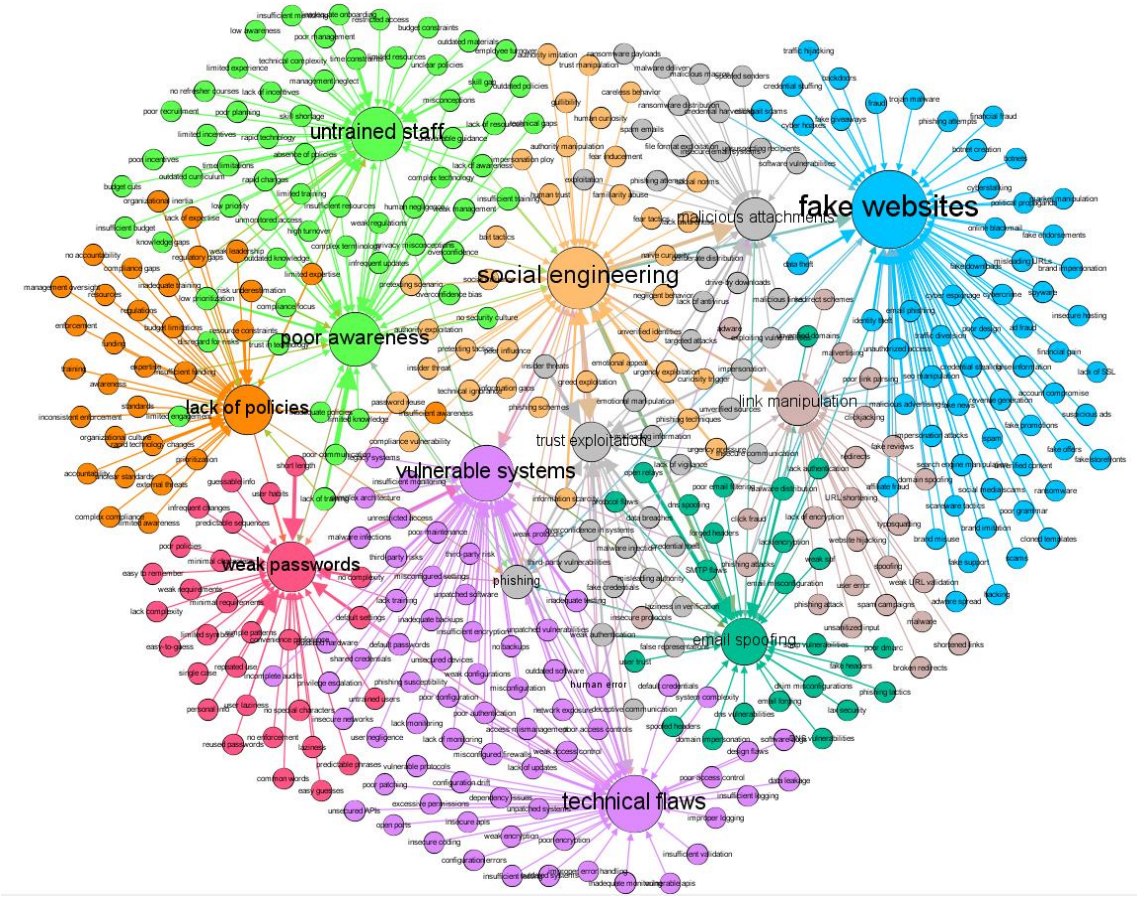

Рис. 5. Спрямована повна каузальна мережа, отримана шляхом ієрархічного звернення РВЕ до системи ГШІ

Всі відповіді агрегуються, а отримані пари понять формують множину:

$$P = \bigcup_{i=1}^{n} P_i.$$

Після отримання відповідей від системи їх можна об’єднати в єдиний CSV-файл для аналізу та візуалізації за допомогою спеціального програмного забезпечення, наприклад, Gephi. Завантаження отриманих даних до програми Gephi дозволяє отримати граф, подібний до наведеного на рис. 5.

На практиці мережа може поповнюватися доти, доки не стане достатньо повною за оцінкою експерта-людини.

З рис. 5 видно, що найбільш впливовими вузлами мережі за заданою тематикою (*Out-Degree*) на наведеному прикладі, є: *social engineering* (8), *phishing* (8), *human error* (6), *poor awareness* (4), *malware distribution* (4), *insider threats* (4). Отже, кількість важливих концептів збільшилася порівняно з попереднім випадком.

Крок 5: формування кінцевої ПНМ.

Формування кінцевої мережі включає в себе вибір порогу значущості для фільтрації незначних зв'язків та формування кінцевої мережі на основі відфільтрованих зв'язків та вузлів. Це дозволяє отримати більш точну та повну ПНМ, яка може бути використана для подальшого сценарного аналізу щодо вирішення задач в сфері кібербезпеки.

Підрахунок частот можна здійснити наступним чином. Кожна пара понять (c,c') з множини P отримує вагу $w_{(c,c')}$, що дорівнює частоті появи цієї пари серед відповідей всіх експертів:

$$w_{(c,c')} = \frac{N(c,c',P)}{n},$$

де $N(c,c',P)$ – кількість появ (c,c') у P.

Встановлюється поріг значущості θ, при якому з множини P залишаються лише ті пари, для яких $w_{(c,c')} > \theta$.

На основі відфільтрованої множини пар понять формується кінцева семантична мережа $G^{\theta} = (V^{\theta}, E^{\theta})$, де V^{θ} – множина вузлів (понять), а E^{θ} – множина ребер, що зв'язують ці поняття.

Поріг значущості θ може бути обраний залежно від бажаної точності та згуртованості семантичної мережі. Для цього можна використовувати такі підходи.

Евристичний поріг. Встановлюється експериментально на основі попередніх знань або аналізу тренувальних даних.

Статистичний підхід. Визначення θ на основі статистичних показників, наприклад, врахування середнього значення ваги або стандартного відхилення. Даний підхід передбачає використання статистичних властивостей ваг, що були отримані для пар понять у множині P.

Якщо $W = \{w_{(c,c')} \mid (c,c') \in P\}$ – множина вагових значень для всіх пар понять, то середнє значення ваги $\mu_w = \frac{1}{|W|}\sum_{w \in W} w$ можна викристати для встановлення базового порогу.

Стандартне відхилення ваги $\sigma_w = \sqrt{\frac{1}{|W|}\sum_{w \in W}(w - \mu_w)^2}$ показує, наскільки значення ваг відрізняються від середнього.

Таким чином, визначення порогу θ може полягати у додаванні множника k до середнього значення ваги:

$$\theta = \mu_w + k \cdot \sigma_w,$$

де k — коефіцієнт, який контролює рівень суворості відбору пар понять. Чим більший k, тим вищий поріг і тим менше зв'язків залишиться у кінцевій мережі.

Адаптивний поріг. Він передбачає динамічне коригування значення θ на основі оцінки якості кінцевої семантичної мережі $G(\theta) = (V(\theta), E(\theta))$, побудованої при даному значенні порогу.

При цьому, початкове значення θ_0 може бути вибране на основі середнього значення ваг або інших евристичних методів, наприклад: $\theta_0 = \mu_w$.

Визначається функція якості мережі $Q(\theta)$, яка може враховувати такі показники як:

- кількість кластерів $C(\theta)$ – кількість окремих підграфів у мережі;
- згуртованість мережі $Coh(\theta)$ – міра внутрішньої злагодженості та кількості зв'язків між вузлами;
- значущість зв'язків $Sig(\theta)$ – середня вага залишених зв'язків.

Відповідно до цього, згуртованість мережі $Coh(\theta)$ враховує щільність зв'язків та їх кластеризацію, а значущість зв'язків $Sig(\theta)$ визначає, наскільки важливими є зв'язки між вузлами, враховуючи середню вагу та медіану ваг зв'язків. Згуртованість мережі враховує щільність зв'язків між вузлами та рівень кластеризації.

Щільність мережі $D(\theta)$ визначається як відношення кількості наявних зв'язків $|E(\theta)|$ до максимально можливої кількості зв'язків у повній мережі:

$$D(\theta) = \frac{2 \cdot |E(\theta)|}{|V(\theta)| \cdot (|V(\theta)| - 1)},$$

де $|V(\theta)|$ – кількість вузлів у мережі, а $|E(\theta)|$ – кількість зв'язків.

Коефіцієнт кластеризації $Clast(\theta)$ визначає ймовірність того, що два сусідні вузли для певного вузла також є сусідами один для одного. Він обчислюється як середнє значення локальних коефіцієнтів кластеризації для всіх вузлів у мережі:

$$Clast(\theta) = \frac{1}{|V(\theta)|} \sum_{v \in V(\theta)} \frac{2 \cdot |e_v|}{k_v \cdot (k_v - 1)},$$

де $|e_v|$ – кількість зв'язків між сусідами вузла v, а k_v – кількість його сусідів.

Згуртованість мережі може бути визначена як комбінація щільності та кластеризації:

$$Coh(\theta) = \alpha \cdot D(\theta) + \beta \cdot Clast(\theta),$$

де α та β – вагові коефіцієнти, що визначають важливість щільності та кластеризації.

Вважатиме, що значущість зв'язків у мережі – це міра того, наскільки зв'язки між вузлами є важливими або сильними з точки зору їхньої ваги. Для її обчислення враховується середня вага зв'язків $\bar{w}(\theta)$ – середнє значення ваг всіх зв'язків у мережі і медіана ваг звязків $Med_w(\theta)$:

$$\bar{w}(\theta) = \frac{1}{|E(\theta)|} \sum_{(c,c') \in E(\theta)} w_{(c,c')},$$

де $w_{(c,c')}$ – вага зв'язку між вузлами c та c'.

Медіана ваг зв'язків $Med_w(\theta)$ визначає центральне значення ваг зв'язків, яке може бути більш стійким до впливу аномальних значень, ніж середня вага.

У свою чергу, значущість зв'язків може бути визначена як комбінація середньої ваги та медіани:

$$Sig(\theta) = \gamma \cdot \bar{w}(\theta) + \delta \cdot Med_w(\theta),$$

де γ та δ – вагові коефіцієнти, що визначають важливість середньої ваги та медіани.

Відповідно до наведеного, загальна функція якості $Q(\theta)$ може бути визначена як:

$$Q(\theta) = \alpha \cdot Coh(\theta) + \beta \cdot Sig(\theta) - \gamma \cdot C(\theta),$$

де α, β, і γ – вагові коефіцієнти, що визначають важливість кожного критерію.

Підсумовуючи зазначимо, що значення порогу θ адаптується таким чином, щоб максимізувати $Q(\theta)$. Це можна зробити шляхом ітераційного пошуку, збільшуючи або зменшуючи θ на крок $\Delta\theta$, поки $Q(\theta)$ не досягне максимального значення:

$$\theta_{opt} = \arg\max_{\theta} Q(\theta).$$

Таким чином, адаптивний поріг θ динамічно підлаштовується в залежності від того, як змінюється якість мережі при його різних значеннях. Це дозволяє досягти оптимального балансу між збереженням значущих зв'язків і уникненням зайвих слабких зв'язків.

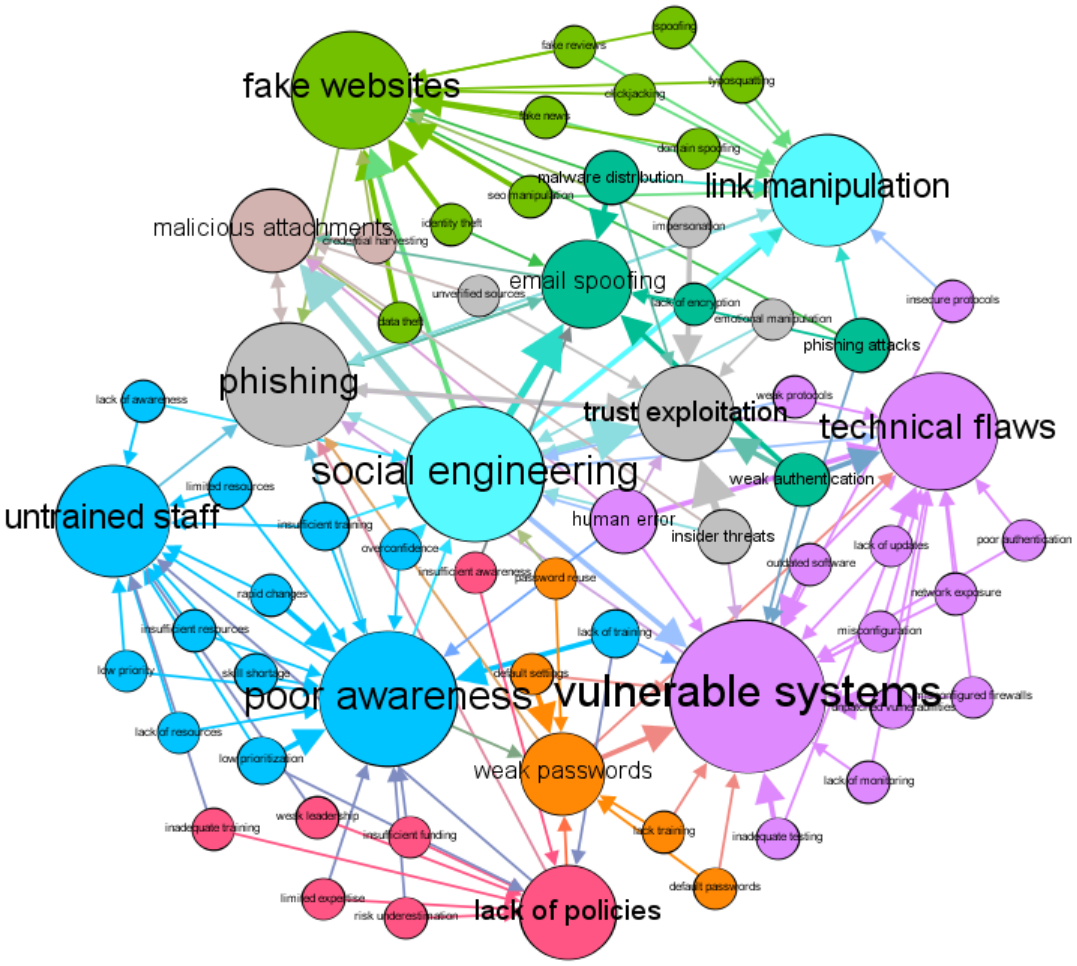

Рис. 6. Спрямована каузальна мережа, отримана шляхом узагальнення ієрархічного звернення PBE до системи ГШІ

Слід зауважити, що граф, сформований у наведеному прикладі, маючи відносно високу повноту понять, може містити неточну інформацію, помилково надану системою ГШІ при обробці окремих промптів. Припускаючи, що існує деяка ймовірність зустрічі з подібними помилками (хоча дослідження вказують на те, що вона відносно мала), можна виключити з

розгляду поняття, які трапляються рідше заданого порогу при побудові мережі. У випадку, представленому на рис. 6, вузли, які мають ступень менше двох, не розглядалися.

Неважко помітити, що найбільш впливовими вузлами цієї мережі (*Out-Degree*) у цьому випадку) є: *social engineering* (8), *phishing* (8), *human error* (6).

Таким чином, первинна ПНМ, отримана шляхом простого ієрархічного запиту до системи ГШІ, охоплює найбільшу кількість понять, які є відносно слабкозв'язаними (мережа близька до ієрархічної), але завдяки повноті її можна розглядати як добру "сировину" для подальшої аналітичної обробки.

З іншого боку, статистично оброблена ПНМ, отримана шляхом ієрархічного запиту РВЕ до системи ГШІ типу ChatGPT, є більш точною, ніж основна мережа, і, нарешті, третя мережа, отримана шляхом узагальнення ієрархічного запиту від РВЕ до системи ГШІ, має найвищий показник кластеризації, тобто більшу степінь взаємодії між окремими концептами, що впливають на цілі в цьому ланцюжку причинності. Цей тип мережі, ймовірно, найбільше підходить для подальшого сценарного аналізу.

Висновки

Проведене дослідження дозволяє зробити наступні висновки:

1. Використання системи ГШІ типу ChatGPT і системи візуалізації графів, наприклад, Gephi для формування і аналізу ПНМ у певних предметних областях, таких як кібербезпека, є цілком конкурентоспроможнім підходом.

2. Запропонована методика формування ПНМ на основі застосування ГШІ дозволяє отримати множину взаємозв'язаних понять з великим ступенем взаємодії між окремими концептами.

3. Експериментально доведено, що інтеграція інструментів текстової аналітики та мережевого аналізу може виявитися корисною для отримання інформації з великих обсягів неструктурованих даних. Одним із результатів наведеного дослідження є доцільність емуляції груп експертів за допомогою системи ГШІ. Цей підхід може підвищити ефективність екстрагування знань і забезпечити глибше розуміння структури та значення текстових документів у різних предметних областях, зокрема, кібербезпеки.

Можна бачити, що використання алгоритмів машинного навчання може допомогти у розблокуванні прихованих в текстових даних закономірностей, а також отримати більш глибоке розуміння складних явищ у сфері кібербезпеки.

Проте, важливо зазначити, що наведений підхід не позбавлений обмежень. Насамперед, інтерпретація результатів вимагає досвіду в досліджуваній області та існує потреба контролювання процесу побудови мережи людиною-експертом для забезпечення точності отриманих результатів.

Запропонована методика формування ПНМ за допомогою ГШІ дозволяє автоматизувати процес формування мереж, зменшуючи залежність від експертів та знижуючи витрати на ресурси. Результати експериментів показали, що цей підхід є ефективним для формування ПНМ у сфері кібербезпеки, що може бути використано для подальшого сценарного аналізу та управління ризиками.

Список використаної літератури

1. Kalyan K. S. A survey of GPT-3 family large language models including ChatGPT and GPT-4 // Natural Language Processing Journal. – 2023. – P. 100048. DOI: 10.1016/j.nlp.2023.100048.

2. Zhang H., Song H., Li S., Zhou M., Song D. A survey of controllable text generation using transformer-based pre-trained language models // ACM Computing Surveys. – 2023. – Vol. 56, No. 3. – P. 1-37. DOI: 10.1145/3617680.

3. Trieu-Do V., Garcia-Lebron R., Xu M., Xu S., Feng Y. Characterizing and leveraging Granger causality in cybersecurity: Framework and case study // ICST Transactions on Security and Safety. – 2021. – Vol. 7, No. 25. DOI: 10.4108/eai.11-5-2021.169912.

4. Zhang H., Yao D. D., Ramakrishnan N., Zhang Z. Causality reasoning about network events for detecting stealthy malware activities // Computers & Security. – 2016. – Vol. 58. – P. 180-198. DOI: 10.1016/j.cose.2016.01.002.

5. Papachristou M., Yuan Y. Network Formation and Dynamics Among Multi-LLMs // arXiv preprint. – 2024. – P. arXiv:2402.10659. DOI: 10.48550/arXiv.2402.10659.

6. Luo K., Zhou T., Chen Y., Zhao J., Liu K. Open Event Causality Extraction by the Assistance of LLM in Task Annotation, Dataset, and Method // In Proceedings of the Workshop: Bridging Neurons and Symbols for Natural Language Processing and Knowledge Graphs Reasoning (NeusymBridge)@ LREC-COLING-2024. – 2024. – P. 33-44.

7. Saha D., Tarek S., Yahyaei K., Saha S. K., Zhou J., Tehranipoor M., Farahmandi F. LLM for SoC Security: A Paradigm Shift // IEEE Access. – 2024. DOI: 10.1109/ACCESS.2024.3427369.

8. Khatibi E., Abbasian M., Yang Z., Azimi I., Rahmani A. M. ALCM: Autonomous LLM-Augmented Causal Discovery Framework // arXiv preprint. – 2024. – P. arXiv:2405.01744. DOI: 10.48550/arXiv.2405.01744.

9. Guo G., Karavani E., Endert A., Kwon B. Causalvis: Visualizations for Causal Inference // Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems. – 2023. – P. 1-20. DOI: 10.1145/3544548.3581236.

10. Пучков О., Ланде Д., Субач І., Рибак О. Інтеграція технологій інформаційного пошуку і штучного інтелекту в галузі кібербезпеки. // Information Technology and Security. – 2023. – Том. 11, № 2. – С. 206–215. DOI: 10.20535/2411-1031.2023.11.2.293789.

11. Lande D., Strashnoy L. Concept Networking Methods Based on ChatGPT & Gephi // SSRN. – 2023. Available at: http://dx.doi.org/10.2139/ssrn.4420452.

12. Ланде Д.В., Страшной Л.Л.. *Ієрархічне формування причинно-наслідкових мереж на основі ChatGPT*: матеріали Першої Всеукр. наук.-практ. конф., присвяченої 100-річному ювілею академіка В.М. Глушкова, м. Київ, 26 травня 2023 р. Київ, 2023. С.24-30.

References

1. Kalyan K. S. A survey of GPT-3 family large language models including ChatGPT and GPT-4 // Natural Language Processing Journal. – 2023. – P. 100048. DOI: 10.1016/j.nlp.2023.100048.

2. Zhang H., Song H., Li S., Zhou M., Song D. A survey of controllable text generation using transformer-based pre-trained language models // ACM Computing Surveys. – 2023. – Vol. 56, No. 3. – P. 1-37. DOI: 10.1145/3617680.

3. Trieu-Do V., Garcia-Lebron R., Xu M., Xu S., Feng Y. Characterizing and leveraging Granger causality in cybersecurity: Framework and case study // ICST Transactions on Security and Safety. – 2021. – Vol. 7, No. 25. DOI: 10.4108/eai.11-5-2021.169912.

4. Zhang H., Yao D. D., Ramakrishnan N., Zhang Z. Causality reasoning about network events for detecting stealthy malware activities // Computers & Security. – 2016. – Vol. 58. – P. 180-198. DOI: 10.1016/j.cose.2016.01.002.

5. Papachristou M., Yuan Y. Network Formation and Dynamics Among Multi-LLMs // arXiv preprint. – 2024. – P. arXiv:2402.10659. DOI: 10.48550/arXiv.2402.10659.

6. Luo K., Zhou T., Chen Y., Zhao J., Liu K. Open Event Causality Extraction by the Assistance of LLM in Task Annotation, Dataset, and Method // In Proceedings of the Workshop: Bridging Neurons and Symbols for Natural Language Processing and Knowledge Graphs Reasoning (NeusymBridge)@ LREC-COLING-2024. – 2024. – P. 33-44.

7. Saha D., Tarek S., Yahyaei K., Saha S. K., Zhou J., Tehranipoor M., Farahmandi F. LLM for SoC Security: A Paradigm Shift // IEEE Access. – 2024. DOI: 10.1109/ACCESS.2024.3427369.

8. Khatibi E., Abbasian M., Yang Z., Azimi I., Rahmani A. M. ALCM: Autonomous LLM-Augmented Causal Discovery Framework // arXiv preprint. – 2024. – P. arXiv:2405.01744. DOI: 10.48550/arXiv.2405.01744.

9. Guo G., Karavani E., Endert A., Kwon B. Causalvis: Visualizations for Causal Inference // Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems. – 2023. – P. 1-20. DOI: 10.1145/3544548.3581236.

10. Puchkov O., Lande D., Subach I., Rybak O. Integration of information search technologies and artificial intelligence in the field of cybersecurity.. // Information Technology and Security. – 2023. – Vol. 11, no 2. – P. 206–215. DOI: 10.20535/2411-1031.2023.11.2.293789.

11. Lande D., Strashnoy L. Concept Networking Methods Based on ChatGPT & Gephi // SSRN. – 2023. Available at: http://dx.doi.org/10.2139/ssrn.4420452.

12. Lande D.V., Strashnoy L.L. *Ієрархічне формування причинно-наслідкових мереж на основі ChatGPT*: Proceedings of the First All-Ukrainian Scientific and Practical Conference dedicated to the 100th anniversary of Academician V.M. Glushkov, Kyiv, May 26, 2023 Kyiv, 2023. P.24-30.