УДК 004.5

Формування каузальної мережі вразливостей в кібербезпеці на основі ChatGPT

Ланде Д.В., Срашной Л.Л.

Постановка проблеми

Серед завдань, які вирішуються сьогодні великими лінгвістичними моделями, такими як ChatGPT [1], можна назвати машинний переклад, реферування та переказ текстів, генерація нових текстів, виділення тематик, питань до текстів. Можливості в екстрагуванні іменних сутностей дозволяють використовувати ChatGPT у фактографічних системах, зокрема, в медицині, економіці. Звичайно, інтелектуальні чати інтегруються із зовнішніми системами, такими як геоінформаційні, системи аналізу та візуалізації графів, мереж. У роботах [2, 3] авторами показано, як можна формувати мережі зв'язків персонажів літературних творів, мережі предметних областей зі зв'язками типу «загальне-часткове».

Ця робота присвячена опису методики формування каузальних мереж з допомогою двох різних методик. Якщо перша методика є варіантом методики, наведеної в [3], то друга методика, що базується на декомпозиції основного поняття, наводиться в цій доповіді вперше.

В рамках цієї роботи також представляється програма візуалізації мереж, побудована на основі бібліотеки візуалізації графів GaphViz лабораторії AT&T. Для завантаження даних у її середовище цілком підходить формат CSV, тому всі запити до ChatGPT супроводжуються вимогою до цього формату.

Мета

Метою доповіді є представлення методів побудови каузальних мереж, що відповідають заданій проблемі. Передбачається, що предметна область у достатній мірі охоплювалася ChatGPT при її навчанні. Ця робота присвячена саме створенню каузальної мережі вразливостей в кібербезпеці (**cyber security vulnerability**).

Виклад основного матеріалу

Методика 1: багаторазове виконання одного і того ж промпту

У рамках першої методики пропонується багаторазове виконання одного й того ж промпту до системи ChatGPT, тобто емуляція діяльності багатьох експертів, які вирішують одну й ту саму задачу. Таким чином, пропонується відпрацювати деякий промпт для ChatGPT, щоб одержимо множину причин первинного поняття (cyber security vulnerability). Система ChatGPT може допомогти в отриманні змісту форматованого файлу (поля, що відповідають іменам персонажів, розділені крапкою з комою). Для цього застосовується, наприклад, такий простий промпт:

```
→ Name 10 reasons for the term "cyber security vulnerability".  Each reason must contain no more than three words. Present the answer in the form: "reason; cyber security vulnerability". Each entry on a separate line.
```

Система видає відповідь такого вигляду:

```
Software bugs; cyber security vulnerability
Misconfigured systems; cyber security vulnerability
```

Weak encryption; cyber security vulnerability
Unpatched software; cyber security vulnerability
Social engineering attacks; cyber security vulnerability
Phishing scams; cyber security vulnerability
Insider threats; cyber security vulnerability
Malware infections; cyber security vulnerability
Network vulnerabilities; cyber security vulnerability
Data leakage; cyber security vulnerability

Система ChatGPT у різні моменти під час обробки тексту може видавати різні варіанти відповідей. Кожну таку відповідь можна сприймати як відповідь деякого віртуального експерта, тому узагальнюючи відповіді такого «рою» експертів можна отримати повнішу відповідь. Об'єднані в одному CSV-файлі відповіді від ChatGPT завантажуються для аналізу та візуалізації у програму, розроблену авторами на основі бібліотеки GraphViz [4]. Інтерфейс програми, розміщений в Інтернеті за адресою http://bigsearch.space/uli.html, дозволяє ввести в поле введення отримані від ChatGPT відповіді та запустити візуалізацію спрямованої мережі.

У результаті 10-кратного виконання наведеного вище промпту було побудовано ієрархічну мережу з 77 вузлів.

Методика 2: первинна декомпозиція проблеми

Одна з основних процедур, що застосовуються у системному аналізі – це декомпозиція. У межах другої методики передбачається декомпозиція основного поняття кілька часткових понять (у межах цієї роботи реалізована декомпозиція на 10 часткових понять). Потім кожного часткового поняття формується однотипний промпт, що дозволяє визначати його основні причини. Авторами запропоновано наступний промпт декомпозиції для виявлення 10 часткових проблем:

→Decompose the concept of "cyber security vulnerability" into 10 partial concepts. Each partial concept must contain no more than three words. Present the answer in the form: "partial concept; cyber security vulnerability". Each entry on a separate line.

У результаті кожної з часткових проблем формується типовий промпт виявлення причини у межах загальної проблеми cyber security vulnerability, наприклад,

1 Reason Prompt: Name 10 reasons for the concept of " **Weak encryption** " as part of the concept of "cyber security vulnerability". Each reason must contain no more than three words. In the format "reason; **Weak encryption** ". Each entry on a separate line.

Отримані від ChatGPT відповіді об'єднуються, в результаті чого формується мережа, яка містить 102 вузли, що значно перевищує мережу, побудовану за методикою 1.

Поєднання результатів

На практиці, очевидно, що мережа, отримана шляхом логічного об'єднання мереж сформованих відповідно до методик 1 і 2 матиме переваги першого та другого підходу. Дійсно, ця мережа (Рис. 1) у предметній області, що розглядається, виявилася найбільш повною, що містить 162 вузли, серед яких 25 брали участь у формуванні мережі частіше 1 разу. Такі вузли, наведені нижче, можуть розглядатися як джерела подальшого розвитку каузальної мережі у предметній галузі, що розглядається.

Вузол	Частота
CYBER SECURITY VULNERABILITY	110
INSIDER THREATS	16
WEAK PASSWORDS	15
SOCIAL ENGINEERING	14
UNPATCHED SOFTWARE	14
PHISHING ATTACKS	14
ZERO-DAY EXPLOITS	13
WEAK ENCRYPTION	13
DATA BREACHES	12
CONFIGURATION ERRORS	12
MALWARE INFECTIONS	12
LACK OF ENCRYPTION	4
SOFTWARE BUGS	3
DATA LEAKAGE	3
UNPATCHED SYSTEMS	3
WEAK ACCESS CONTROLS	3
SOCIAL ENGINEERING ATTACKS	3

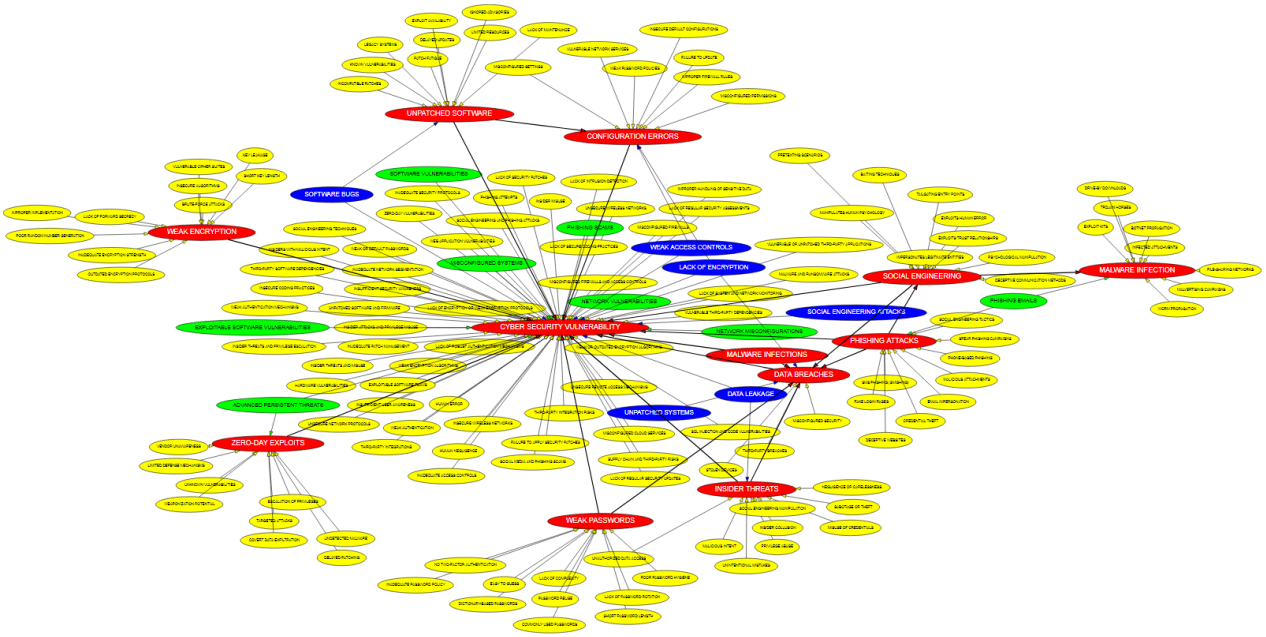

Рисунок 1. Мережа, отримана шляхом об'єднань методик 1 та 2.

Висновки

Запропоновано, продемонстровано та порівняно методи формування каузальних мереж причин на основі використання ChatGPT. Такі мережі можуть надалі використовуватися в рамках завдань системного та сценарного аналізу.

У рамках першої методики реалізовано концепцію віртуальних експертів, показано особливості мереж, сформованих за допомогою цього підходу. На підставі аналізу такої мережі можна зробити висновок, що вона охоплює не найбільшу кількість причин первинного поняття, які відносно слабко пов'язані, але, повторення тих самих причин у різних «віртуальних експертів» підтверджує їх точність, отже вони можуть бути непоганою «сировиною» для подальшої аналітичної обробки.

Друга мережа, отримана шляхом первинної декомпозиції проблеми та подальшого

виявлення причин кожної часткової проблеми, є більшою при однаковій кількості звернень до ChatGPT, як і в першому випадку. Це дає широке поле для дослідження, однак, при цьому повторюваність понять у такій мережі мінімальна, мережа близька до ієрархічної. Тому адекватність причин часткових проблем доводиться ретельно перевіряти експертним шляхом.

Незважаючи на суттєвий виграш у ресурсах (як часових, так і людських), важливо зазначити, що як сам процес побудови каузальних мереж, так і інтерпретація результатів, вимагають участі досвідченого експерта в предметній галузі, що вивчається.

Література

1. St. Wolfram. "What Is ChatGPT Doing ... and Why Does it Work?". – Wolfram Media, Inc. March 9, 2023. 112 p.
2. Lande, Dmitry and Strashnoy, Leonard. Concept Networking Methods Based on ChatGPT & Gephi (April 17, 2023). SSRN. Available at http://dx.doi.org/10.2139/ssrn.4420452
3. Lande, Dmitry and Strashnoy, Leonard. Hierarchical Formation of Causal Networks Based on ChatGPT (May 8, 2023). SSRN. Available at http://dx.doi.org/10.2139/ssrn. 4440629
4. Lambert M. Surhone, Mariam T. Tennoe, Susan F. Henssonow. Graphviz. VDM Publishing, 2010. 108 p.