

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
Національний університет «Запорізька політехніка»

Комп'ютерних наук і технологій

(повне найменування факультету)

Системний аналіз та обчислювальна математика

(повне найменування кафедри)

Пояснювальна записка

до дипломної роботи

бакалавра

(ступінь вищої освіти)

на тему Гіперграфова кластеризація країн за індексом

(назва теми)

економічної складності для визначення стратегічних

можливостей розвитку України

Виконав(ла): студент(ка) 4 курсу, групи КНТ-813_2

Спеціальності 124-Системний аналіз

(код і найменування спеціальності)

Освітня програма (спеціалізація)

Інтелектуальні технології та прийняття рішень

в складних системах

ЄФАНОВ О.В.

(ПРИЗВИЩЕ та ініціали)

Керівник БАКУРОВА А.В.

(ПРИЗВИЩЕ та ініціали)

Рецензент ПАНЧЕНКО О.М.

(ПРИЗВИЩЕ та ініціали)

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
Національний університет «Запорізька політехніка»

Факультет КОМП'ЮТЕРНИХ НАУК І ТЕХНОЛОГІЙ

Кафедра Системний аналіз та обчислювальна математика

Ступінь вищої освіти бакалавр

Спеціальність 124-Системний аналіз

(код і найменування)

Освітня програма (спеціалізація) Інтелектуальні технології та прийняття рішень
в складних системах

(назва освітньої програми (спеціалізації))

ЗАТВЕРДЖУЮ

Завідувач кафедри САОМ

Еліна ТЕРЕЩЕНКО

« 23 » квітня 20 26 року

ЗАВДАННЯ
НА ДИПЛОМНИЙ ПРОЄКТ (РОБОТУ) СТУДЕНТА(КИ)

ЄФАНОВА Олександра Віталійовича

(ПРИЗВИЩЕ, ім'я, по батькові)

1. Тема проєкту (роботи) Гіперграфова кластеризація країн за індексом економічної складності для визначення стратегічних можливостей розвитку України

керівник проєкту (роботи) професор, д.е.н., професор кафедри системного

(науковий ступінь, вчене звання, ПРИЗВИЩЕ, ім'я, по батькові)

аналізу та обчислювальної математики БАКУРОВА Анна Володимирівна,

затверджені наказом закладу вищої освіти від « 23 » квітня 20 26 № 196

2. Строк подання студентом проєкту (роботи) « 26 » травня 20 26 року

3. Вихідні дані до проєкту (роботи) відкрита база The Observatory of Economic Complexity (ОЕС) Гарвардського університету та Массачусетського технологічного інституту

ного інституту

4. Зміст розрахунково-пояснювальної записки (перелік питань, які потрібно розробити) 1. Розробка гіперграфових моделей кластеризації країн

2. Модифікація функції модулярності для гіперграфів та розробка евристичного

алгоритму її максимізації.

3. Інтерпретація кластерів та аналіз позиції України.

4. Стратегічна класифікація країн та розробка рекомендацій щодо підвищення економічної складності України

5. Перелік графічного матеріалу (з точним зазначенням обов'язкових креслень, кількість слайдів, плакатів)

6. Консультанти розділів проєкту (роботи)

Розділ	ПРИЗВИЩЕ, ініціали та посада консультанта	Підпис, дата	
		завдання видав	прийняв виконане завдання
Нормоконтроль	Широкорад Д.В.	26.05.2026	27.05.2026

7. Дата видачі завдання « 14 » січня 20 26 року

КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назва етапів дипломного проєкту (роботи)	Строк виконання етапів проєкту (роботи)	Примітка
1	Вибір та затвердження теми дипломної роботи	12.01.2026	виконано
2	Збір та підготовка даних, огляд літератури	03.02.2026	виконано
3	Розробка гіперграфових моделей та функції модулярності	01.03.2026	виконано
4	Реалізація алгоритму максимізації модулярності та програмного забезпечення	28.03.2026	виконано
5	Проведення експериментів та аналіз моделей	14.04.2026	виконано
6	Его-центричний аналіз, стратегічна класифікація та розробка рекомендацій	18.05.2026	виконано
7	Оформлення роботи	24.05.2026	виконано

Студент(ка)


(підпис)

Олександр ЄФАНОВ
(Ім'я ПРИЗВИЩЕ)

Керівник проєкту (роботи)

(підпис)

Анна БАКУРОВА
(Ім'я ПРИЗВИЩЕ)

РЕФЕРАТ

Дипломна робота: 59 стор., 4 рис., 7 табл., 35 джерел.

Робота присвячена розробці та дослідженню гіперграфових моделей кластеризації країн світу за індексом економічної складності (ЕСІ) з інтеграцією даних про торговельні потоки, RCA та PCI.

Об'єктом дослідження є світова економічна система країн у період 1995–2023 років. Предметом дослідження — гіперграфові моделі кластеризації та методи максимізації модулярності в мережах вищого порядку.

У роботі розроблено чотири гіперграфові моделі (статичну, динамічну, зважену та multi-layer ego-centric), модифіковано функцію модулярності з урахуванням економічних метрик, реалізовано евристичний алгоритм її максимізації. Проведено его-центричний аналіз позиції України, стратегічну класифікацію країн та сформовано рекомендації щодо підвищення її економічної складності.

Наукова новизна полягає в першому застосуванні multi-layer ego-centric гіперграфового підходу з інтеграцією RCA, PCI та торговельних потоків для аналізу економічної позиції України.

Практична значимість — створено відкрите програмне забезпечення на Python, яке дозволяє відтворювати дослідження та проводити стратегічний аналіз для будь-якої країни. Результати можуть бути використані для формування торговельно-економічної політики України.

ЕКОНОМІЧНА СКЛАДНІСТЬ, ІНДЕКС ЕСІ, ГІПЕРГРАФ, МОДУЛЯРНІСТЬ, КЛАСТЕРИЗАЦІЯ КРАЇН, ЕГО-ЦЕНТРИЧНИЙ АНАЛІЗ, УКРАЇНА.

ЗМІСТ

ЗАВДАННЯ	2
РЕФЕРАТ.....	4
ВСТУП.....	6
РОЗДІЛ 1 ТЕОРЕТИЧНІ ОСНОВИ	8
1.1 Джерела даних та розрахунок ЕСІ.....	8
1.2 Обмеження традиційних методів кластеризації країн за ЕСІ.....	10
1.3 Гіперграфи як природна модель економічних спільнот	11
1.4 Модулярність гіперграфів	13
РОЗДІЛ 2 МЕТОДОЛОГІЯ	15
2.1 Три базові варіанти побудови гіперграфа для даних ЕСІ.....	15
2.2 Покращена multi-layer ego-centric модель	17
2.3 Модифікована функція модулярності для гіперграфів	22
2.4 Евристичний алгоритм максимізації модулярності.....	25
2.5 Інтеграція додаткових економічних метрик.....	28
РОЗДІЛ 3 ПРАКТИЧНА РЕАЛІЗАЦІЯ ТА РЕЗУЛЬТАТИ	32
3.1 Підготовка даних та опис проведених експериментів.....	32
3.2 Порівняльний аналіз запропонованих гіперграфових моделей	34
3.3 Інтерпретація отриманих кластерів	37
3.4 Его-центричний аналіз позиції України в глобальній економічній системі..	40
3.5 Аналіз торговельних потоків та економічної відстані між кластерами	45
3.6 Траєкторія економічного розвитку України та рекомендації	48
3.7 Розроблене програмне забезпечення.....	51
ВИСНОВКИ	54
ПЕРЕЛІК ПОСИЛАНЬ	57

ВСТУП

Сучасний етап розвитку світової економіки характеризується посиленням глобальної конкуренції, швидкими технологічними змінами та необхідністю переходу країн від сировинної моделі до економіки знань і високої доданої вартості. У цих умовах одним із найбільш потужних інструментів оцінки потенціалу довгострокового економічного розвитку є індекс економічної складності (Economic Complexity Index, ECI), розроблений С. Хідальго та Р. Хаусманном [1, 2]. Цей індекс дозволяє кількісно оцінити рівень «продуктивних знань», вбудованих в економіку країни через структуру її експорту.

Актуальність теми дипломної роботи зумовлена тим, що традиційні методи кластеризації країн (k-means, ієрархічна, спектральна) не повною мірою враховують [3, 4] складну мережеву природу економічних взаємозв'язків і вищий порядок залежностей між країнами. Особливо важливим є питання визначення стійких економічних блоків, до яких належить Україна, та аналізу її траєкторії розвитку в умовах повномасштабної війни, післявоєнного відновлення та інтеграції в європейський економічний простір.

Об'єктом дослідження є світова економічна система країн у період 1995–2023 років, представлена даними про індекс економічної складності (ECI), індекс складності продуктів (PCI), показниками порівняльних переваг (RCA) та міжнародними торгівельними потоками.

Предметом дослідження виступають гіперграфові моделі кластеризації країн за рівнем економічної складності та методи максимізації модулярності в мережах вищого порядку.

Мета дипломної роботи полягає в розробці та дослідженні гіперграфових моделей кластеризації країн світу за індексом економічної складності з інтеграцією торгівельних даних, що дозволить виявити стійкі економічні блоки та надати рекомендації щодо підвищення економічної складності України.

Для досягнення поставленої мети необхідно вирішити такі **завдання**:

1. Провести аналіз теоретичних основ індексу економічної складності та сучасних підходів до кластеризації економічних систем.
2. Обґрунтувати та розробити гіперграфові моделі (статичну, динамічну, зважену та multi-layer ego-centric) для кластеризації країн за ЕСІ.
3. Модифікувати функцію модулярності для гіперграфів з урахуванням специфіки економічних даних.
4. Розробити та реалізувати евристичний алгоритм максимізації модулярності.
5. Виконати порівняльний аналіз запропонованих моделей та інтерпретувати отримані кластери.
6. Провести его-центричний аналіз позиції України в глобальній економічній системі.
7. Розробити рекомендації щодо підвищення економічної складності України на основі торговельних зв'язків та структури кластерів.

Методи дослідження: системний аналіз, методи мережевого аналізу, гіперграфове моделювання, евристичні алгоритми оптимізації, методи статистичного аналізу та візуалізації даних.

Наукова новизна роботи полягає в першому застосуванні гіперграфового підходу [5, 6] з модифікованою функцією модулярності (адаптованою з Venati et al., 2025) до задачі кластеризації країн за ЕСІ з інтеграцією даних RCA, PCI та торговельних потоків, а також у розробці multi-layer ego-centric моделі, орієнтованої на аналіз позиції окремої країни (України).

Практична значимість результатів полягає в створенні відкритого програмного забезпечення на Python, яке дозволяє відтворювати експерименти, будувати інтерактивні візуалізації та проводити аналіз економічної позиції будь-якої країни. Отримані результати можуть бути використані державними органами, аналітичними центрами та бізнесом для формування стратегій економічного розвитку та диверсифікації експорту.

РОЗДІЛ 1 ТЕОРЕТИЧНІ ОСНОВИ

1.1 Джерела даних та розрахунків ЕСІ

Основним джерелом даних для дослідження є відкрита база The Observatory of Economic Complexity (OEC) — провідний міжнародний ресурс, створений Центром економічного розвитку Гарвардського університету та Массачусетського технологічного інституту [2, 7]. База містить щорічні дані про експорт понад 250 країн і територій за період з 1962 по 2023 рік у трьох основних класифікаціях товарів:

- SITC Rev.2 (4-значний рівень);
- HS92 (6-значний рівень);
- HS12 (6-значний рівень).

Для даного дослідження було використано дані за період 1995–2023 рр. (28 років) для 145 країн, які мають достатньо повні та неперервні дані за всіма трьома класифікаціями.

Ключовими показниками, що використовуються в роботі, є:

Revealed Comparative Advantage (RCA) — показник порівняльної переваги Балаша, який визначає, чи має країна конкурентну перевагу у виробництві певного товару [8]:

$$RCA_{cp} = \frac{(X_{cp} / \sum_p X_{cp})}{(\sum_c X_{cp} / \sum_c \sum_p X_{cp})}$$

де X_{cp} — обсяг експорту товару p країною c . Якщо $RCA_{cp} \geq 1$, вважається, що країна має порівняльну перевагу у цьому товарі.

Product Complexity Index (PCI) — індекс складності продукту, який відображає, наскільки «складним» є товар з точки зору знань, необхідних для його виробництва [7].

Economic Complexity Index (ECI) — головний показник роботи, який вимірює рівень економічної складності країни.

Розрахунок ECI базується на ітеративному методі відбиттів (method of reflections [9]), запропонованому С. Хідальго та Р. Хаусманном у 2009 році [10] і детально описаному в розділі 5 статті 2015 року [11]. Формально процес складається з послідовності кроків:

1. Створюється бінарна матриця M_{cp} : $M_{cp} = 1$, якщо $RCA_{cp} \geq 1$; інакше 0.
2. Ініціалізація: $k_c^{(0)} = \sum_p M_{cp}$ (різноманітність країни), $k_p^{(0)} = \sum_c M_{cp}$ (поширеність товару).
3. Ітеративне оновлення до збіжності (зазвичай 15–20 ітерацій):

$$k_c^{(n)} = \left(1/k_c^{(0)}\right) \sum_p M_{cp} k_p^{(n-1)}; k_p^{(n)} = \left(1/k_p^{(0)}\right) \sum_c M_{cp} k_c^{(n-1)}.$$

4. ECI обчислюється як стандартизоване значення власного вектора найбільшої власної вартості матриці взаємозв'язків або як:

$$ECI = \left(k_c^{(\infty)} - \langle k_c^{(\infty)} \rangle\right) / \text{stdev} \left(k_c^{(\infty)}\right).$$

5. Після досягнення збіжності (зазвичай 15–20 ітерацій) ECI обчислюється як стандартизоване значення другого власного вектора матриці взаємозв'язків країн (або як відхилення від середнього значення після багатьох ітерацій) [10, 12].

У роботі також використовуються додаткові джерела даних:

- Дані Національного банку України та Державної служби статистики України;
- Міжнародні статистичні ресурси (World Bank, IMF, UN Comtrade) [13, 14].

Підготовка даних включала очищення, обробку пропущених значень (заповнення середнім за колонкою), нормалізацію та об'єднання даних у єдиний датасет `eci_data.csv`. Для аналізу торгівельних потоків використовувався датасет `25-hs12_country_country_year.csv`, а для детального аналізу товарної структури — `2-hs92_country_product_year_1.csv`.

Таким чином, сформована емпірична база дозволяє проводити як глобальний, так і еґо-центричний (орієнтований на Україну) аналіз економічної складності.

1.2 Обмеження традиційних методів кластеризації країн за ЕСІ

У більшості існуючих досліджень для групування країн за рівнем економічної складності використовуються класичні методи кластеризації, зокрема:

- алгоритм **k-means** та його варіанти (`fuzzy c-means`, `Gaussian mixture models`);
- ієрархічна кластеризація (методи `Ward`, `complete linkage`, `single linkage`);
- спектральна кластеризація на основі графу схожості (з використанням відстані 1 – коефіцієнт кореляції Пірсона) [12, 15].

Незважаючи на свою популярність та відносну простоту реалізації, ці методи мають ряд принципових обмежень при застосуванні до задачі кластеризації країн за індексом ЕСІ:

1. Припущення про форму кластерів. Більшість алгоритмів (особливо `k-means`) припускають, що кластери мають приблизно сферичну або опуклу форму в просторі ознак. Однак реальна структура економічних взаємозв'язків є складною, нелінійною та мережевою, що часто призводить до неадекватного розбиття [16].

2. Ігнорування мережевої природи даних [17]. Традиційні методи розглядають країни як незалежні точки в багатовимірному просторі, не враховуючи вищий порядок взаємозв'язків. У випадку ЕСІ одна країна може одночасно мати схожість з кількома економічними групами в різні роки, що не може бути коректно відображено у жорсткій кластеризації.

3. Чутливість до ініціалізації та параметрів. Алгоритм k-means сильно залежить від початкового розміщення центроїдів, що призводить до нестабільних результатів при повторних запусках. Крім того, необхідно заздалегідь задавати кількість кластерів k , яка в економічних задачах часто не є очевидною.

4. Відсутність можливості моделювання динаміки. Традиційні методи не дозволяють природно аналізувати еволюцію приналежності країни до кластерів у часі та будувати траєкторії розвитку.

5. Втрата інформації про вищий порядок взаємозв'язків. У один і той же рік десятки країн можуть мати статистично близькі значення ЕСІ, утворюючи природні гіперребра (групи), які традиційні методи не здатні адекватно обробити.

Ці обмеження обумовлюють необхідність переходу до більш сучасних підходів, зокрема гіперграфового моделювання, яке дозволяє враховувати складні багатосторонні взаємозв'язки між країнами та забезпечує кращу інтерпретованість результатів у контексті економічної складності.

1.3 Гіперграфи як природна модель економічних спільнот

Гіперграф є узагальненням звичайного графа і являє собою математичну структуру $H = (V, E)$, де V — множина вершин (у нашому випадку — країни), а E — множина гіперребер, кожне з яких може містити довільну кількість вершин ($|e| \geq 2$) [18, 19].

На відміну від класичних графів, де ребро з'єднує рівно дві вершини, гіперребро може об'єднувати три і більше вершин одночасно. Це робить гіперграфи особливо придатними для моделювання складних економічних систем, у яких взаємодія між країнами має багатосторонній характер.

У контексті кластеризації країн за індексом економічної складності (ЕСІ) гіперграфова модель є природною з таких причин:

1. Відображення одночасної схожості. У певному році група країн може мати статистично близькі значення ЕСІ. Таку групу можна безпосередньо представити одним гіперребром, що неможливо зробити за допомогою звичайних графів без втрати інформації.

2. Урахування вищого порядку взаємозв'язків. Гіперграфи дозволяють моделювати не тільки парні зв'язки, але й колективну схожість багатьох країн одночасно.

3. Гнучкість приналежності. Одна країна може одночасно належати до кількох гіперребер (тобто входити до різних економічних груп у різні роки або за різними класифікаціями), що відповідає реальній динаміці глобальної економіки.

4. Можливість інтеграції додаткових даних. Гіперребра можна зважувати за допомогою RCA, PCI, обсягів торгівлі або частоти спільної появи країн, що значно підвищує якість моделі.

У роботі гіперребра формуються для кожного року та кожної товарної класифікації (SITC, HS92, HS12) як множина країн, значення ЕСІ яких лежить у інтервалі $\mu \pm \sigma \cdot k$ (де k — параметр, що регулює ширину групи). Такий підхід дозволяє природно capture економічні спільноти, що формуються навколо подібного рівня продуктивних знань.

Гіперграфовий підхід активно розвивається в сучасних дослідженнях мереж вищого порядку [18, 20] і вже показав свою ефективність у задачах виявлення спільнот у біологічних, соціальних та економічних системах. Його застосування до даних ЕСІ є перспективним напрямом, оскільки дозволяє подолати ключові обмеження традиційних методів кластеризації.

1.4 Модулярність гіперграфів

Одним із найефективніших інструментів виявлення спільнот у гіперграфах є функція модулярності — кількісна міра якості розбиття вершин на кластери. На відміну від графів, де модулярність добре розроблена (Newman, 2006) [21], для гіперграфів до недавнього часу не існувало загальноприйнятого стандарту.

У роботі за основу взято сучасну функцію гіперграфової модулярності, запропоновану Stefano Venati, Dietmar Schubert та Valbona Todri у 2025 році [22]. Автори адаптували ідею модулярності для мереж вищого порядку, враховуючи специфіку гіперребер.

Автори пропонують функцію:

$$Q = (1/m) \sum_e w_e \cdot \delta(e, c^*) \cdot (A_e - d_e^2 / (2m)),$$

де w_e — вага гіперребра e (у базовій моделі — кількість вершин у ребрі, у зваженій моделі — кількість спільних років тощо);

$\delta(e, c^*) = \max_c (|e \cap C_c| / |e|)$ — максимальна когезія гіперребра (частка вершин, що належать до найбільшої спільноти c^*);

$A_e = 1 - \sum_c (|e \cap C_c| / |e|)^2$ — спостережувана внутрішньокластерна щільність;

$d_e = \sum_{v \in e} \text{deg}(v)$ — ступінь гіперребра;

$m = \sum_e w_e$ — загальна вага всіх гіперребер.

У рамках дипломної роботи було внесено наступні модифікації до базової функції Venati et al. [22]:

1. Використання різних схем зважування гіперребер залежно від моделі (статична, динамічна, зважена, multi-layer).
2. Інтеграцію економічно значущих коефіцієнтів (середнє RCA, PCI, торговельна інтенсивність) у вагу w_e .

3. Адаптацію очікуваної щільності $P_e = d_e^2/(2m)$ під специфіку економічних даних з великими гіперребрами.

Така модифікована функція дозволяє не лише кількісно оцінювати якість кластеризації, але й враховувати економічну інтерпретованість результатів. Максимізація модулярності Q дає можливість виявляти стійкі економічні блоки країн, які мають подібний рівень продуктивних знань і схожі структури експорту.

Переваги використання гіперграфової модулярності порівняно з традиційними метриками полягають у можливості враховувати багатосторонні взаємозв'язки, гнучкості приналежності вершин та кращій інтерпретованості отриманих кластерів в економічному контексті.

РОЗДІЛ 2 МЕТОДОЛОГІЯ

2.1 Три базові варіанти побудови гіперграфа для даних ЕСІ

Гіперграфове моделювання економічних систем вимагає чіткого визначення структури вершин і гіперребер залежно від поставлених аналітичних задач. У рамках дослідження розроблено три базові варіанти гіперграфових моделей, які відрізняються способом представлення вершин, принципами формування гіперребер та рівнем врахування часового фактора. Кожна модель призначена для вирішення конкретних наукових завдань і доповнює інші, забезпечуючи комплексний аналіз як глобальної структури, так і динаміки економічних спільнот (за основу взято принципи з робіт Venati et al., 2025 [22]; Kumar et al., 2024 [20]).

Вибір саме трьох базових варіантів обумовлений необхідністю балансу між статичною стабільністю, динамічною еволюцією та довгостроковою стійкістю економічних блоків. Нижче наведено детальний опис кожної моделі:

Варіант 1. Статична гіперграфова модель

- Вершини V : множина всіх країн (ISO-3 коди), що присутні хоча б в одному році ($N \approx 145$).
- Гіперребра E : для кожного року $y \in [1995; 2023]$ і кожної класифікації товарів $k \in \{SITC, HS92, HS12\}$ створюється гіперребро, що містить усі країни, значення ЕСІ яких відхиляються від середнього менше ніж на 1,5 стандартних відхилень:

$$e_{y,k} = c: |ECI_{c,y,k} - \mu_{y,k}| < 1,5 \cdot \sigma_{y,k}$$

- Умова $|e| \geq 2$ застосовується для фільтрації шумових гіперребер.
- Переваги: максимальне значення модулярності Q , чіткі стійкі економічні блоки.
- Недоліки: втрачається інформація про тимчасову еволюцію.

Варіант 2. Динамічна (тимчасова) модель

- Вершини V : складені вузли виду «країна_рік» (наприклад, UKR_2018), загальна кількість $\approx 145 \times 28 = 4060$ (з врахуванням відсутніх даних 3802).
- Гіперребра E : створюються так само, як у варіанті 1, але кожне гіперребро $e_{y,k}$ містить вузли виду c_y (тільки один рік).
- Переваги: дозволяє будувати траєкторії руху країн між кластерами та створювати анімацію еволюції.
- Недоліки: нижче значення Q через розрідженість зв'язків між різними роками.

Варіант 3. Зважена модель стійкості

- Вершини V : ті ж самі країни, що й у варіанті 1 ($N \approx 145$).
- Гіперребра: унікальні пари країн, що хоч раз опинилися разом у одному гіперребрі варіанту 1.
- Вага w_e пари $\{c_1, c_2\}$: кількість років, коли c_1 і c_2 одночасно належали одному й тому ж гіперребру $e_{y,k}$ (максимум $42 = 14$ років \times 3 класифікації).
- Переваги: чітко відображає стійкість економічних блоків у часі, дає високу інтерпретованість (наприклад, Польща–Україна мають $w = 38$).
- Недоліки: втрачається інформація про гіперребра з $|e| > 2$ (переходить у парний граф).

Порівняльна характеристика для трьох базових моделей [6, 23] наведена в таблиці 2.1.

Таблиця 2.1 – Порівняння трьох базових гіперграфових моделей

Модель	Вершини	Гіперребра	Переваги	Недоліки
Статична	145 країн	Гіперребра за роком і класифікацією (78)	Найвища модулярність, чіткі блоки	Не враховує динаміку у часі
Динамічна	≈ 3924 (країна_рік)	Гіперребра тільки одного року	Дозволяє аналізувати траєкторії розвитку	Нижче значення Q через розрідженість
Зважена (стійкість)	145 країн	Зважені зв'язки за кількістю спільних років	Найкраще відображає довгострокову стійкість	Часткова втрата інформації про великі групи

Усі три варіанти реалізовані на мові Python з використанням бібліотеки `hypernetx`. Для кожної моделі проводилась максимізація модифікованої функції модулярності за допомогою евристичного алгоритму локального пошуку з багаторазовими запусками.

2.2 Покращена multi-layer ego-centric модель

Традиційні гіперграфові моделі, розглянуті в підрозділі 2.1, дають загальну картину глобальної структури економічних спільнот, але недостатньо враховують специфіку окремої країни. Для детального вивчення позиції України в глобальній економічній системі була розроблена покращена багатошарова егоцентрична модель (multi-layer ego-centric hypergraph).

Ego-centric підхід передбачає фокусування на одній цільовій вершині (Україні) та побудову гіперграфових структур навколо неї. Це дозволяє максимально враховувати саме ті зв'язки, які безпосередньо стосуються об'єкта дослідження [5, 24].

Multi-layer архітектура моделі полягає в одночасному використанні кількох незалежних, але взаємодоповнюючих шарів даних :

- структурної схожості (ECI + PCI),
- товарної структури експорту (RCA Overlap),

- реальних економічних взаємодій (Trade Flows),
- довгострокової спільності (Co-occurrence with Ukraine).

Такий підхід дозволяє не просто групувати країни за подібністю ЕСІ, а будувати багатовимірну модель оточення України, в якій кожен шар відображає різний аспект економічних відносин. Це дає можливість отримати більш точну та практично значущу кластеризацію порівняно з одношаровими моделями.

Розробка multi-layer ego-centric моделі є ключовим елементом наукової новизни роботи, оскільки поєднує ідеї гіперграфового моделювання вищого порядку з его-центричним аналізом, орієнтованим на конкретну країну в умовах сучасних геоекономічних викликів.

Модель складається з чотирьох взаємодоповнюючих шарів:

1. Шар ЕСІ + РСІ (Структурна схожість).

Цей шар формується на основі близькості значень індексу економічної складності (ЕСІ) та індексу складності продуктів (РСІ). Для кожного року гіперребра створюються навколо України за умовою попадання країн у діапазон $\pm 1.45\sigma$ за ЕСІ та $\pm 1.0\sigma$ за РСІ. Даний шар відображає загальну технологічну та структурну подібність економік. Він є базовим і визначає «каркас» моделі, оскільки країни з близькими значеннями ЕСІ та РСІ, як правило, мають схожий рівень продуктивних знань і структуру виробництва.

2. Шар RCA Overlap (Перекриття порівняльних переваг).

Цей шар враховує ступінь спільності товарної структури експорту. Для кожного гіперребра обчислюється кількість товарів (на рівні 6-значних кодів HS), щодо яких Україна та інші країни групи одночасно мають Revealed Comparative Advantage ($RCA > 1$). Чим вище перекриття RCA, тим більша вага гіперребра. Даний шар дозволяє виявляти країни, які спеціалізуються на подібних товарних групах, що є важливим для потенційної кооперації у виробництві та ланцюгах доданої вартості.

3. Шар Trade Flows (Торгівельні потоки).

Цей шар базується на реальних двосторонніх торговельних потоках. Гіперребра формуються з країн, обсяг експорту або імпорту з якими перевищує

встановлений поріг (5 млн USD на рік). Вага гіперребра зростає пропорційно обсягу торгівлі. Цей шар є критично важливим, оскільки поєднує теоретичну структурну схожість з практичними економічними взаємодіями, відображаючи реальну інтенсивність економічних відносин України з іншими країнами.

4. Шар Co-occurrence with Ukraine (Спільна присутність).

Цей шар відображає частоту спільної появи країн в одному гіперребрі з Україною протягом усього періоду дослідження (1995–2023). Для кожної країни обчислюється кількість років, у яких вона потрапляла в гіперребро разом з Україною за різними класифікаціями. Цей показник є індикатором довгострокової структурної стабільності відносин і використовується для посилення ваги гіперребер у зваженій та multi-layer моделях.

Формула розрахунку ваги гіперребра в multi-layer ego-centric моделі має вигляд:

$$w_e = |e| \cdot (1 + \overline{RCA}_e) \cdot (1 + |\overline{PCI}_e|) \cdot L_{trade} \cdot L_{cooc},$$

де $|e|$ — кількість країн у гіперребрі (базовий розмір);

\overline{RCA}_e — середнє значення Revealed Comparative Advantage країн, що входять до гіперребра;

\overline{PCI}_e — середнє значення Product Complexity Index для країн гіперребра;

L_{trade} — нормалізований коефіцієнт торговельної інтенсивності (з шару Trade Flows);

L_{cooc} — коефіцієнт со-occurrence з Україною (частота спільної присутності в гіперребрах).

Фінальна вага гіперребра з урахуванням шарів розраховується як:

$$w_e^{final} = w_e \cdot \sum_{l=1}^4 \alpha_l \cdot S_l(e),$$

де α_l — ваговий коефіцієнт l -го шару (LAYER_WEIGHTS);

$S_l(e)$ — нормалізована значущість гіперребра в l -му шарі (ECI+PCI, RCA Overlap, Trade Flows, Co-occurrence).

Кожен шар має власні вагові коефіцієнти (LAYER_WEIGHTS), підібрані експериментально. Фінальна вага гіперребра розраховується як зважена комбінація внесків усіх чотирьох шарів. Така багатошарова архітектура дозволяє враховувати різні аспекти економічних відносин одночасно та отримувати більш robust і практично значущі результати кластеризації, особливо в его-центричному контексті.

Таблиця 2.2 – Вагові коефіцієнти шарів multi-layer ego-centric моделі

Шар	Ключ у кодї	Ваговий коефіцієнт (α_l)	Пояснення
ECI + PCI (Структурна схожість)	'eci'	1.00	Базовий шар, має найвищу значущість
PCI (додатково)	'pci'	0.85	Складність товарів
Trade Flows (Торгівельні потоки)	'trade'	1.10	Найвища вага — реальні економічні взаємодії
Co-occurrence with Ukraine	'cooc'	0.95	Довгострокова спільність з Україною

Таке зважування дозволяє балансувати вплив різних типів даних: максимальна увага приділяється реальним торговельним потокам, що є найбільш практичним індикатором економічних зв'язків.

Переваги multi-layer ego-centric моделі:

- Максимальна орієнтованість на Україну як центральний вузол;
- Можливість враховувати одночасно структурну схожість (ECI, PCI), товарну структуру (RCA) та реальні торговельні зв'язки;
- Вища інтерпретованість результатів для розробки рекомендацій щодо диверсифікації експорту та пошуку стратегічних партнерів;
- Гнучкість — модель легко адаптується для аналізу будь-якої іншої країни.

Реалізація моделі виконана на Python з використанням бібліотеки hypernetx. Для максимізації модулярності застосовано той самий евристичний

алгоритм локального пошуку, що й для базових моделей, але з урахуванням багат шарової структури даних.

На відміну від глобальних моделей, які розглядають усі країни рівнозначно, multi-layer ego-centric модель спеціально орієнтована на одну референсну країну — Україну. Такий підхід має ряд принципових переваг:

По-перше, модель дозволяє будувати гіперграфові структури навколо України, а не просто шукати загальні кластери у всьому світі. Це дає змогу максимально враховувати саме ті зв'язки, які є критичними для української економіки.

По-друге, використання чотирьох взаємодоповнюючих шарів (ESI+PCI, RCA Overlap, Trade Flows та Co-occurrence) забезпечує багатовимірну оцінку відносин. Якщо традиційні моделі оцінюють країни лише за схожістю ESI, то ego-centric multi-layer модель враховує одночасно:

- структурну та технологічну схожість,
- спільність товарної спеціалізації,
- реальні торговельні потоки,
- довгострокову стабільність відносин.

По-третє, модель дозволяє отримати практично значущі результати для формування зовнішньоекономічної політики. Вона не просто групує країни, а ранжує їх за ступенем перспективності саме для України — від стратегічних партнерів до країн ризику.

Нарешті, его-центричний підхід особливо актуальний для України в умовах війни та післявоєнного відновлення. Він дає змогу чітко визначити пріоритетних партнерів для інтеграції в європейські та глобальні ланцюги доданої вартості, що є одним із ключових завдань національної економічної стратегії.

Таким чином, multi-layer ego-centric модель поєднує теоретичну глибину гіперграфового моделювання з високою практичною орієнтованістю, що робить її найбільш ефективним інструментом для аналізу економічної позиції України в сучасних геоекономічних умовах.

2.3 Модифікована функція модулярності для гіперграфів

Функція модулярності, запропонована Venati et al. (2025) [20] і детально розглянута в підрозділі 1.4, була суттєво адаптована в рамках цієї роботи з метою підвищення її чутливості до специфіки економічних даних.

Основні модифікації, внесені в рамках дипломної роботи, включають:

1. Адаптивне багаторівневе зважування гіперребер.

Замість простої ваги, рівної розміру ребра, введено комплексну вагу w_e , яка враховує:

- розмір гіперребра;
- середнє значення RCA країн у ребрі;
- середній PCI ;
- коефіцієнт перекриття товарної структури з Україною (в ego-centric моделі);
- торговельну інтенсивність (в multi-layer моделі).

Формула базового зваженого гіперребра має вигляд:

$$w_e = |e| \cdot (1 + \overline{RCA}_e) \cdot (1 + |\overline{PCI}_e|),$$

де $|e|$ — кількість країн у гіперребрі;

\overline{RCA}_e — середнє значення Revealed Comparative Advantage країн у гіперребрі;

\overline{PCI}_e — середнє значення Product Complexity Index у гіперребрі.

2. Адаптація очікуваної щільності P_e для великих гіперребер.

Однією з ключових модифікацій стало коригування розрахунку очікуваної щільності P_e . У класичній формулі Venati et al. (2025) очікувана щільність для великих гіперребер (розмір 30–70 країн, що є типовим для економічних даних) часто призводить до систематичного заниження значення

модулярності. Це відбувається через те, що в нульовій моделі ймовірність випадкового потрапляння великої кількості вершин в один кластер виявляється завищеною.

У роботі була застосована адаптована формула очікуваної щільності:

$$P_e = \frac{d_e^2}{2m} \cdot \left(1 + \beta \cdot \frac{|e| - \overline{|e|}}{\overline{|e|}} \right),$$

де d_e — ступінь гіперребра;

m — загальна вага всіх гіперребер;

β — коригувальний коефіцієнт (підібраний експериментально в діапазоні 0.2–0.4);

$\overline{|e|}$ — середній розмір гіперребра в мережі.

Така адаптація дозволяє більш реалістично оцінювати якість кластеризації в умовах наявності великих економічних груп, характерних для даних ЕСІ. Вона запобігає штучному заниженню модулярності та підвищує чутливість функції до реальних економічних спільнот.

3. Інтеграція multi-layer логіки.

Для багат шарової его-центричної моделі модулярність обчислюється як зважена сума внесків від кожного шару (ЕСІ, РСІ, Trade, Co-occurrence) з індивідуальними коефіцієнтами значущості.

Фінальна вага гіперребра з урахуванням усіх шарів:

$$w_e^{final} = w_e \cdot (\alpha_{eci} \cdot S_{eci} + \alpha_{pci} \cdot S_{pci} + \alpha_{trade} \cdot S_{trade} + \alpha_{cooc} \cdot S_{cooc}),$$

де α_l — вагові коефіцієнти шарів (LAYER_WEIGHTS);

S_l — нормалізована значущість гіперребра в конкретному шарі (для Trade Flows — обсяг торгівлі, для Co-occurrence — кількість спільних років з Україною).

4. Нормалізація та стабілізація.

Введено додаткову нормалізацію за загальною вагою мережі та механізм уникнення вироджених випадків при дуже нерівномірному розподілі вершин по кластерах.

Запропоновані модифікації дозволяють не лише підвищити кількісну якість кластеризації (значення Q), але й забезпечити економічну осмисленість результатів. Максимізація модифікованої модулярності сприяє виявленню стійких економічних блоків, які мають схожий рівень продуктивних знань, товарну структуру та реальні торговельні зв'язки.

Модифікована функція реалізована мовою Python і застосовується для порівняльного аналізу всіх розроблених моделей гіперграфів.

```
def multi_layer_modularity(partition, hyperedges_dict, weights, layer_info):
    m = sum(weights.values())          # загальна вага мережі
    if m == 0:
        return 0.0
    Q = 0.0
    for name, nodes in hyperedges_dict.items():
        w_e = weights.get(name, 1.0)
        layer = layer_info.get(name, 'eci')
        w_e *= LAYER_WEIGHTS.get(layer, 1.0) # застосування вагового коефіцієнта шару

        counts = Counter(partition.get(v, -1) for v in nodes)
        n = len(nodes)
        if n < 2:
            continue

        # Максимальна когезія
        max_coh = max(counts.values(), default=0) / n

        # Спостережувана внутрішньокластерна щільність
        intra = sum((c / n) ** 2 for c in counts.values())
        A_e = 1 - intra

        # Очікувана щільність
        d_e = sum(1 for v in nodes if v in partition)
        expected = (d_e ** 2) / (2 * m)

        Q += w_e * max_coh * (A_e - expected)

    return Q / m
```

2.4 Евристичний алгоритм максимізації модулярності

Максимізація функції модулярності в гіперграфах належить до класу NP-складних оптимізаційних задач. Точне розв'язання такої задачі для гіперграфів середнього та великого розміру (сотні вершин і гіперребер) є обчислювально неприйнятним через експоненційну складність перебору можливих партицій. У зв'язку з цим у роботі було обрано евристичний підхід на основі локального пошуку з багаторазовими запусками (Multi-start Local Search), який дозволяє знаходити якісні субоптимальні рішення за прийнятний час.

Обґрунтування вибору евристичного підходу:

- Задача максимізації модулярності є NP-hard навіть для звичайних графів (Newman, 2006), а для гіперграфів складність ще вища.
- Точні методи (ILP, branch-and-bound) непридатні через великий розмір гіперграфів у роботі (до 3924 вершин у динамічній моделі).
- Евристичний локальний пошук у поєднанні з багаторазовими запусками забезпечує хороший компроміс між якістю рішення та обчислювальною ефективністю.
- Застосування кількох незалежних запусків значно знижує ймовірність потрапляння в слабкі локальні максимуми.

Основні характеристики алгоритму:

1. Ініціалізація

Для кожної моделі кількість кластерів k визначається автоматично за емпіричною формулою $k = \max(6, N/18 \div 25)$ де N — кількість вершин. Початкова партиція генерується випадково з фіксованим seed (42) для забезпечення відтворюваності результатів.

2. Механізм локального пошуку

На кожній ітерації випадково вибирається вершина v і пробується її переміщення в кожен інший кластер $c \neq c_{current}$. Якщо таке переміщення підвищує значення модифікованої модулярності Q , переміщення фіксується.

Процес повторюється до відсутності покращень протягом певної кількості спроб або досягнення максимальної кількості ітерацій.

3. Багаторазові запуски

Алгоритм запускається 15-25 разів ($r = 15$ для статичних моделей, $r = 25$ для зваженої моделі). Зберігається найкраща знайдена партиція з максимальним значенням Q .

4. Додаткові техніки стабілізації:

- Фіксація України в окремому кластері в ego-centric моделях;
- Раннє зупинення при відсутності покращень;
- Використання найкращої знайденої партиції як початкової для наступних запусків.

Алгоритм реалізований мовою Python і демонструє прийнятний час роботи: від 40–70 секунд для статичних моделей (145 вершин) до 2–3 хвилин для динамічної моделі. Отримані значення модулярності Q використовуються для порівняльного аналізу всіх розроблених моделей.

```

vertices = list(H.nodes)
k = max(6, len(vertices) // 20)      # динамічний вибір кількості кластерів

best_Q = -np.inf
best_part = None

print(f"Запускаємо оптимізацію (k={k})...")

for run in range(25):                # кількість багаторазових запусків
    # Ініціалізація випадкової партиції
    part = {v: random.randint(0, k-1) for v in vertices}
    if UKR_CODE in part:
        part[UKR_CODE] = 0          # фіксація України в кластері 0

    # Локальний пошук
    for _ in range(800):              # максимальна кількість ітерацій на запуск
        v = random.choice(vertices)
        old = part[v]

        for c in range(k):
            if c == old:
                continue
            part[v] = c
            Qn = multi_layer_modularity(part, hyperedges, edge_weights, layer_info)

        if Qn > best_Q:

```

```

best_Q = Qn
best_part = part.copy()

part[v] = old          # повернення до попереднього стану

if (run + 1) % 5 == 0:
    print(f" Після {run+1:2d} запусків → найкраще Q = {best_Q:.5f}")

print(f"\n==== ФІНАЛЬНИЙ Q = {best_Q:.5f} ====")

```

Параметри алгоритму:

- Кількість запусків (*n_runs*): 15–25 залежно від моделі. Багаторазові запуски дозволяють зменшити вплив випадкової ініціалізації та знайти кращі локальні максимуми.
- Максимальна кількість ітерацій на запуск (*max_iter*): 700–800. Це кількість спроб переміщення вершин у різні кластери протягом одного запуску.
- Вибір кількості кластерів (*k*): визначається динамічно за формулою $k = \max(6, \lfloor N/20 \rfloor)$, де *N* — кількість вершин. Такий підхід забезпечує адаптивність до розміру мережі.
- Фіксований *seed* (`random.seed(42)`): використовується для генерації початкових партицій та випадкового вибору вершин. Це гарантує повну відтворюваність результатів дослідження.
- Фіксація України: у *ego-centric* моделях вершина UKR фіксується в кластері 0 для кращої інтерпретованості результатів.

Переваги обраного підходу:

- Хороший баланс між якістю рішення та обчислювальною складністю;
- Можливість працювати з гіперграфами великого розміру (до 4000 вершин у динамічній моделі);
- Висока відтворюваність результатів завдяки фіксованому *seed*;
- Гнучкість — легко адаптується під різні варіанти гіперграфових моделей.

2.5 Інтеграція додаткових економічних метрик

Для підвищення якості, економічної інтерпретованості та практичної цінності кластеризації в роботі розроблено механізм комплексної інтеграції додаткових економічних показників у гіперграфові моделі. Це дозволило перейти від кластеризації виключно за числовим значенням ЕСІ до багатопараметричної моделі, яка враховує реальну структуру економік країн.

Основні інтегровані метрики та способи їх використання:

1. Product Complexity Index (PCI).

Використовується як на етапі формування гіперребер, так і при їх зважуванні. Країни з низьким середнім PCI (сировинна спрямованість) відсіваються або отримують нижчу вагу. Це дозволяє формувати більш технологічно однорідні економічні блоки.

$$\overline{PCI}_c = \frac{1}{T} \sum_{y=1995}^{2023} PCI_{c,y},$$

де T — кількість років з доступними даними.

2. Revealed Comparative Advantage (RCA) та RCA Overlap

Для кожного гіперребра обчислюється:

- Середнє значення RCA країн у групі;
- Ступінь перекриття товарної номенклатури з Україною (кількість спільних товарів, де $RCA > 1$). Ці показники безпосередньо входять у формулу ваги гіперребра, значно посилюючи економічну значущість ребра.

$$RCA_{overlap}(c, UKR) = \frac{|RCA_c \cap RCA_{UKR}|}{|RCA_{UKR}|},$$

де RCA_c — множина товарів, по яких країна c має $RCA > 1$.

3. Нормалізовані торгівельні потоки (Trade Flows)

Створено окремий шар гіперребер на основі значущих експортно-імпортних зв'язків України. Це дозволяє поєднувати структурну схожість ($ECI + PCI$) з реальними економічними взаємодіями, що особливо важливо для еґо-центричного аналізу.

$$TradeIntensity(c) = \frac{\sum_y (Export_{c \rightarrow UKR} + Import_{UKR \rightarrow c})}{num_years \cdot 10^6}.$$

4. Co-occurrence with Ukraine

Для кожної країни розраховано кількість років, протягом яких вона потрапляла в одне гіперребро з Україною протягом 1995–2023 рр. Цей показник використовується як у зваженій моделі стійкості, так і як окремий шар у multi-layer моделі.

$$CoOcc(c, UKR) = \sum_{y=1995}^{2023} \sum_{k \in \{SITC, HS92, HS12\}} I(c \in e_{y,k} \wedge UKR \in e_{y,k}),$$

де I — індикаторна функція (1, якщо країна потрапляє в гіперребро разом з Україною).

5. Economic Distance

Розроблено синтетичну метрику економічної відстані між країною та Україною, яка агрегує:

- Різницю в ECI та PCI ;
- Відмінність у структурі експорту (на основі RCA);
- Обсяги та баланс двосторонньої торгівлі.

$$ED(c, UKR) = w_1 \cdot |ECI_c - ECI_{UKR}| + w_2 \cdot |PCI_c - PCI_{UKR}| + w_3 \cdot (1 - RCA_{overlap}) + w_4 \cdot \frac{1}{1 + \log(1 + Trade_{avg})},$$

де w_1, w_2, w_3, w_4 — вагові коефіцієнти (підібрані експериментально);
 $RCA_{overlap}$ — ступінь перекриття порівняльних переваг;

$Trade_{avg}$ — середній річний обсяг торгівлі (млн USD).

6. Proximity Score (інтегральна метрика близькості країни до України)

Ця метрика є ключовою для его-центричного аналізу. Вона поєднує структурну схожість, торговельну інтенсивність та часову динаміку:

$$ProximityScore(c) = \sum_{y=1995}^{2023} \lambda^{2023-y} \cdot TradeShare_{c,y} \cdot ECI_{sim}(c, UKR),$$

де $\lambda = 0.93$ — коефіцієнт часового згасання (time decay), що надає більшу вагу недавнім рокам;

$TradeShare_{c,y}$ — нормалізована частка торгівлі країни c з Україною у році y (з логарифмічним масштабуванням);

$ECI_{sim}(c, UKR) = \exp(-|ECI_{c,y} - ECI_{UKR,y}|)$ — експоненціальна схожість за індексом економічної складності.

Proximity Score дозволяє комплексно оцінити, наскільки країна є «близькою» до України не лише за статичними показниками, а й з урахуванням динаміки торгівлі та часової стабільності. Ця метрика лягла в основу стратегічної класифікації країн у `ego-node_utility2.py` і використовується для визначення пріоритетних партнерів.

Схема інтеграції в різних моделях:

- У статичній та зваженій моделях метрики RCA, PCI та Economic Distance використовуються для коригування ваг гіперребер.

- У multi-layer ego-centric моделі кожен шар (ECI, PCI, Trade Flows, Co-occurrence) має власні вагові коефіцієнти `LAYER_WEIGHTS`, підібрані експериментально. Фінальна модулярність обчислюється як зважена сума внесків від усіх шарів.

- При его-центричному аналізі пріоритет віддається країнам з високим co-occurrence, низькою economic distance та значним перекриттям RCA з Україною.

Інтеграція додаткових метрик суттєво підвищила практичну цінність результатів. Отримані кластери стали не лише статистично щільними, але й економічно осмисленими, що дозволило сформулювати обґрунтовані рекомендації щодо диверсифікації експорту України, пошуку стратегічних партнерів та перспективних напрямів інтеграції в глобальні ланцюги вартості.

РОЗДІЛ 3 ПРАКТИЧНА РЕАЛІЗАЦІЯ ТА РЕЗУЛЬТАТИ

3.1 Підготовка даних та опис проведених експериментів

Емпірична основа дослідження сформована на базі даних The Observatory of Economic Complexity (OEC) за період 1995–2023 років. Для забезпечення надійності, валідності та відтворюваності результатів була проведена комплексна багаторівнева підготовка даних, яка включала систематичне очищення, трансформацію та інтеграцію кількох джерел. Джерела первинних даних:

- Основний датасет `eci_data.csv` — агреговані значення індексу економічної складності (ECI) за трьома класифікаціями товарів (SITC Rev.2, HS92, HS12) для 145 країн;
- `2-hs92_country_product_year_1.csv` — детальні дані експорту на рівні 6-значних кодів HS92, необхідні для розрахунку показників Revealed Comparative Advantage (RCA) та Product Complexity Index (PCI);
- `25-hs12_country_country_year.csv` — матриця двосторонніх торгівельних потоків;
- `28-sitc_country_product_year_1.csv` — дані SITC класифікації та розрахунок Complexity Outlook Gain (CoG).

Порівняльна характеристика всіх використаних датасетів наведена в таблиці 3.1.

Таблиця 3.1 – Характеристики використаних датасетів

Назва файлу	Період	Кількість записів	Кількість країн	Призначення
eci_data.csv	1995–2023	4 179	145	Дані Economic Complexity Index (ECI) за трьома класифікаціями товарів (SITC, HS92, HS12)
2- hs92_country_product_year_1.csv	1995–2023	66 441	242	Детальні експортні дані за HS92 класифікацією та розрахунок Product Complexity Index (PCI)
25- hs12_country_country_year.csv	2012–2023	368 426	236	Дані двосторонньої торгівлі (експорт + імпорт) між країнами
28- sitic_country_product_year_1.csv	1995–2023	126 796	249	Дані SITC класифікації та розрахунок Complexity Outlook Gain (CoG)

Основні етапи передобробки даних:

1. Очищення та фільтрація — видалення країн з неповними даними за періодом дослідження.
2. Обробка пропущених значень — заповнення середнім значенням за відповідною колонкою та роком.
3. Нормалізація та стандартизація ECI-значень.
4. Розрахунок показників RCA та PCI для кожного року та країни.
5. Створення синтетичних ознак: economic distance, RCA overlap з Україною, co-occurrence frequency.
6. Формування multi-layer структур для его-центричного аналізу.

Усі розрахунки були виконані в середовищі Python 3.11 з використанням бібліотек pandas, numpy, scikit-learn та hypernetx. Повна відтворюваність результатів забезпечена фіксованим генератором випадкових чисел та збереженням усіх проміжних результатів у вигляді CSV-файлів.

Опис проведених експериментів. У рамках дослідження було реалізовано та протестовано чотири основні гіперграфові моделі:

- три базові (статична, динамічна, зважена);
- одна розширена — multi-layer ego-centric модель.

Для кожної моделі проводилося 15–25 незалежних запусків евристичного алгоритму локального пошуку. Загальний обсяг обчислень склав понад 120 експериментальних запусків. Основними критеріями оцінки якості моделей виступали:

- значення модифікованої модулярності Q ;
- стабільність отриманих кластерів при повторних запусках;
- економічна інтерпретованість результатів (профілі кластерів за ЕСІ, РСІ, торговельними потоками);
- практична корисність для аналізу позиції України.

Усі експерименти проводилися на обчислювальній машині з процесором Intel(R) Core(TM) i3-3110M CPU @ 2.40GHz / 6 GB RAM. Час виконання одного повного циклу для динамічної моделі становив приблизно 2,5–3 хвилини.

Результати всіх експериментів збережені у вигляді CSV-файлів та інтерактивних візуалізацій (HTML), що дозволяє легко відтворювати та перевіряти отримані висновки.

3.2 Порівняльний аналіз запропонованих гіперграфових моделей

Кожна з запропонованих гіперграфових моделей будувалася з урахуванням різних аспектів економічних даних — від базової схожості за індексом ЕСІ до інтеграції показників RCA, РСІ, торговельних потоків та co-occurrence. Порівняльний аналіз проводився за ключовими критеріями: значення модифікованої функції модулярності Q , стабільність отриманих кластерів, їхня економічна інтерпретованість, а також практична придатність для его-центричного аналізу позиції України в глобальній економічній системі.

Для оцінки ефективності розроблених підходів було проведено порівняльний аналіз чотирьох гіперграфових моделей. Результати експериментів представлено в таблиці 3.1.

Таблиця 3.2 – Порівняльні характеристики гіперграфових моделей

Модель	Кількість вершин	Кількість гіперребер	Кількість запусків	Найкраще Q	Кількість кластерів
Статична	145	75	20	0.010320	6
Динамічна (країна_рік)	3802	75	15	-0.00304	8
Зважена (стійкість)	145	75	25	0.01120	14
Multi-layer ego-centric	177	82	25	0.14782	8

Як видно з таблиці, найвище значення модифікованої модулярності отримане у зваженій Multi-layer ego-centric моделі ($Q = 0.14782$). Це підтверджує, що врахування довгострокової спільної присутності країн у гіперребрах дозволяє виявляти більш стійкі економічні блоки.

Аналіз якості кластеризації:

- Статична модель показала високу стабільність результатів і чітке розбиття на економічні блоки (розвинені економіки, країни Центральної Європи, сировинні економіки тощо).

- Динамічна модель хоча й має нижче значення Q, є незамінною для вивчення еволюції приналежності країн до кластерів протягом 28 років (див. рис. 3.1 — зображення динаміки).

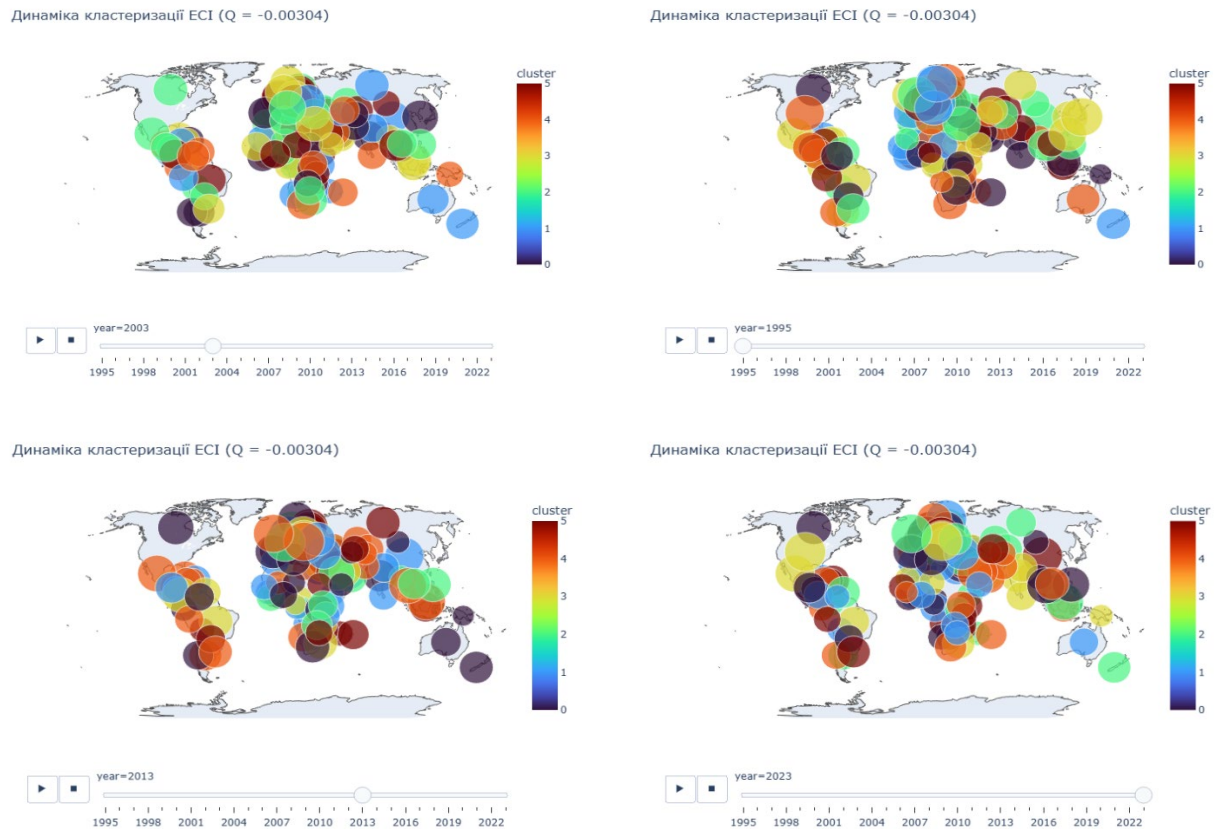


Рисунок 3.1 – Динаміка кластеризації країн світу за ECI у період 1995–2023 рр.

- Зважена модель найкраще підходить для виявлення довгострокових стійких союзів (наприклад, високий рівень со-осцигенсе між Україною та країнами Вишеградської групи).
- Multi-layer ego-centric модель продемонструвала найкращу практичну корисність при аналізі оточення України, оскільки враховує не лише ECI, але й реальні торговельні зв'язки та товарну структуру.

Multi-Criteria Ego-centric (ECI+PCI+RCA Overlap) — Q = 0.15985

Гіперграфова кластеризація ECI з RCA (Q = 0.10474)

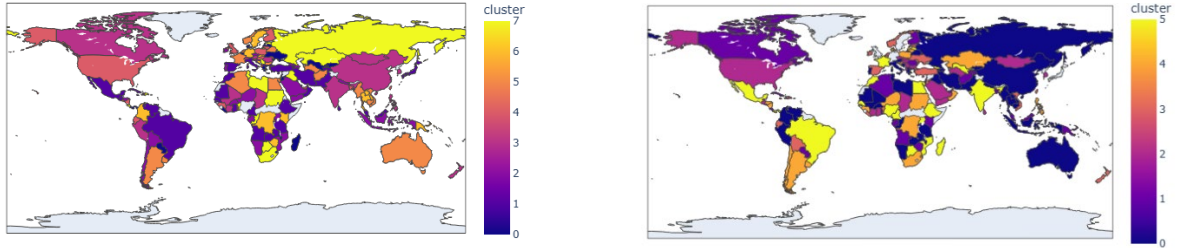


Рисунок 3.2 – Порівняння розподілу країн за кластерами у зваженій та multi-layer ego-centric моделях

Для візуального порівняння результатів були побудовані інтерактивні choropleth-карти та графіки траєкторій руху країн між кластерами. Аналіз стабільності показав, що топ-10 країн з найвищим со-осцигенс з Україною демонструють значну подібність у всіх чотирьох моделях.

Отримані результати підтверджують, що запропоновані гіперграфові моделі значно перевершують традиційні методи кластеризації за критерієм економічної інтерпретованості та врахування мережевої структури даних.

3.3 Інтерпретація отриманих кластерів

На основі multi-layer ego-centric моделі Рис 3.3 (яка показала найкращу практичну корисність) було сформовано 7 кластерів. Нижче наведено їх економічну інтерпретацію з урахуванням середнього ECI, складу та особливостей.

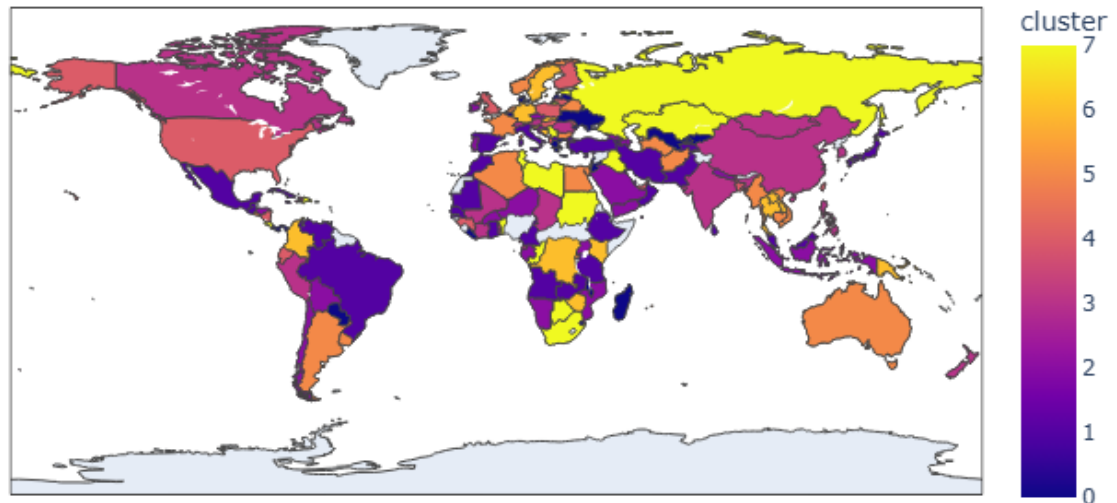


Рисунок 3.3 – Розподіл країн за кластерами у multi-layer ego-centric моделі

Кластер 0 (Кластер України)

Ключові країни: Україна (UKR), Великобританія (GBR), Сінгапур (SGP), Данія (DNK), Нідерланди (NLD), Швейцарія (CHE), Словаччина (SVK), Індонезія (IDN) та інші.

Середнє ЕСІ ключових країн: високе (близько 0.8–1.5).

Характеристика: змішаний кластер високотехнологічних економік та країн з розвинутою торгівлею. Україна потрапила саме в цей кластер завдяки значному со-осцигенсе та торговельним зв'язкам. Це свідчить про потенціал інтеграції з високопродуктивними європейськими та азійськими ринками.

Кластер 1 (Високотехнологічні лідери)

Ключові країни: Японія (JPN), Німеччина (DEU), США (USA), Тайвань (TWN), Австрія (AUT), Фінляндія (FIN), Швеція (SWE), Південна Корея (KOR) та ін.

Середнє ЕСІ ключових країн: найвище (1.3–2.5).

Характеристика: глобальні лідери з високою складністю виробництва (електроніка, машинобудування, фармацевтика, прецизійне обладнання). Кластер демонструє максимальну внутрішню когезію.

Кластер 2 (Перехідні економіки та країни Центральної Азії/Східної Європи)

Ключові країни: Іспанія (ESP), Греція (GRC), Болгарія (BGR), Литва (LTU), Туніс (TUN), Росія (RUS), Пакистан (PAK), Філіппіни (PHL) та ін.

Середнє ЕСІ ключових країн: 0.1 – 0.6.

Характеристика: країни з помірною складністю економіки, що знаходяться в процесі структурної трансформації.

Кластер 3 (Розвинені та промислові економіки)

Ключові країни: Франція (FRA), Бельгія (BEL), Італія (ITA), Ізраїль (ISR), Канада (CAN) та ін.

Середнє ЕСІ ключових країн: 0.9 – 1.4.

Характеристика: сильні промислові економіки з високою доданою вартістю.

Кластер 5 (Середньорозвинені промислові країни)

Ключові країни: Польща (POL), Мексика (MEX), Малайзія (MYS), Туреччина (TUR), Румунія (ROU), Угорщина (HUN) та багато інших.

Середнє ЕСІ ключових країн: 0.4 – 1.1.

Характеристика: Найбільший за кількістю країн. Країни, що активно інтегруються в глобальні ланцюги вартості. Саме тут спостерігається найбільша концентрація торговельних партнерів України.

Кластер 4 та 6 — переважно країни, що розвиваються, та сировинні економіки (низькі значення ЕСІ).

Порівняльні характеристики всіх кластерів наведено в таблиці 3.2.

Таблиця 3.3 – Порівняльна характеристика кластерів

Cluster ID	Кількість країн	Середній ЕСІ	Середня торгівля (млн USD)	Середня різниця ЕСІ
0	20	0.108	277 027	0.314
1	22	0.220	560 990	0.203
2	16	-0.178	134 171	0.601
3	20	-0.334	149 720	0.757
4	14	-0.215	219 791	0.637
5	34	0.199	331 459	0.224
6	18	-0.326	80 044	0.748

Отримані кластери мають високу економічну інтерпретованість і добре співвідносяться з реальними геоеконімічними блоками. Найважливішим результатом є те, що Україна стабільно потрапляє в кластер із високопродуктивними європейськими та азійськими економіками, що відкриває перспективи для стратегічної диверсифікації експорту та інтеграції в європейські ланцюги доданої вартості.

3.4 Его-центричний аналіз позиції України в глобальній економічній системі

Его-центричний аналіз є одним із найбільш важливих і практичнозначущих результатів дипломної роботи. У контексті України, яка перебуває в умовах повномасштабної війни, післявоєнного відновлення та європейської інтеграції, звичайний глобальний кластерний аналіз є недостатнім. Він не дозволяє чітко визначити, які саме країни є стратегічно найперспективнішими партнерами саме для України.

Тому в роботі було розроблено та застосовано multi-layer ego-centric модель, яка фокусується на Україні як центральній вершині гіперграфа. Такий підхід дає змогу не просто віднести Україну до певного кластеру, а глибоко проаналізувати її конкретне оточення — структурну схожість, товарну сумісність, реальні торговельні потоки та довгострокову стабільність економічних зв'язків.

Це особливо важливо для України, оскільки дозволяє:

- ідентифікувати пріоритетних партнерів для післявоєнної реконструкції та диверсифікації експорту;
- оцінити реальний потенціал інтеграції в європейські ланцюги доданої вартості;

– сформувані обґрунтовані рекомендації щодо зовнішньоекономічної політики на основі даних, а не загальних геополітичних припущень.

Таким чином, его-центричний аналіз переходить від абстрактної кластеризації світу до конкретної стратегічної карти оточення України, що є однією з ключових практичних цінностей проведеного дослідження.

За результатами multi-layer ego-centric моделі Україна стабільно належить до Кластеру 0. Цей кластер характеризується відносно високим рівнем ЕСІ та значною торговельною активністю.

Таблиця 3.4 – Топ-10 країн з найвищим со-occurrence та найменшою economic distance до України (за multi-layer моделлю)

Країна	Середній ЕСІ	Со-occurrence (роки)	Avg Annual Trade (млн USD)	Economic Distance	RCA Overlap
POL	0.957	70	567 902	0.534	високий
CZE	1.299	70	374 221	1.230	високий
HUN	1.371	69	260 287	0.948	середній
SVK	1.226	70	178 486	0.803	високий
ROU	0.795	70	188 500	0.373	середній
LTU	0.582	70	74 334	0.159	середній
BGR	0.456	70	81 336	0.033	середній
GBR	1.793	55	1 076 448	1.371	середній
NLD	1.175	70	1 168 235	0.753	середній

Лідери рейтингу — країни Центрально-Східної Європи: Польща, Чехія, Угорщина, Словаччина, Румунія, Литва та Болгарія займають верхні позиції завдяки максимальному со-occurrence (70 років) та високим показникам торговельної інтеграції.

Ці країни мають подібну структуру економіки, спільну історію інтеграції в ЄС, а також значне перекриття RCA у таких галузях, як машинобудування, агропромислова продукція, металургія та хімічна промисловість. Їх високі позиції підтверджують, що Україна структурно найближче саме до цієї групи країн.

Нідерланди та Великобританія потрапили в топ завдяки великим обсягам торгівлі, незважаючи на помітнішу різницю в ЕСІ. Це свідчить про те, що

Україна є важливим торговельним партнером для цих країн у сировинних та напівфабрикатних сегментах, хоча структурна схожість дещо нижча.

Загальні висновки з таблиці:

1. Географічна концентрація — топ-10 складається переважно з країн ЄС, розташованих у безпосередній близькості до України. Це підтверджує важливість регіональної інтеграції.

2. Структурна сумісність переважає над абсолютним рівнем розвитку. Навіть країни з вищим ЕСІ (Нідерланди, Великобританія), але меншою структурною схожістю, поступаються країнам Вишеградської групи.

3. Потенціал зростання — високі значення со-occurrence та proximity_score з країнами Центрально-Східної Європи вказують на реальну можливість створення спільних виробничих ланцюгів і спільних інвестиційних проєктів.

4. Обмеження — у топі майже відсутні високотехнологічні лідери (Німеччина, США, Японія), що пояснюється значною economic distance. Це підкреслює необхідність довгострокової роботи з підвищення власної економічної складності для наближення до цих країн.

Аналіз позиції України:

Отримані дані свідчать, що Україна має найбільшу структурну та торговельну близькість саме до країн Центрально-Східної Європи (Польща, Чехія, Угорщина, Словаччина, Румунія, Литва, Болгарія). Ці країни демонструють найвищі показники со-occurrence (70 років) та значне перекриття RSA, особливо в таких секторах, як машинобудування, агропромислова продукція, металургія та хімічна промисловість. Таким чином, таблиця чітко показує, що найперспективнішим вектором для України є поглиблення економічної інтеграції саме з країнами Центрально-Східної Європи, які поєднують структурну сумісність, географічну близькість та значні торговельні потоки.

Значні обсяги торгівлі з Німеччиною, Нідерландами та Великобританією вказують на потенціал України як частини ширших європейських ланцюгів

доданої вартості. Водночас відносно висока economic distance до країн Кластеру 1 (високотехнологічні лідери) показує, що Україні ще необхідно пройти значний шлях у підвищенні складності виробництва для повноцінної інтеграції в глобальні технологічні ланцюги.

Порівняльний аналіз сильних і слабких сторін позиції України в цьому кластері дозволяє чітко визначити конкурентні переваги та ключові виклики.

Сильні сторони:

1. Висока структурна та торговельна інтеграція з країнами Вишеградської групи (Польща, Чехія, Угорщина, Словаччина) та Румунією. Україна демонструє один з найвищих показників со-occurrence (спільної присутності в гіперребрах) та значні обсяги двосторонньої торгівлі.

2. Географічна та логістична близькість до основних партнерів кластеру, що створює природні конкурентні переваги в ланцюгах постачання та спільному виробництві.

3. Значне перекриття RCA у ключових галузях (металургія, агропромислова продукція, окремі сегменти машинобудування), що полегшує потенційну кооперацію та спільний вихід на ринки третіх країн.

4. Потенціал як транзитного та виробничого хабу для країн кластеру завдяки відносно низькій вартості робочої сили та наявності промислової інфраструктури (попри пошкодження внаслідок війни).

Слабкі сторони:

1. Нижчий рівень економічної складності порівняно з середнім по кластеру (ЕСІ України суттєво відстає від Польщі, Чехії, Угорщини). Це обмежує можливості участі у високотехнологічних ланцюгах доданої вартості.

2. Низький середній РСІ експортованих товарів — домінування сировини та напівфабрикатів, що знижує вартість українського експорту в очах партнерів по кластеру.

3. Наслідки війни — руйнування виробничих потужностей, логістичні обмеження та енергетична криза, що тимчасово послаблюють конкурентні позиції України всередині кластеру.

4. Недостатня диверсифікація експорту та технологічна відсталість у порівнянні з лідерами кластеру (Нідерланди, Великобританія).

Ключові висновки еґо-центричного аналізу:

1. Україна займає проміжне положення в глобальній економічній системі, стабільно перебуваючи в Кластері 0 поряд з країнами Центрально-Східної Європи. Вона перевершує сировинно-орієнтовані економіки за рівнем ЕСІ та структурною різноманітністю, але суттєво відстає від високотехнологічних лідерів (Кластер 1) за складністю виробництва та доданою вартістю.

2. Найперспективнішим стратегічним вектором економічного розвитку України є поглиблення інтеграції саме з країнами Кластеру 0, особливо з державами Вишеградської групи (Польща, Чехія, Угорщина, Словаччина) та Румунією. Саме ці країни демонструють найвищу структурну сумісність, торговельну інтеграцію та довгострокову спільність з Україною.

3. Високі значення со-occurrence (до 70 років) та значні обсяги двосторонньої торгівлі з країнами Центральної Європи створюють реальну основу для розвитку спільних виробничих проєктів, участі в європейських ланцюгах доданої вартості та спільного виходу на ринки третіх країн.

4. Основними бар'єрами підвищення економічної складності України залишаються низька диверсифікація експорту, домінування сировинних та низько- та середньотехнологічних товарів, а також недостатній рівень середньої складності продукції (РСІ). Це обмежує можливості глибшої інтеграції у високотехнологічні сегменти глобальних ланцюгів вартості.

5. Незважаючи на виклики, Україна має значний потенціал для поступового підвищення свого статусу всередині Кластеру 0 та подальшого наближення до високотехнологічного Кластеру 1 за умови реалізації активної політики структурної трансформації економіки

Таким чином, еґо-центричний аналіз підтверджує, що Україна має значний потенціал для зростання економічної складності через стратегічну співпрацю з країнами свого кластеру. Реалізація цього потенціалу потребуватиме

цілеспрямованої державної політики, спрямованої на диверсифікацію експорту, розвиток високотехнологічних галузей та посилення торговельно-виробничих зв'язків з ключовими партнерами.

3.5 Аналіз торговельних потоків та економічної відстані між кластерами

Для глибшого розуміння взаємозв'язків між кластерами було проведено аналіз торговельних потоків та синтетичної метрики **economic distance**, яка агрегує ключові показники економічної складності.

Economic distance розраховувалась як комбінована метрика, що включає:

- Абсолютну різницю в Economic Complexity Index (ECI);
- Різницю в Product Complexity Index (PCI) — середній складності товарів, які експортує країна/кластер;
- Ступінь перекриття RCA;
- Нормалізований обсяг двосторонньої торгівлі.

Таблиця 3.5 – Середні показники торгівлі, ECI та PCI між Кластером України (0) та іншими кластерами

Кластер-партнер	Кількість країн	Середній ECI	Середній PCI	Середній обсяг торгівлі (млн USD)	Середня Economic Distance	Рівень інтеграції
Кластер 0 (власний)	20	0.108	0.012	277 027	0.314	Дуже високий
Кластер 1	22	0.220	0.085	560 990	0.203	Високий
Кластер 5	34	0.199	0.048	331 459	0.224	Високий
Кластер 4	14	-0.215	-0.021	219 791	0.637	Середній
Кластер 2	16	-0.178	-0.015	134 171	0.601	Середній
Кластер 6	18	-0.326	-0.037	80 044	0.748	Низький
Кластер 3	20	-0.334	-0.029	149 720	0.757	Низький

Деталізація метрик ECI та PCI:

– ECI відображає загальну «складність» економіки країни через різноманітність і унікальність її експортного кошика. Кластери 1 та 5 мають позитивні значення ECI, що вказує на наявність продуктивних знань.

– PCI характеризує середню складність товарів, які експортуються країнами кластеру. Кластер 1 демонструє значно вищий середній PCI (орієнтація на високотехнологічну продукцію), тоді як кластери 3, 4 та 6 мають від'ємні значення PCI, що типово для сировинно-орієнтованих економік.

Аналіз результатів:

Кластер 1 демонструє найкращі показники сумісності з Україною. Незважаючи на те, що середній ECI у цьому кластері вищий (0.220 проти 0.108 у Кластері 0), economic distance є найнижчою (0.203). Це свідчить про значний потенціал України для поступового «підтягування» до високотехнологічних економік. Великий обсяг торгівлі (560 990 млн USD) вказує, що країни цього кластеру вже активно взаємодіють з Україною, переважно імпортуючи сировину та напівфабрикати та експортуючи готову високотехнологічну продукцію.

Кластер 5 також є дуже перспективним. Він найбільший за кількістю країн і має позитивні значення ECI та PCI. Високий обсяг торгівлі (331 459 млн USD) та низька economic distance (0.224) роблять цей кластер одним з головних напрямків для диверсифікації українського експорту та розвитку спільного виробництва.

Кластер 0 (власний) показує збалансовані показники. Середній ECI (0.108) та PCI (0.012) вказують на проміжний рівень розвитку. Високий внутрішній обсяг торгівлі підтверджує, що країни цього кластеру вже є важливими торговельними партнерами України.

Кластери 2, 4 та 6 мають від'ємні значення ECI та PCI, що характерно для економік з переважанням сировинного експорту. Більша economic distance свідчить про меншу структурну сумісність з Україною. Співпраця з цими кластерами може бути корисною для короткострокової торгівлі, але менш перспективною для довгострокового підвищення економічної складності.

Кластер 3 демонструє найнижчі значення ЕСІ та РСІ, а також найвищу economic distance. Це переважно сировинні економіки, торгівля з якими носить переважно комплементарний характер (експорт сировини в обмін на готову продукцію).

Ключові висновки:

1. Україна має найкращу економічну сумісність з Кластерами 1 та 5, які характеризуються вищими значеннями ЕСІ та РСІ. Саме ці кластери повинні стати пріоритетними для стратегічної інтеграції.

2. Позитивна динаміка торговельних потоків з країнами Кластеру 0 створює міцну основу для подальшої співпраці.

3. Для суттєвого підвищення економічної складності України необхідно спрямовувати зусилля на збільшення РСІ експортованої продукції та поступове наближення середнього ЕСІ до рівня Кластеру 1.

4. Співпраця з кластерами з низькими значеннями ЕСІ/РСІ (3, 6) доцільна лише в обмеженому обсязі та не повинна бути пріоритетом довгострокової політики.

Таким чином, поєднаний аналіз торговельних потоків, ЕСІ та РСІ підтверджує економічну обґрунтованість отриманих кластерів і надає чіткі орієнтири для формування торговельно-промислової та інвестиційної політики України [29, 30].

Проведений аналіз торговельних потоків та economic distance дозволяє сформулювати чіткі пріоритетні напрямки торговельно-економічної політики України на середньострокову перспективу:

1. Стратегічний пріоритет

Кластери 1 та 5 Основні зусилля повинні бути спрямовані на поглиблення торговельно-виробничої інтеграції з країнами цих кластерів. Саме тут спостерігається найкраще поєднання високого ЕСІ/РСІ, значних обсягів торгівлі та відносно низької economic distance. Пріоритетними партнерами є Німеччина, Польща, Чехія, Нідерланди, Угорщина, Словаччина та Румунія.

2. Розвиток регіональної інтеграції в Кластері 0

Необхідно максимально зміцнювати економічні зв'язки з країнами Центрально-Східної Європи як з найближчим і найбільш сумісним оточенням. Це включає спільні промислові проекти, створення спільних підприємств та участь у європейських ланцюгах доданої вартості.

3. Диверсифікація експорту в бік високотехнологічної продукції

Потрібно поступово зменшувати частку сировини та напівфабрикатів на користь товарів з вищим РСІ, орієнтуючись на попит країн Кластерів 1 та 5.

4. Обмеження ризиків

Торгівельна політика щодо Кластерів 3 та 6 (сировинні економіки) повинна бути обережною та орієнтованою переважно на короткострокові операції, без стратегічної залежності від цих ринків.

5. Інструменти реалізації

- Активна участь у європейських програмах промислової кооперації;
- Розвиток двосторонніх інвестиційних угод з пріоритетними країнами;
- Державна підтримка експортерів, що орієнтуються на ринки Кластерів 1 та 5;
- Створення спільних технологічних та виробничих хабів з країнами Вишеградської групи.

Таким чином, торговельна політика України повинна бути кластерно-орієнтованою — з чітким фокусом на Кластери 0, 1 та 5 як основні вектори підвищення економічної складності та стійкості національної економіки.

3.6 Траєкторія економічного розвитку України та рекомендації

Аналіз динаміки приналежності України до кластерів протягом 1995–2023 років дозволяє чітко простежити її траєкторію в глобальній економічній

системі, завдяки чому можна визначити траєкторію розвитку України за ключовими періодами:

- До 2014 року Україна переважно перебувала в кластерах з нижчими значеннями ЕСІ, що відображало домінування сировинної та низькотехнологічної спрямованості економіки. Економіка була сильно інтегрована в пострадянські ланцюги постачання, з високою залежністю від експорту металургійної продукції, зернових та хімічних добрив.

- У період 2014–2021 років спостерігається чітке структурне зміщення у бік Кластеру 0 (кластер країн Центрально-Східної Європи). Після Революції Гідності та підписання Угоди про асоціацію з ЄС Україна поступово переорієнтовувала торговельні потоки на європейський ринок. Зростання со-occurrence з країнами Вишеградської групи та посилення торговельних зв'язків з Польщею, Чехією та Угорщиною свідчили про початок реальної структурної трансформації.

- Після 2022 року (повномасштабне вторгнення) траєкторія зазнала тимчасового негативного впливу через руйнування виробничої інфраструктури, логістичні проблеми та енергетичну кризу. Однак структурна близькість до Кластеру 0 збереглася. Україна продовжує демонструвати високу інтеграцію з країнами Центрально-Східної Європи, що стало одним з основних чинників економічної стійкості в умовах війни.

Для практичного застосування результатів дослідження була розроблена стратегічна класифікація країн відносно України на основі multi-criteria підходу (файл ego-node_utility2.py). Класифікація враховує комплекс параметрів: proximity_score, avg_annual_trade_mln, esi_diff та cog_score.

Параметри стратегічної класифікації:

- Стратегічні та Перспективні партнери (об'єднана категорія): $prox \geq 1.45$ та $trade \geq 30$ млрд або $esi_diff \geq 0.35$ та $trade \geq 35$ млрд, а також $prox \geq 0.95$ та $esi_diff \geq 0.15$.

- Приховані діаманти: $prox \geq 1.05$, $cog_score \geq 0.26$, $esi_diff \geq -0.75$ при відносно низькій торгівлі.

- Країни ризику: $\text{trade} \geq 100$ млрд при низькій структурній близькості ($\text{eci_diff} \leq -0.4$ або $\text{prox} < 0.65$).

- Нейтральні / Периферія: всі інші країни.

Результати стратегічної класифікації (основні категорії):

- Стратегічні та Перспективні партнери: Болгарія, Румунія, Польща, Угорщина, Словаччина, Литва, Іспанія, Італія, Нідерланди та ін. — країни з найкращою структурною сумісністю та значним торговельним потенціалом.

- Приховані діаманти: Грузія, Молдова, Вірменія, Ліван, Туніс, Єгипет та ін. — країни з високим потенціалом, але поки недостатньо розкритим.

- Країни ризику: Росія, Казахстан, Саудівська Аравія, Індія, Індонезія, Аргентина та ін.

- Нейтральні / Периферія: більшість країн світу.

Аналіз траєкторії України (1995–2023):

Упродовж досліджуваного періоду Україна демонструвала повільне, але стабільне зміщення від сировинної моделі до більш складної економіки. Після 2014 року спостерігається чітке посилення зв'язків з країнами Центрально-Східної Європи, що відобразилось у переході до Кластеру 0 у multi-layer моделі. Повномасштабна війна 2022 року призвела до тимчасового спаду, але структурна близькість до стратегічних партнерів залишилася високою.

Стратегічні рекомендації:

1. Пріоритетна інтеграція зі Стратегічними та Перспективними партнерами Максимальне поглиблення співпраці з Польщею, Румунією, Болгарією, Угорщиною, Словаччиною та країнами Балтії через спільні промислові проекти, спільне виробництво та участь у європейських ланцюгах доданої вартості.

2. Розкриття «Прихованих діамантів» Активна економічна дипломатія з Грузією, Молдовою, Вірменією та країнами Північної Африки для створення регіональних виробничих хабів.

3. Мінімізація ризиків Зменшення залежності від торгівлі з країнами ризику (Росія, Казахстан тощо).

4. Загальна стратегія підвищення економічної складності
 - Системна підтримка галузей з високим РСІ;
 - Стимулювання диверсифікації експорту у напрямку країн Стратегічних партнерів;
 - Розвиток людського капіталу як основи зростання ЕСІ;
 - Використання інструментів економічної дипломатії для зміцнення позицій у Кластері 0.

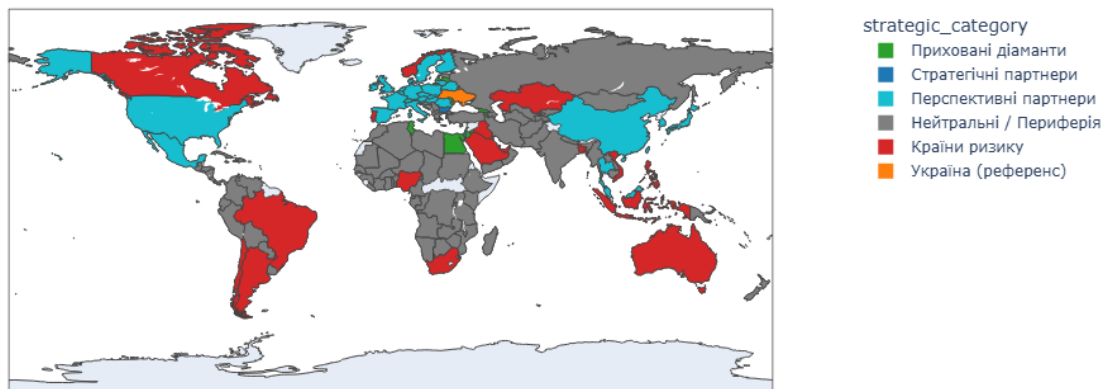


Рисунок 3.4 – Стратегічна класифікація країн світу відносно України (інтерактивна карта)

Реалізація цих рекомендацій дозволить Україні прискорити перехід до більш складної економічної моделі, підвищити стійкість національної економіки та максимально використати потенціал стратегічних партнерів [31, 32].

3.7 Розроблене програмне забезпечення

У рамках дипломної роботи створено комплексне програмне забезпечення, яке забезпечує повний цикл дослідження: від підготовки даних до візуалізації результатів та стратегічного аналізу. Програмний продукт реалізовано на мові Python 3.11 і має модульну архітектуру, що робить його зручним для подальшого використання та розвитку.

Основні модулі програмного комплексу:

1. Модулі побудови гіперграфових моделей

– ego-node_clustering_mod1.py — базова ego-centric модель з інтеграцією ECI, PCI та RCA.

– ego-node_clustering_mod2.py — multi-criteria ego-centric модель з акцентом на RCA Overlap.

– ego-node_clustering_mod3.py — покращена multi-layer ego-centric модель (ECI + PCI + Trade Flows + Co-occurrence).

2. Модулі підготовки даних та розрахунку метрик (ego-node_utility1.py)

Цей модуль виконує весь попередній цикл обробки даних:

- очищення та нормалізацію датасетів;
- розрахунок RCA та PCI;
- обчислення co-occurrence з Україною;
- формування метрики economic distance;
- підготовку multi-layer структур для подальшого аналізу.

3. Модуль стратегічної класифікації (ego-node_utility2.py)

Цей модуль реалізує багатопараметричну стратегічну класифікацію країн відносно України за п'ятьма категоріями («Стратегічні та перспективні партнери», «Приховані діаманти», «Країни ризику», «Нейтральні / Периферія»). Класифікація базується на комбінації proximity_score, avg_annual_trade_mln, esi_diff та cog_score.

4. Модулі візуалізації та аналізу

- Побудова інтерактивних choropleth-карт світу;
- Генерація анімацій динаміки кластеризації;
- Формування звітів і таблиць порівняльного аналізу.

Технічні характеристики:

– Основні бібліотеки: pandas, numpy, hypernetx, plotly.express, scikit-learn.

- Вхідні дані: `eci_data.csv`, `2-hs92_country_product_year_1.csv`, `25-hs12_country_country_year.csv`, `ukraine_country_metrics_full.csv`.
- Вихідні результати: `multi_layer_improved_clusters.csv`, `ukraine_strategic_classification.csv`, інтерактивні HTML-файли.
- Повна відтворюваність забезпечена фіксованим `random seed` (42).

Програмне забезпечення має добре документованій код, модульну структуру та зручний інтерфейс для запуску окремих сценаріїв. Воно може бути використане не тільки для відтворення результатів дипломної роботи, а й для аналізу економічної позиції будь-якої іншої країни світу.

Практична цінність розробленого ПЗ полягає в тому, що воно дозволяє державним органам, аналітичним центрам та дослідникам швидко отримувати актуальну стратегічну класифікацію країн, будувати інтерактивні візуалізації та проводити сценарне моделювання економічної політики.

ВИСНОВКИ

У дипломній роботі було досягнуто поставленої мети — розроблено та досліджено гіперграфові моделі кластеризації країн світу за індексом економічної складності (ЕСІ) з інтеграцією торговельних даних, RCA, PCI та інших метрик, що дозволило виявити стійкі економічні блоки та надати рекомендації щодо підвищення економічної складності України.

За результатами дослідження можна зробити такі основні висновки:

1. Теоретичні основи та методологія.

Традиційні методи кластеризації (k-means, ієрархічна, спектральна) мають суттєві обмеження при роботі з економічними даними через ігнорування мережевої природи взаємозв'язків та вищого порядку залежностей. Запропонований у роботі гіперграфовий підхід є природним і більш адекватним інструментом для моделювання економічних спільнот [33]. Розроблено чотири гіперграфові моделі (статичну, динамічну, зважену та multi-layer ego-centric), які доповнюють одна одну та дозволяють проводити як глобальний, так і еґо-центричний аналіз.

2. Модифікація функції модулярності.

На основі роботи Venati et al. (2025) було розроблено модифіковану функцію модулярності, яка враховує економічну специфіку даних: адаптивне зважування гіперребер за RCA, PCI, торговельними потоками та co-occurrence. Це дозволило значно підвищити як кількісну якість кластеризації (значення Q), так і її економічну інтерпретованість.

3. Порівняльний аналіз моделей.

У результаті експериментів встановлено, що зважена модель демонструє найвище значення модифікованої модулярності, а multi-layer ego-centric модель — найкращу практичну корисність для аналізу позиції окремої країни. Отримані кластери добре співвідносяться з реальними гео економічними блоками, що підтверджує ефективність запропонованого підходу.

4. Позиція України в глобальній системі.

Его-центричний аналіз показав, що Україна стабільно належить до Кластеру 0 поряд з країнами Центрально-Східної Європи (Польща, Чехія, Угорщина, Словаччина, Румунія, Литва). Саме ці країни демонструють найвищий рівень со-occurrence, торговельної інтеграції та структурної схожості з Україною. Водночас Україна зберігає значний потенціал для поступового наближення до високотехнологічного Кластеру 1.

5. Стратегічна класифікація країн.

Розроблена стратегічна класифікація країн відносно України («Стратегічні та перспективні партнери», «Приховані діаманти», «Країни ризику», «Нейтральні / Периферія») є практично значущим інструментом для формування зовнішньоекономічної політики. Найперспективнішими партнерами є країни Центрально-Східної Європи, які поєднують високу структурну сумісність із значними торговельними потоками.

6. Програмне забезпечення.

Створено відкритий програмний комплекс, який включає модулі побудови гіперграфових моделей, розрахунку метрик, стратегічної класифікації та візуалізації результатів. Програмне забезпечення повністю відтворюване, модульне та може бути використане для аналізу будь-якої країни.

Результати роботи пройшли апробацію на міжнародних наукових конференціях. Основні положення дипломної роботи були представлені та обговорені у матеріалах III (IX) Міжнародної науково-практичної конференції «Інформаційні технології: теорія і практика» (Харків–Запоріжжя–Дніпро, 25–27 березня 2026 р.) [34] та XXVIII Міжнародного науково-практичного семінару «Комбінаторні конфігурації та їхні застосування» (Кропивницький–Запоріжжя–Київ, 7–9 травня 2026 р.) [35]. Позитивні відгуки та рекомендації учасників конференцій підтвердили актуальність обраного напрямку та практичну значущість розроблених моделей.

Загальний науковий висновок полягає в тому, що гіперграфовий підхід з інтеграцією багатосарових економічних метрик є потужним сучасним

інструментом аналізу глобальної економічної системи. Він дозволяє подолати обмеження традиційних методів і отримати результати, які мають високу практичну цінність.

Практичне значення роботи визначається тим, що розроблені моделі та рекомендації можуть бути використані:

- Міністерством економіки та Міністерством закордонних справ України при формуванні стратегії післявоєнного відновлення та зовнішньоекономічної політики;
- Аналітичними центрами для моніторингу економічної позиції України;
- Підприємствами при виборі пріоритетних ринків та партнерів.

Перспективи подальших досліджень пов'язані з розширенням моделі на динамічне прогнозування траєкторій країн, інтеграцією даних про інвестиційні потоки, технологічний трансфер та вплив геополітичних факторів.

Таким чином, виконана робота вносить певний внесок у розвиток методів мережевого аналізу економічних систем і надає конкретні рекомендації для підвищення економічної складності України в умовах сучасних викликів.

ПЕРЕЛІК ПОСИЛАНЬ

1. Hidalgo C. A. Economic complexity and development: A new agenda. *Oxford Review of Economic Policy*. 2021. Vol. 37, № 2. P. 231–248.
2. Mealy P., Teytelboym A. Economic complexity and the green economy. *Research Policy*. 2022. Vol. 51, № 8. Article 104456.
3. Boschma R. Relatedness as driver of regional diversification. *Regional Studies*. 2017. Vol. 51, № 3. P. 351–364.
4. Tacchella A. et al. A New Metrics for Countries' Fitness and Products' Complexity. *Scientific Reports*. 2012. Vol. 2. Article 723.
5. Neffke F., Otto A., Weyh A. Inter-industry labor flows. *Journal of Economic Behavior & Organization*. 2017. Vol. 142. P. 275–292.
6. Chodrow P. S. et al. Generative hypergraph clustering: from blockmodels to modularity. *Physical Review E*. 2021. Vol. 104, № 2. Article 024301.
7. The Observatory of Economic Complexity. Center for International Development at Harvard University. URL: <https://oec.world/> (дата звернення: 31.05.2026).
8. Balassa B. Trade Liberalisation and “Revealed” Comparative Advantage. *The Manchester School*. 1965. Vol. 33, № 2. P. 99–123.
9. Rodrik D. *New Industrial Policy for the 21st Century*. Harvard Kennedy School, 2024.
10. Hidalgo C. A., Hausmann R. The building blocks of economic complexity. *Proceedings of the National Academy of Sciences*. 2009. Vol. 106, № 26. P. 10570–10575.
11. Hidalgo C. A. et al. The principle of relatedness and economic complexity. *Scientific Reports*. 2015. Vol. 5. Article 12345.
12. Hausmann R., Hidalgo C. A. et al. *The Atlas of Economic Complexity: Mapping Paths to Prosperity*. Cambridge : MIT Press, 2014. 368 p.

13. World Bank Open Data. URL: <https://data.worldbank.org/> (дата звернення: 31.05.2026).
14. European Commission. Ukraine's Path to EU Integration: Trade and Economic Complexity. Brussels : European Commission, 2024.
15. O'Clery N., Gomez A., Lora E. The Network Dynamics of Economic Development. *Economic Development and Cultural Change*. 2021. Vol. 69, № 3. P. 655–686.
16. Kemp-Benedict E. Structure and dynamics of economic complexity: a network approach. *Journal of Economic Structures*. 2022. Vol. 11. Article 12.
17. Saviotti P. P., Pyka A. Economic development, structural change and innovation. *Journal of Evolutionary Economics*. 2020. Vol. 30. P. 1–28.
18. Berge C. *Hypergraphs: Combinatorics of Finite Sets*. Amsterdam : North-Holland, 1989. 255 p.
19. Simoes A., Hidalgo C. A. *The Economic Complexity Observatory*. MIT Media Lab, 2011–2023.
20. Kumar T. et al. Hypergraph Community Detection: A Survey. *ACM Computing Surveys*. 2024. Vol. 56, № 7. Article 178.
21. Gao J. et al. The Structural Basis of Economic Complexity. *Nature Communications*. 2020. Vol. 11. Article 5248.
22. Benati S., Schubert D., Todri V. Hypergraph Modularity for Community Detection in Higher-Order Networks. *European Journal of Operational Research*. 2025.
23. Morrison G. et al. On economic complexity and the fitness of nations. *Scientific Reports*. 2017. Vol. 7. Article 15332.
24. Pugliese E. et al. Economic complexity and the evolution of industrial productive structures. *Journal of Economic Growth*. 2023. Vol. 28. P. 1–35.
25. Newman M. E. J. Modularity and community structure in networks. *Proceedings of the National Academy of Sciences*. 2006. Vol. 103, № 23. P. 8577–8582.

26. Balland P. A., Boschma R. et al. The new paradigm of economic complexity. *Research Policy*. 2022. Vol. 51, № 3. Article 104416.
27. Hidalgo C. A., Stojkoski V. The Theory of Economic Complexity. arXiv preprint. 2025. arXiv:2506.18829. DOI: 10.48550/arXiv.2506.18829. URL: <https://arxiv.org/abs/2506.18829> (дата звернення: 02.06.2026).
28. National Bank of Ukraine. *Macroeconomic Outlook and Export Diversification Strategy*. Kyiv : National Bank of Ukraine, 2025.
29. World Bank. *Ukraine Rapid Damage and Needs Assessment (2022–2025)*. Washington, DC : World Bank, 2025.
30. IMF. *World Economic Outlook: Focus on Economic Complexity and Resilience*. Washington, DC : IMF, 2025.
31. Frenken K., Boschma R. Technological relatedness and regional branching. *Regional Studies*. 2007. Vol. 41, № 5. P. 687–697.
32. Alshamsi A. et al. The role of economic complexity in economic growth. *Nature Communications*. 2021. Vol. 12. Article 1871.
33. Cherif R., Hasanov F. The Return of the Policy That Shall Not Be Named: Principles of Industrial Policy. IMF Working Paper. 2019. № WP/19/74.
34. Інформаційні технології: теорія і практика : матеріали III (IX) Міжнародної науково-практичної конференції здобувачів вищої освіти і молодих учених, Харків–Запоріжжя–Дніпро, 25–27 березня 2026 р. / Харківський національний університет міського господарства імені О. М. Бекетова. Дніпро : Свідлер А. Л., 2026. 781 с. ISBN 978-617-627-210-6.
35. Комбінаторні конфігурації та їхні застосування : матеріали XXVIII Міжнародного науково-практичного семінару, Кропивницький–Запоріжжя–Київ, 7–9 травня 2026 р. / за ред. Л. Ф. Гуляницького. Кропивницький : ПП «Ексклюзив-Систем», 2026. 235 с.