НАЦІОНАЛЬНА АКАДЕМІЯ НАУК УКРАЇНИ
ІНСТИТУТ ПРОБЛЕМ РЕЄСТРАЦІЇ ІНФОРМАЦІЇ НАН УКРАЇНИ

ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ ТА БЕЗПЕКА

МАТЕРІАЛИ XXV МІЖНАРОДНОЇ НАУКОВО-ПРАКТИЧНОЇ КОНФЕРЕНЦІЇ

ВИПУСК 25

Київ – 2025

Рекомендовано до друку Вченою радою
Інституту проблем реєстрації інформації НАН України
(протокол № 9 від 23 грудня 2025 р.)

Інформаційні технології та безпека. Матеріали XXV Міжнародної науково-практичної конференції ІТБ-2025. – Київ: Інжиніринг. – 285 с. ISBN: 978-617-8180-04-1

До збірника увійшли матеріали доповідей, представлених на XXV Міжнародній науково-практичній конференції «Інформаційні технології та безпека» (ІТБ-2025, 11 грудня 2025 року, м. Київ, Україна).

У збірнику представлені матеріали, присвячені питанням створення і впровадження інформаційних технологій, актуальним проблемам забезпечення інформаційної та кібербезпеки, протидії інформаційним операціям і кібертероризму, інтелектуальним технологіям підтримки прийняття рішень, проведенню аналітичних досліджень на основі сучасних методів інтелектуального аналізу даних.

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

ISBN 978-617-8180-04-1

УДК 004.822: 004.912

Методологія аналізу потенціалу співавторства в кібербезпеці через структурно-семантичні гіперграфи на даних ArXiv

Вікторія Полуциганова[1] [0000-0002-9729-5786],
Сергій Смирнов[1] [0000-0003-4190-5204], Дмитро Ланде[1,2] [0000-0003-3945-1178],
Андрій Снарський[1,2] [0000-0002-4468-4542]
[1]*Національний технічний університет України "Київський політехнічний інститут імені Ігоря Сікорського"*
[2] *Institute for Information Recording of the National Academy of Sciences of Ukraine, Kyiv, Ukraine*
medvika@ukr.net, sergsmr@gmail.com, d.lande@kpi.ua, asnarskii@gmail.com

Анотація. У роботі запропоновано новий методологічний підхід до оцінки потенціалу наукового співавторства в галузі кібербезпеки, який базується на поєднанні структурного аналізу мереж співавторства з оцінкою тематичної близькості дослідників. На відміну від традиційних графових моделей, що враховують лише попарні взаємодії, запропонована методика використовує гіперграфи – математичні структури, здатні зберігати багатосторонні зв'язки, наприклад, одночасну участь кількох авторів у спільній публікації. Для повноцінного аналізу використовується двоїста модель: перший гіперграф відображає структуру співавторства, другий – семантичну близькість через ключові слова, витягнуті з анотацій публікацій. Це дозволило одночасно враховувати як історичний досвід співпраці, так і тематичну сумісність дослідників, навіть якщо вони ще не працювали разом. Розроблено комплекс метрик, що вимірюють частоту спільних публікацій, кількість посередників у мережі співавторства, часову динаміку публікаційної активності, а також спільність і вагомість тематичних термінів. На основі цих метрик обчислюється узагальнений індекс потенціалу співпраці для пари авторів.

***Ключові слова*:** гіперграфи, аналіз потенціалу співавторства, кібербезпека, ArXiv, структурно-семантичний аналіз, наукометрія, комплексні мережі.

Вступ

Сучасні конфлікти у сфері безпеки та оборони характеризуються гібридністю, поліконтекстністю та високою швидкістю трансформації. Традиційні методи аналізу — SWOT, PMESII, експертні оцінки – страждають від когнітивної інерції, низької адаптивності до нових типів загроз та відсутності строгих механізмів альтернативного мислення. Генеративні LLM, хоча й дозволяють масштабувати аналіз, схильні до галюцинацій, невиконання інструкцій та втрати пояснюваності. Необхідна система, яка формалізує альтернативне мислення, гарантує контроль якості та забезпечує інтуїтивно зрозуміле, але математично строге представлення конфлікту.

Наукова співпраця, зокрема, в галузі кібербезпеки, реалізується через спільні публікації, що охоплюють групи з багатьох авторів. Традиційні методи аналізу співавторства, засновані на графах, моделюють лише парні зв'язки, ігноруючи багатосторонні взаємодії [1]. Гіперграфи, навпаки, зберігають цілу структуру, представляючи кожну публікацію як гіперребро, що з'єднує всіх авторів. Це дозволяє точно аналізувати непрямі ланцюги співпраці через спільних співавторів.

Теоретичне підґрунтя

В основі моделі взаємодії авторів використовуються гіперграфи [2], адже вони перевершують бінарні графи в трьох ключових аспектах:

1. Моделювання багатосторонніх взаємодій у гіперграфах відображається точніше, тому що у графах:

- втрачається інформація про групову структуру;
- збільшується шум, оскільки парні зв'язки можуть бути випадковими;
- не моделюються непрямі зв'язки через спільних авторів, тоді як гіперграфи дозволяють обчислювати відстані через послідовність гіперребер.

Парні зв'язки в графах не не дозволяють аналізувати непрямі шляхи співпраці, оскільки кожна пара вершин з'єднується лише за наявності прямої спільної публікації. Гіперграфи визначають

відстань через послідовність гіперребер, що забезпечує точніший аналіз структурної близькості.

У гіперграфі кожна публікація p представляється як одне гіперребро e_p, що зберігає групову структуру.

2. Можливість аналізу співпраці через відстані в гіперребрах. Відстань між авторами A та B в гіперграфі H визначається як мінімальна кількість гіперребер для переходу від A до B:

$$d_H(A,B) = \min|\gamma|,\ \gamma=\{e1,...,en\},\ e_{i\cap}e_{i+1}=\emptyset.$$

Цей підхід відображає реальні ланцюги співпраці: якщо A працював з C, а C – з B, відстань $d_H(A,B)=2$, навіть якщо A і B не мають спільних публікацій..

3. Двоїста модель, що базується на гіперграфах дозволяє врахувати одночасно структурну та тематичну близькість:

- гіперграф співавторства H_C відповідає за історичну активність між авторами;
- гіперграф ключових слів H_K моделює тематичну схожість публікацій.

Методика розрахунку метрик гіперграфів

Методика оцінки потенціалу співпраці у майбутніх публікаціях складається з декількох етапів. 1-м з них є підготовка даних. Для того щоб оцінити можливість подальшої співпраці або використання історичної інформації про публікації, необхідно зібрати певний набір даних.

2-м етапом є побудова гіперграфів співавторства H_c, а саме:

- вершини якого – автори, які опублікували роботи в межах заданого набору даних;
- гіперребро, що об’єднує всіх її авторів, це публікація

Ці дані можна отримати автоматичним виділенням списків авторів з метаданих масивів публікацій, наприклад, з ArXiv.

3-м етапом є створення гіперграфу, який буде використовуватися в даній методології – це гіперграф ключових слів $H_к$. Відповідно компонентами його виступатимуть:

- вершини: автори;
- гіперребрами є ключові слова, витягнуті з абстрактів публікацій.

4-м етапом є виділення основних метрик гіперграфів, що у подальшому обрахунку наддасть оцінку потенціалу співпраці авторів в майбутніх роботах. Фактори, що впливають на оцінку:

1. Частота спільних публікацій.
2. Кількість співавторів у спільних публікаціях.
3. Часова близькість.
4. Структурна близькість в гіперграфі.

Запропонована методологія дозволяє аналізувати не лише потенціал співпраці авторів, які уже працювали разом, але й враховує. наскільки тематично сумісні автори, навіть якщо вони ще не мали прямої співпраці.

Висновки

Запропоновано та апробовано методологію оцінки потенціалу співавторства в кібербезпеці на основі двоїстої моделі гіперграфів – структурного та семантичного. На відміну від класичних графів, модель зберігає багатосторонні взаємодії та дозволяє прогнозувати співпрацю навіть за відсутності прямої історії спільних публікацій. Наукова новизна полягає в тому, що вперше запропоновано двоїсту гіперграфову модель, яка враховує структурні та семантичні зв'язки між авторами, зберігаючи багатосторонню природу наукової співпраці. Апробація на 5 000 публікаціях ArXiv підтвердила гіперболічний розподіл ймовірностей співпраці та точність прогнозів (понад 80%). Подальші дослідження спрямовані на розширення моделі до динамічних багатошарових гіперграфів.

Бібліографічні посилання

[1] D. Lande, A. Snarskii, et al. Network Analysis of Co-authorship and Research Directions in Thermoelectrics. (2025). Network Analysis of Co-authorship and Research Directions in Thermoelectrics. Journal of Thermoelectricity, (3), 71.91. DOI: 10.63527/1607-8829-2025-3-71-91

[2] Aksoy, S. G., et al. (2020). Hypernetwork science via high-order hypergraph walks. *EPJ Data Science*, 9(16). DOI: 10.1140/epjds/s13688-020-00231-0