

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
КРИВОРІЗЬКИЙ ДЕРЖАВНИЙ ПЕДАГОГІЧНИЙ УНІВЕРСИТЕТ
Кафедра інформатики та прикладної математики

«Допущено до захисту»

В.о. завідувача кафедри

_____ Моїсеєнко Н.В.

Реєстраційний № _____

«___» _____ 2023 р.

«___» _____ 2023 р.

АНАЛІЗ ЕВОЛЮЦІЇ СТРУКТУРИ СОЦІАЛЬНОЇ МЕРЕЖІ У GERPI

Кваліфікаційна робота
студента групи І-19
ступінь вищої освіти «магістр»
спеціальності
014.09 Середня освіта (Інформатика)
Васильченка Антона Вікторовича

Керівник: кандидат фізико-
математичних наук, доцент
Тарасова Олена Юріївна

Оцінка:

Національна шкала _____

Шкала ECTS ___ Кількість балів ___

Голова ЕК _____

Члени ЕК _____

ЗАПЕВНЕННЯ

Я, Васильченко Антон Вікторович, розумію і підтримую політику Криворізького державного педагогічного університету з академічної доброчесності. Запевняю, що ця кваліфікаційна робота виконана самостійно, не містить академічного плагіату, фабрикації, фальсифікації. Я не надавав і не одержував недозволену допомогу під час підготовки цієї роботи. Використання ідей, результатів і текстів інших авторів мають покликання на відповідне джерело. Із чинним Положенням про запобігання та виявлення академічного плагіату у роботах здобувачів вищої освіти Криворізького державного педагогічного університету ознайомлений. Чітко усвідомлюю, що у разі виявлення у кваліфікаційній роботі порушення академічної доброчесності робота не допускається до захисту або оцінюється незадовільно.



ЗМІСТ

ВСТУП.....	4
РОЗДІЛ 1. АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ.....	6
1.1. Теоретичні основи аналізу даних соціальних мереж	6
1.2. Моделі соціальних графів та опис основних метрик	12
1.3. Моделі динамічного аналізу мереж	16
1.4. Порівняльна характеристика програмних засобів для аналізу соціальних мереж	18
Висновки до розділу 1	23
РОЗДІЛ 2 МОДЕЛЮВАННЯ ТА ВІЗУАЛІЗАЦІЯ СОЦІАЛЬНИХ МЕРЕЖ	24
2.1. Вибір програмних засобів моделювання для аналізу та візуалізації.....	24
2.2. Практична реалізація моделей соціальних мереж.....	26
Висновки до розділу 2	34
ВИСНОВКИ.....	35
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ	36

ВСТУП

Актуальність теми. В епоху інформаційного глобалізму спостерігається стрімкий розвиток соціальних мереж, адже з кожним днем темпи їх поширення і популярності зростають, Україна не лишається осторонь цих процесів. Сьогодні соціальні мережі – це не тільки інструмент спілкування, обміну думками й отримання інформації, а ще й важливий елемент у технологічному, політичному, соціальному та інформаційному протистоянні. Дані соціальних мереж – це, передусім, невичерпне джерело дослідницьких можливостей. Графи є ефективним інструментом побудови безлічі інформаційних моделей для задач різних предметних галузей. Це невід’ємний компонент аналізу соціальних мереж, який також дозволяє і ефективно їх візуалізувати. Візуальне представлення медіа даних є важливим та нетривіальним завданням, зважаючи на їх різноманітність та великий обсяг. Візуалізація дозволяє виявити закономірності, дати загальне уявлення про структуру мережі.

Аналіз соціальних мереж, як правило, реалізовується у рамках ресурсозберігаючих, нормативних та динамічних напрямків досліджень. Кожен з них розв’язує досить велике коло завдань та використовує різноманітний комплекс методів. Завдяки доступному інструментарію можна ефективно проводити структурний аналіз та аналіз поведінки зв’язків, статистичну оцінку соціальних мереж з урахуванням їх можливого масштабування, контент-аналіз соціальної мережі (аналіз тексту, заголовків, медіа-даних), визначення спільнот у рамках соціальної мережі.

Метою роботи є дослідження аналізу еволюції структури соціальної мережі за допомогою візуалізації у Gephi.

Для досягнення мети сформульовані **задачі**:

1. теоретичний огляд наукової та методичної літератури;
2. дослідження основних алгоритмів та методів аналізу соціальних мереж;
3. побудова, візуалізація та аналіз соціальної мережі за допомогою Python та Gephi.

Об'єкт дослідження: динамічні характеристики цільових груп соціальної мережі.

Предмет дослідження методи та моделі аналізу еволюційних характеристик соціальних графів.

Методи дослідження – методи математичної логіки, теорії графів, математичної статистики, системного аналізу, методи аналізу соціальних мереж, апарату математичного аналізу, лінійної алгебри, методи математичного моделювання, теорія алгоритмів.

Практичне значення роботи у тому, що запропонований алгоритм дозволить простежити динаміку зміни показників параметрів метрик, порівнюючи їх зі структурними змінами у соціальному графі. Запропонований прототип програмного продукту може використовуватися для аналізу та вивчення соціальних мереж.

Структура та обсяг: робота містить вступ, два розділи, висновки. Загальний обсяг роботи – 38 сторінок, з них 35 сторінок основного тексту. Текст містить 11 рисунків. Список використаних джерел на 3 сторінках складається з 28 позицій. Оформлення кваліфікаційної було здійснено згідно Положення про кваліфікаційні роботи в КДПУ.

РОЗДІЛ 1. АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ

1.1. Теоретичні основи аналізу даних соціальних мереж

Сьогодні мільйони людей по всьому світу не уявляють життя без соціальних мереж і відвідують одну або декілька з них щодня. Для пересічних користувачів соціальні мережі дають передусім величезні можливості для спілкування, пошуку інформації, моніторингу актуальних новин, підтримки зв'язку з друзями та рідними у будь-якому навіть віддаленому куточку світу, а також для створення та просування власного бізнесу, продажу товарів чи інфопродуктів.

Аналіз соціальних мереж зародився у 1934 році, коли Джейкоб Леві Морено створив соціограми – абстракції соціальних взаємодій для візуального представлення комунікаційної мережі [1]. Зокрема, соціограма – це граф, у якому кожен вузол представляє людину, а ребра – взаємодію між учасниками. Морено використовував соціограми для вивчення поведінки невеликих груп людей, оскільки в епоху, в яку він працював, було важко отримати детальну інформацію про велику кількість особистих взаємодій. Усе змінилося з появою онлайн-соціальних мереж, таких як Facebook, Twitter, Instagram. Сьогодні будь-хто може безкоштовно завантажити об'ємні дані, що відкриває можливості для їх всебічного аналізу.

Аналіз соціальних мереж (social network analysis, SNA) є міждисциплінарним інструментарієм тому, що знаходиться у фокусі предметних областей різних наук: інформатики, соціології, математики (теорія графів), статистики, теорії складних систем, системного аналізу, психології. Так, у [2] проаналізована історія виникнення та методологія мережного аналізу (рис. 1.1), розглядаються основні концепції, пов'язані з аналізом соціальних мереж, їх історичний розвиток і сучасні тенденції у дослідженні цієї області. Автор [2] наголошує, що аналіз соціальних мереж також тісно пов'язаний, переважно на емпіричному та прикладному рівні, з мікросоціологією та соціометрією.

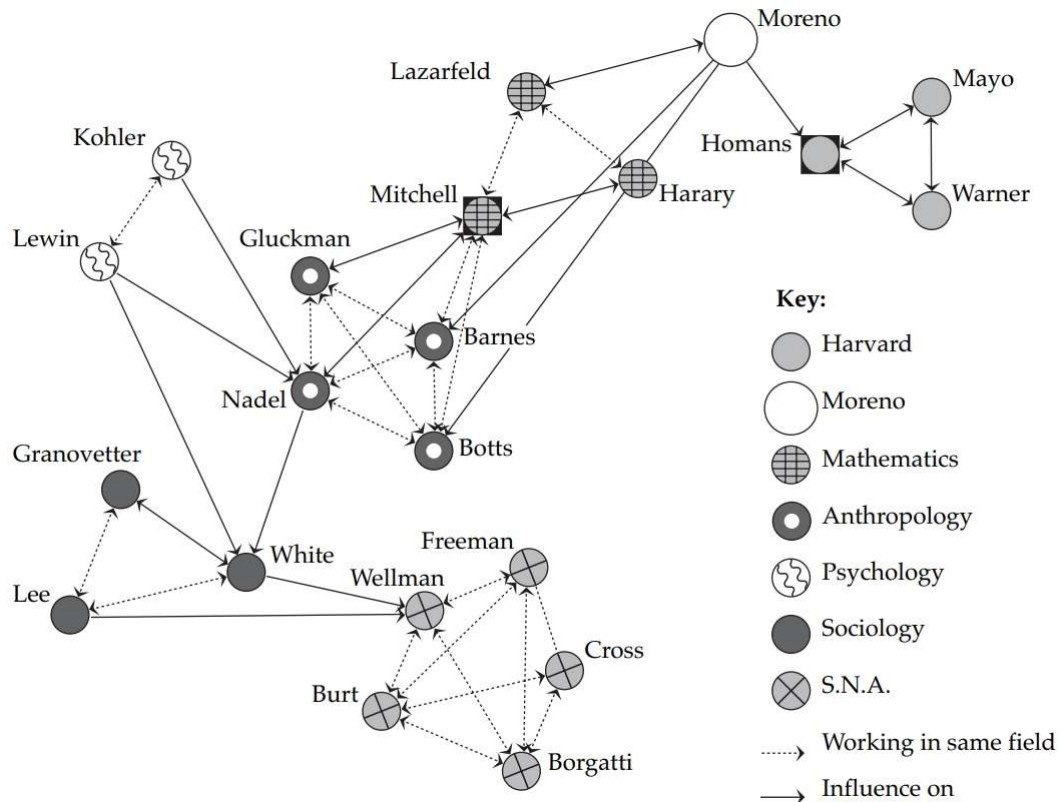


Рис. 1.1 Історія та впливові особи в аналізі соціальних мереж (SNA) [3]

П. Ердеш, А. Рен'ї та А. Барабаші [4] є відомими вченими, які вивчали та створили класичні математичні моделі соціальних мереж. Сучасні дослідження соціальних мереж також активно проводяться науковцями з групи «Social Network Analysis» [5], які зробили значний внесок у розвиток цієї галузі. Підтримка розглядуваної проблематики постійно присутня на освітній платформі технологічної компанії, що працює в галузі освіти – Coursera [6]. Про що свідчить результат опрацювання запиту «social network analysis» – 345 пов'язаних матеріалів освітніх курсів. Наукова корпорація INSNA [6] займається публікацією спеціалізованих наукових журналів «Journal of Social Structure» та «Social Networks Journal», а також формує список запитів SOCNET.

За допомогою Google Trends [7] проаналізуємо динаміку популярності запиту «social network analysis» у світі серед користувачів Google за останні 5 років. Бачимо стійкий тренд, який тримається орієнтовно на відмітці близько 80 відносних одиниць. Зазначимо, що 100 відсоткових одиниць відповідає найвищому рівню попиту на цей запит у відносних одиницях.

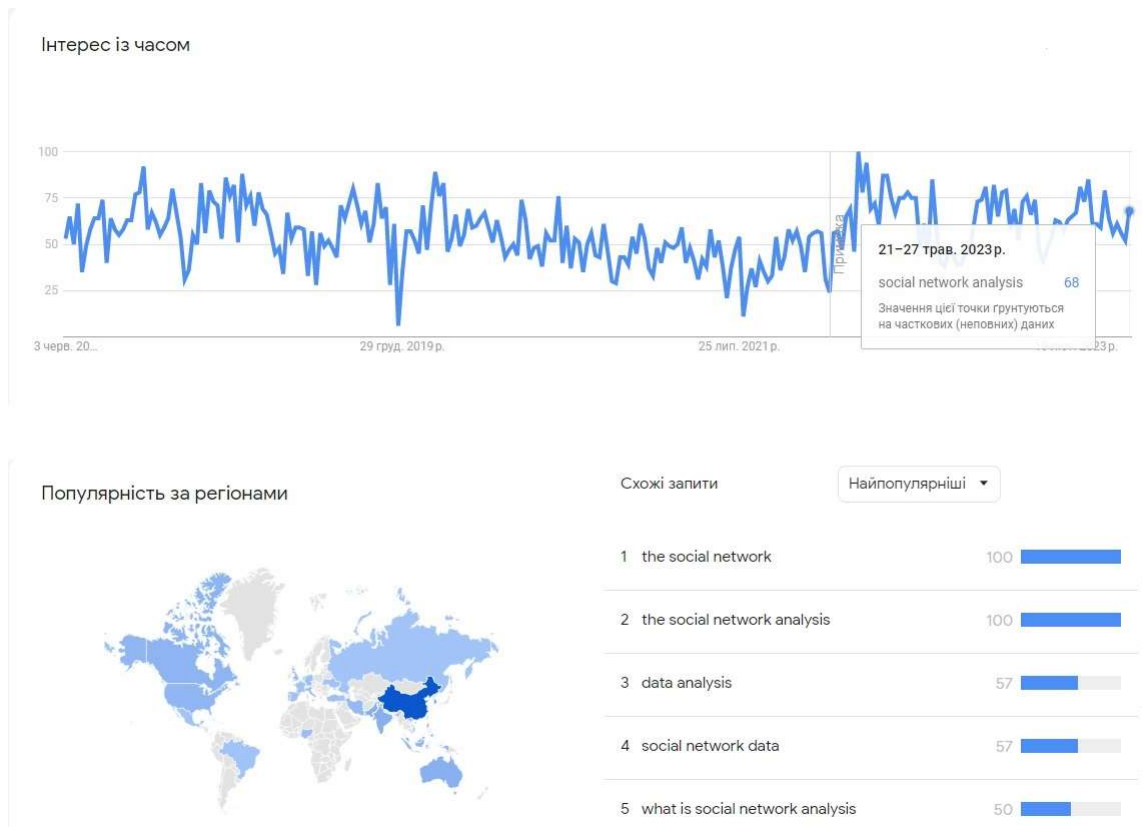


Рис. 1.2 Демонстрація динаміки використання запита «social network analysis» проаналізованого Google Trends

Термін «соціальна мережа» уперше запровадив соціолог манчестерської школи Джеймс Барнс у 1954 році, задовго до появи Інтернету і сучасних мереж. Барнс прийшов до висновку, що розмір соціальної мережі навколо однієї людини становить приблизно 150 осіб [8].

Стосовно Інтернет спільноти – поняття «соціальна мережа» вперше використав Тім О’Рейлі – основоположник концепції Web 2.0 у 2005 році [9].

Classmates.com вважається першою соціальною онлайн мережою, що з’явилася у 1995 році (США). На початку 2000-х років народились MySpace, LinkedIn і Facebook, які і сприяли розвитку, поширенню та популяризації тренду соціальних мереж.

Як відомо, соціальні мережі в основному відрізняються одна від одної складом аудиторії (віковий параметр, гендерна приналежність, сфера інтересів). З початком повномасштабної війни в Україні, за даними GlobalLogic [10] значно зросла кількість користувачів соціальних мереж. Це пояснюється тим, що соціальні мережі використовують перш за все як джерело новин. Так, на основі відкритих даних

з'ясовано, що у липні 2022 року соцмережами користувалися приблизно 76,6% українців, серед них 66% є користувачами Telegram, 61% – YouTube, 58% – Facebook. Більшість українців надають переваги розважальним соцмережам, і, лише 3,6 млн аккаунтів зареєстровані у LinkedIn – платформі для бізнес-комунікацій та пошуку роботи. Важливо зазначити, що опитувані користувачі могли обирати одразу декілька варіантів відповідей [10].

У світовому масштабі, згідно даних Meltwater за 2022 рік [11] переможцями серед усіх соціальних мереж є – Facebook, Twitter, Instagram, TikTok, LinkedIn та YouTube. Фактично, 90% користувачів використовують Facebook як частину своєї соціальної поведінки, що робить цю мережу найпопулярнішою. За ним у порядку спадання популярності використання слідує LinkedIn, Instagram, YouTube і Twitter [11].

У поняття «соціальні мережі» (Social Networks) зараз перш за все вкладають зміст, що це сервіси Internet, які забезпечують формування, впорядкування, відображення відношень між учасниками мережі. Дані сервіси дозволяють створювати профілі для спілкування з іншими. Соціальні мережі характеризуються тим, що мають безліч різного контенту та дуже великі масиви даних про користувачів та їх зв'язки, які доцільно використовувати для аналізу.

Соціальну мережу часто представляють просто як множину вершин – соціальних суб'єктів. У ролі суб'єктів можуть бути користувачі, соціальні групи, зв'язки між ними.

Головними особливостями соціальних мереж зазвичай є наступні:

- широкий спектр можливостей для обміну інформацією між користувачами мережі;
- профілі користувачів містять певний персональний контент;
- реальні друзі у соціальних мережах переважають віртуальних.

Веб-ресурс соціальної мережі забезпечує доступ користувачам до основних функціональних можливостей:

- активне спілкування;

- створення профілю користувача, який може бути публічним або закритим, і включає особисті дані;
- створення та управління користувачем списком осіб, які він вважає своїми друзями і з якими має певні соціальні зв'язки;
- взаємодія користувачів у процесі обміну повідомленнями, перегляд профілів друзів, надсилання файлів, коментарів;
- можливість перегляду взаємозв'язків між користувачами всередині цієї соціальної платформи;
- можливість кооперації, створення груп на основі спільних інтересів для реалізації певних цілей і досягнення спільної мети, наприклад, ведення групового блогу;
- управління вмістом у межах власного профілю;
- синдикація (тобто одночасне поширення інформації на різні телеканали, вебсторінки або вебсайти) контенту;
- підключення різних додатків;
- обмін ресурсами (наприклад, покликаннями на дописи чи на сайти).

Онлайн соціальна мережа є новітнім видом мережі, який в області досліджень практично витіснив традиційні інструменти соціальних наук, такі як опитування, інтерв'ю та анкетування.

Соціальна мережа представляє собою узагальнену структуру, яка складається із множини, що містить серії агентів, у якості яких виступають суб'єкти – особи, спільноти, групи осіб або організації. Ще один компонент соціальної мережі визначається множиною відносин – набір взаємозв'язків між учасниками (агентами), такими як знайомства, комунікація, дружба, та інші форми взаємодії.

Описувати та аналізувати соціальні мережі дуже зручно та ефективно саме за допомогою графа, так як граф дозволяє розв'язувати багато різноманітних задач, виявляти зв'язки між користувачами та аналізувати ступінь їх взаємодії. Доступність особистих даних користувачів дає можливість розширити спектр інформації, отримати додаткові дані для поглибленого аналізу і формування певної статистики.

Графи набувають усе більшої популярності і головною їх перевагою є зручність та ефективність їх використання для вирішення різноманітних задач з аналізу та обробки даних, дослідження зв'язків, пошуку найоптимальніших шляхів та наглядного представлення даних [12].

Граф – це математичне зображення мережі, що складається з деяких точок і ліній і описує взаємні зв'язки між ними (рис. 1.1).

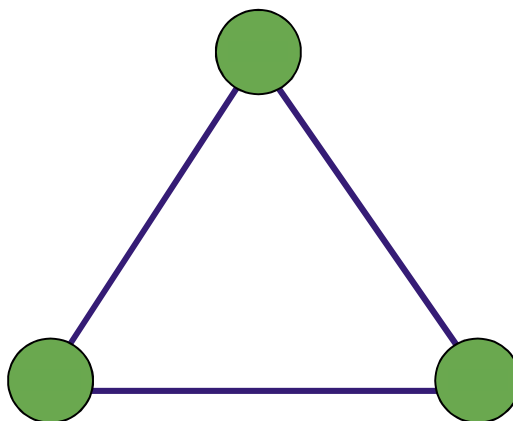


Рис. 1.1. Простий граф з трьома вершинами

Математичний опис графа виглядає так: «граф «G» – це впорядкована пара двох множин: множина «V» складається з набору вершин, які зазвичай називаються вузлами; множина «E» представляє собою з'єднані ребра, які ще називають посиленнями»:

$G = (V, E)$ – це графічне представлення, де V представляє собою вершини або вузли, які можуть бути особами, організаціями або продуктами, а E – це ребра або зв'язки, що відображають зв'язки між двома вузлами, які з'єднуються шляхом взаємодій або шаблонів спілкування. Також використовується термін «край».

Суб'єкти соціальної мережі групуються за принципом схожості зв'язків (наприклад, спільні інтереси, спільні соціальні групи, спільне місце роботи чи проживання). Принципи подібності зв'язків можуть бути будь-якими. На основі даних із соціальних мереж можна побудувати соціальний граф користувача або соціальний граф інтересів. Дані графи схожі за принципами, але мають деякі принципові відмінності. У соціальному графі зв'язки формуються з урахуванням друзів користувача, тоді як у графі інтересів зв'язки формуються за подібністю

захоплень. Зв'язки у такому графі можуть бути як «користувач – користувач» або «користувач – інтерес», так і «цікавість – інтерес».

Одна з найпопулярніших соціальних мереж Facebook є яскравим прикладом ненаправленого графу, де вершинами (вузлами) вважають користувачів, а зв'язками («краєм») називають наявність зв'язку «дружби» між ними. Прикладом направленою графу можна вважати алгоритм рейтингу сторінок Google.

Усі дані у соціальній мережі поділяють на структуровані та неструктуровані. Структуровані дані в соціальні мережі найпростіше описати саме за допомогою графа, тому саме ці дані часто представляють графічною структурою або моделюють у вигляді графа. Аналіз соціальних мереж та використання графічних аналітичних додатків є найпоширенішими методами для вимірювання структурованих даних.. Неструктуровані дані (користувацький контент (UGC)) вимірюються за допомогою контент – аналізу, а також методів, що використовують алгоритми структурування даних.

1.2 Моделі соціальних графів та опис основних метрик

Для опису, візуалізації та відображення загальної структури і взаємозв'язків усіх членів соціальної мережі доцільно використовувати контекстну соціограму – соціальний граф [15]. Кампанія Facebook вперше розробила соціальний граф, що описує їх соціальну мережу.

Соціальний граф визначається певними особливостями, які можна охарактеризувати за допомогою певних метрик, які допомагають вирішувати різноманітні задачі. Дані метрик використовують для аналізу соціальних мереж, так як вони відображають характеристики соціальних об'єктів у числовому форматі.

Для вирішення різноманітних завдань аналізу на соціальному графі зазвичай використовують групи метрик: метрики взаємовідносин, метрики зв'язків та метрики сегментації.

- Гомофілія – це міра, в якій користувач утворює зв'язки з іншими подібними користувачами, залежно від спільних характеристик, таких як вік, стать, соціальний стан, освіта тощо.

- Множинність – це числовий показник "множинних" зв'язків, в яких знаходиться користувач.
- Взаємність – визначається ступенем взаємодії між користувачами і реагуванням на дії один одного.
- Мережева закритість – це ступінь, в якому друзі користувача також є друзями один одного.
- Сусідство – це тенденція користувачів мати багато зв'язків з іншими користувачами, які знаходяться географічно близько один до одного.

Метрики зв'язків відображають характеристики взаємозв'язків як окремих соціальних об'єктів, так і всієї графової структури:

- Міст – це користувач або кластер, який створює спеціальний унікальний зв'язок між іншими користувачами або кластерами, і через нього проходить найкоротший маршрут.
- Центральність – це показник, що визначає ступінь «важливості» або «впливу» певного користувача або кластера користувачів всередині мережі.
- Густина – це відношення кількості прямих зв'язків у мережі до загальної кількості можливих зв'язків.
- Відстань – це мінімальна кількість зв'язків, які потрібні для встановлення стійкого взаємозв'язку між двома користувачами.
- «Структурні діри» або «чорні діри» – це відсутність зв'язків між двома частинами мережі.
- Сила зв'язку – це метрика, що визначається лінійною комбінацією часу, «близькості» та «взаємності». Сильні зв'язки визначаються через «сусідство», «гомофілію», «сусідство» або «транзитивність», тоді як слабкі зв'язки визначаються через «мости».

Метрики сегментації показують аспекти соціального графа, який поділений на різні сегменти з унікальними характеристиками та відмінними рисами.

- Кліка – це група користувачів, де кожен користувач має прямі зв'язки з усіма іншими учасниками цієї групи. Іншими словами, вона складається з вершин, які всі з'єднані ребром між собою.
- Соціальне коло – групова спільнота, де «прямі» зв'язки між користувачами не є обов'язковими.
- Коефіцієнт кластеризації – це ймовірність того, що два різних актори мережі, які мають зв'язок з конкретною особою, також будуть між собою пов'язані. Високий коефіцієнт кластеризації вказує на високу ступінь замкнутості, іншими словами, група може мати характеристики кліки.
- Згуртованість – це міра, яка вказує на те, наскільки користувачі взаємопов'язані одним загальним зв'язком. Структурна згуртованість має особливу властивість – навіть найменша кількість видалених користувачів може призвести до розриву групи.

Завдання, що вирішуються за допомогою соціальних графів досить різноманітні. Серед них: соціальний пошук, генерація персоналізованих рекомендацій, ідентифікація користувачів, виявлення «реальних» зв'язків між користувачами.

Для створення моделі графів збір інформації здійснюється шляхом парсингу веб-сервісів і побудови соціального графа, використовуючи отримані дані. Для визначення результатів, дуже важливо враховувати критерії оцінювання та фактори обходу.

На рис. 1.2 представлена класифікація моделей графів [15, 13]. Зробимо короткий опис даної класифікації:

Функціонально-керовані моделі відтворюють статистичні характеристики графа, такі як динамічні зміни щільності графа і поступовий розподіл. До цієї групи моделей належать дві моделі:

1. Модель Барабаші–Альберта, що генерує безмасштабні мережі і включає важливі характеристики: першою з них є ріст, а другою – привілейованість.

2. Forest Fire або «Палаючий ліс» – динамічна модель, що демонструє самоорганізовану критичність.

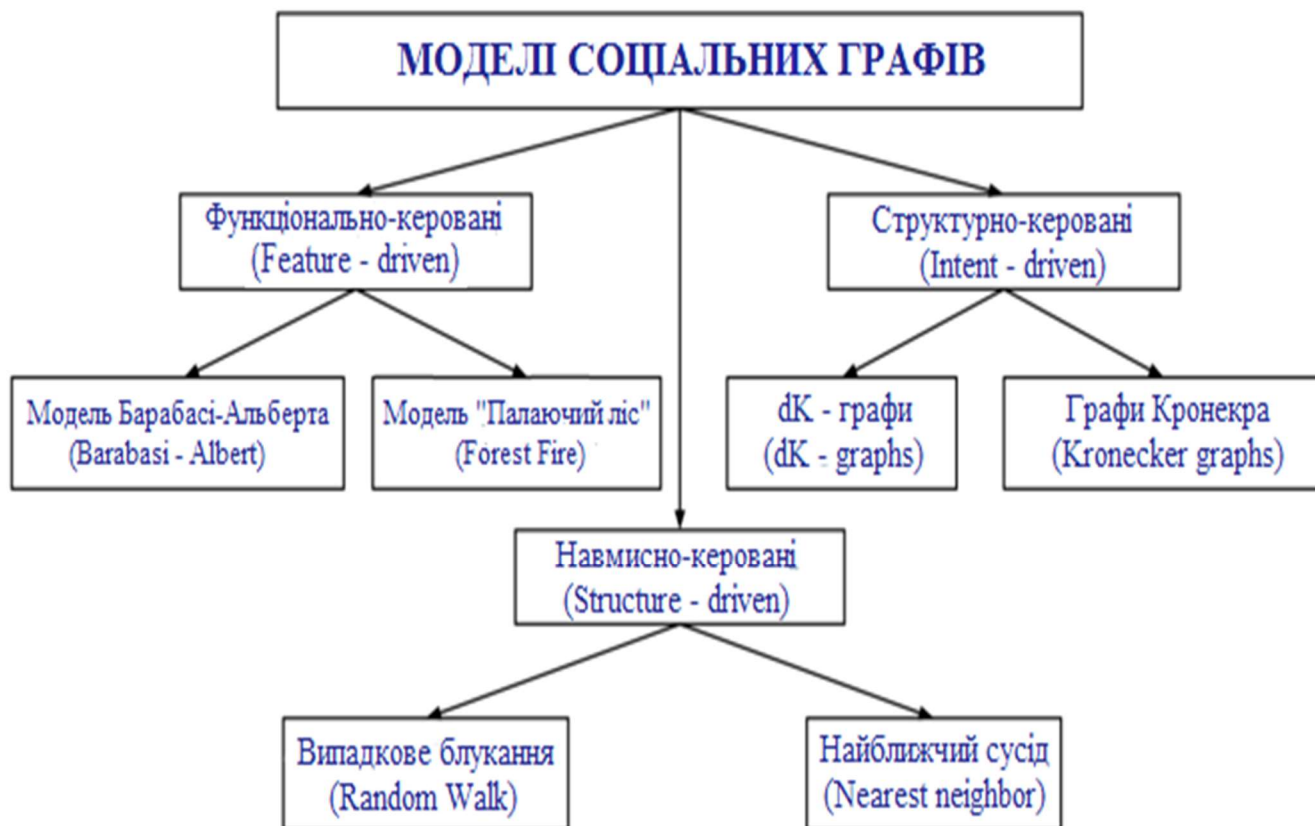


Рис. 1.2. Моделі соціальних графів [15]

Навмисно-керовані моделі зосереджуються на емуляції процесу створення оригінальної структури графа. До цієї групи моделей належать:

1. Модель випадкових блукань (Random Walk) реалізує випадковий процес, де об'єкт виконує послідовні переходи між вузлами, обираючи їх випадковим чином, поки всі вузли не будуть відвідані. Під час цього процесу краї та вершини можуть бути відвідані декілька разів.

2. Модель найближчого сусіда (Nearest Neighbor) – направлений граф.

Структурно-керовані моделі зводять статистичні дані про структуру графа, що дозволяє генератору створювати випадкові графи з такими ж структурними обмеженнями. Ось декілька прикладів таких моделей:

1. Модель графів Кронекера (Kronecker graphs) - ця модель дозволяє генерувати послідовність графів шляхом повторення базового графу Кронекера. Базовий граф поступово збільшується шляхом взаємного поширення його елементів.

2. Модель DK-графів (dK-graphs) - ця модель створює графі з заданим розподілом ступенів вузлів. Вона дозволяє контролювати розподіл ступенів в графі, що може бути корисним для відтворення реальних мереж з відомими властивостями ступенів вузлів.

Таким чином, структурно-керовані моделі використовують статистичні дані про структуру графа для генерації випадкових графів з аналогічними структурними характеристиками.

1.3 Моделі динамічного аналізу мереж

Модель мережі можна визначити як певний об'єкт – замітник об'єкта-оригіналу, де у нашому випадку об'єктом-оригіналом є соціальна мережа. Застосування моделей викликано тим, що мережа може мати надзвичайно складну структуру та для її аналізу ефективніше використовувати математичну модель цієї мережі, яка описує основні елементи структури мережі. За допомогою правильно обраної моделі можна досить ефективно імітувати поведінку цієї мережі та аналізувати її поведінку протягом часу [16].

Динамічний мережевий аналіз (від англ. «Dynamic network analysis» (DNA)) – наукова галузь, що містить традиційний аналіз соціальних мереж (SNA), аналіз посилянь, соціальне моделювання та багатоагентні системи у межах мережевої науки та теорії мереж [17] динамічний мережевий аналіз реалізується за двома напрямками: статистичний аналіз мережевих даних та моделювання для вирішення проблем мережної динаміки. Мережі DNA відрізняються від традиційних соціальних мереж тим, що вони більші, динамічніші, а також можуть містити різні рівні невизначеності. Основна відмінність DNA від SNA полягає у тому, що DNA враховує взаємодію соціальних особливостей, що зумовлюють структуру та поведінку мереж, а також пов'язаний із темпоральним аналізом. Метод базується на ідеї вивчення взаємодій соціальних явищ шляхом порівняння їх часових характеристик.

Статистичні інструменти DNA, як правило, оптимізовані для великомасштабних мереж і допускають аналіз декількох мереж одночасно, в яких існує декілька типів вузлів та декілька типів зв'язків. Багатовузлові мультиплексні

мережі зазвичай називаються метанережами або багатовимірними мережами. На відміну від цього, статистичні інструменти аналізу соціальних мереж фокусуються на одиничних даних та полегшують аналіз використанням лише одного типу зв'язку за один раз.

Статистичні інструменти DNA надають більше метрик для вивчення користувача, тому що вони включають метрики, які використовують дані, отримані з кількох мереж одночасно. Для вивчення еволюційних характеристик соціальної мережі використовують різні системні підходи [17].

Опишемо існуючі відомі моделі для вивчення еволюції соціальної мережі:

– Гравітаційна модель мережі (GM) [19] – модель взаємодії між просторовими об'єктами графа у різних модифікаціях. Для цих моделей характерна особливість, що сила взаємодії (інтенсивність взаємодії) залежатиме від ступеня значущості та міри відстані між об'єктами графа, що розглядаються.

– Експоненційні моделі випадкових графів (STERGM, ERGM) [20, 21]. В експоненційних моделях випадкових графів існують механізми, які можуть відповісти на питання про причини виникнення мережі та її еволюцію. Еволюція та формування такої мережі розглядається як дискретний випадковий процес. Основне завдання аналізу у вивченні таких систем полягає у виявленні причинно-наслідкових зв'язків, що породили зміну.

– Стохастичні акторно-орієнтовані моделі (SAOM) [22]. У стохастичній акторно-орієнтованій мережі, процеси, що протікають, повністю залежні від попереднього стану мережі. У такій мережі її динамічні зміни розглядаються як безперервний процес, де кожен користувач (об'єкт) контролює свою взаємну позицію та властивості. Тут вивчаються не структура мережі та умова її формування у початковий момент часів, а метаморфізм досліджуваних змін.

Аналіз робіт, спрямованих на дослідження динамічних характеристик соціальної мережі показує, що кількість ребер зростає за лінійним законом із появою нових вершин, тоді як щільність графа змінюється за поліноміальною залежністю.

1.4 Порівняльна характеристика програмних засобів для аналізу соціальних мереж

Аналіз соціальних мереж є важливим інструментом для дослідження, зокрема для розуміння поведінки користувачів, виявлення трендів та здійснення прийняття рішень. Існує багато програмних засобів, призначених для аналізу соціальних мереж, і у кожного з них є свої унікальні властивості і переваги.. Розглянемо порівняльний аналіз декількох популярних програмних засобів для аналізу соціальних мереж.

Ражек [23] – програмний засіб, що використовується для візуалізації масштабних мереж. Ця програма використовується для пошуку масштабних графів, що мають десятки тисяч вершин. Ражек підтримує формати файлів «.raj» і «.net». Програмний засіб є безкоштовним для некомерційного використання. Ражек пропонує різноманітні функції візуалізації, особливо для машинозчитуваних мереж, таких, наприклад, як інтелектуальний аналіз даних, мережі цитування, мережі співпраці.

Функціонал Ражек:

- полегшує та забезпечує спосіб візуалізації даних;
- корисний для впровадження великих масштабних мереж;
- здатний розбивати велику мережу на менші мережі для використання з метою аналізу;
- не вимагає навичок програмування;
- корисний для детального аналізу, наприклад, для виявлення порожнин у структурі або обчислення різних мережевих показників тощо.

Gerhi [24] – це одна із найпопулярніших програм із відкритим вихідним кодом, яка часто використовується для дослідження мереж та їх візуалізації. Вбудовані модулі, які доступні в Gerhi, можна імпортувати, маніпулювати, фільтрувати, візуалізувати та експортувати.

Перевага цієї програми перед іншими полягає у тому, що вона розширювана завдяки наявності різноманітних плагінів для розробки. Мета Gerhi — висунути гіпотезу та знайти шаблон і помилки на етапі отримання даних. Типовим режимом Gerhi є 2D, але він також може використовуватися із залученням механізмів 3-D візуалізації.

Спільноти у Gephi можна виявляти за допомогою методів «serval modularity». Крім того, користувач, взаємодіючи з Gephi, може визначати приховані властивості, маніпулюючи структурою графу.

IGraph [25] – безкоштовна програма для створення графу. За замовчуванням містить остовне дерево графу та функції мережевого потоку. Ці особливості є частиною теорії графів. IGraph дозволяє користувачам розгортати мільйони вузлів і ребер без особливих труднощів. IGraph можна легко встановити на будь-якій машині, використовує Ruby, R, C++ і Python [18].

Ключові функції IGraph:

- містить декілька алгоритмів і має просту структуру реалізації;
- IGraph побудований на основі мови високого рівня R;
- дозволяє швидко створювати прототипи;
- може легко обробляти велику кількість об’ємних графів, які мають мільйони вершин і ребер;
- надає низку функцій для алгоритмів виявлення спільноти, таких як метод випадкового блукання та жадібні алгоритми.

NetMiner [26] – програма, яка використовується для візуалізації мережі. NetMiner допомагає виявити мережеву структуру та її основний шаблон. Широко використовується у соціальних мережах а, отже, у мережевому аналізі.

NetMiner складається з різних алгоритмів, які використовуються для виявлення спільноти, наприклад, розповсюдження міток, «edge Betweenness» і «Blondel». Крім того, повністю підтримує різні макети, такі як кластерний шаруватий і круговий.

Основні характеристики:

- добре підходить для аналізу великої мережі;
- реалізовує функції мережевого моделювання.
- гарна якість та надійність функцій керування даними.
- корисний у візуальній аналітиці.

Узагальнене порівняння розглянутих програмних засобів представлено у табл. 1.2.

Загальне порівняння

Програмне забезпечення	Rajek	Gephi	IGraph	NetMiner
Тип	автономний	автономний	бібліотека	автономний
Написано на	.net	Java, OpenGL	Python	Python
Ліцензія	Безкоштовна ліцензія для некомерційного використання	Відкритий вихідний код	Відкритий вихідний код	Відкритий вихідний код
Платформа	Windows	Windows	C/R/Python	Windows
Час обчислення	Швидко	Швидко	Швидко	Швидко
Кількість вузлів	1 млн.	0.15 млн.	10 млн.	1 млн.

У табл. 1.3 узагальнене порівняння основних мережевих показників та функцій для аналізу графів [28]:

Degree Centrality (ступінь центральності) – визначається як кількість зв’язків, що потрапляють на вузол. Більш центральним вважається вузол з високим ступенем. У випадку спрямованої мережі є дві міри; «вихідний ступінь» і «вхідний ступінь». «Out-degree» – це кількість з’єднань, що ведуть до цього вузла, «in» – це кількість з’єднань, що ведуть до вузла.

Clustering coefficient (коефіцієнт кластеризації) є мірою вузлів у графі, це міра кількості трикутників.

Betweenness Centrality (центральність) є мірою центральності та базується на комунікаційному потоці та найкоротших шляхах, доступних у мережі. Найцікавішою частиною графа є кількість вузлів, які мають високу центральність, оскільки вони лежать на різних шляхах зв’язку. Не існує стандартного способу обчислення Betweenness Centrality.

Closeness Centrality (центральність близькості) – визначає наскільки вузол знаходиться близько до інших вузлів у мережі.

Network Diameter (діаметр мережі) – діаметр мережі графа насправді є геодезичною мірою між вузлами, зазвичай він приймає значення від 1 до $G-1$ на графі.

Page Rank (рейтинг сторінки) – алгоритм, що використовується пошуковою системою Google і, крім того, використовується для ранжування сторінки.

Partition (розбиття) – полягає у тому, що граф «G» ділиться на менші компоненти з певними властивостями.

HITS (The Hyperlink-Induced Topic Search) – алгоритм, що дозволяє аналізувати посилання, використовуючи пошук за темою за гіперпосиланням (HITS). Найчастіше цей алгоритм в основному застосовують для оцінювання веб-сторінки. Веб-сторінки ранжуються шляхом обробки усіх вихідних посилань.

Таблиця 1.3

Порівняння мережевих показників та функцій для аналізу графів

Метрики мережі	Pajek	Gephi	IGraph	NetMiner
Degree Centrality	☑	☑	☑	☑
Clustering coefficient	☑	☑	☑	☑
Betweenness Centrality	☑	☑	☑	☑
Closeness Centrality	☑	☑	☑	☑
Network Diameter	☑	☑	☑	☑
Page Rank	✗	☑	☑	☑
Partition / розділення	☑	☑	✗	✗
HITS	✗	☑	✗	☑

Однією з найважливіших властивостей соціальної мережі є структура спільнот. Для того, щоб виявити присутність спільноти у графі, на даний момент розроблено ряд алгоритмів. У табл. 1.4 описане порівняння кількох алгоритмів виявлення спільнот, а саме:

Statistical Method (статистичні методи): моделювання відпау і методи spin – glass використовуються для виявлення присутності різних спільнот на графі.

Louvain Method (метод Лувена) є відомим алгоритмом на основі аналізу модульності. У цьому алгоритмі вершини об'єднуються у вигляді спільноти.

Greedy Method (жадібний алгоритм) є одним із популярних класичних методів, які використовуються для визначення наявності структури спільнот. Цей алгоритм самостійно реалізує структуру жадібної оптимізації.

Таблиця 1.4

Порівняння алгоритмів виявлення спільнот

Алгоритм	Pajek	Gephi	IGraph	NetMiner
Statistical Method	✗	✗	✗	✗
Louvain Method	☑	☑	☑	☑
Greedy Method	✗	✗	☑	✗
Leiden Method	✗	☑	☑	✗

У табл. 1.5 запропоноване порівняння «функцій алгоритму», включаючи час завантаження, модульність, рейтинг сторінки та діаметр. У цьому порівнянні користуємось даними, де використовували набір даних egonets – Facebook.

Таблиця 1.5

Час виконання для різних інструментів

Алгоритм	Pajek	Gephi	IGraph	NetMiner
Degree centrality	2 с	4 с	6.199 с	15 с
Load time	3 с	29 с	3.707 с	40 с
Network diameter	Не застосовується	120 с	3.51 с	15 с
Page Rank	Не застосовується	10 с	9.81 с	Не застосовується
Modularity	6 с	30 с	9 с	14 с

Розглянувши різні складні мережеві інструменти, такі як: Gephi, Pajek, IGraph і NetMiner, після проведеного порівняльного аналізу можна зробити висновок, що для аналізу соціальних мереж існує ряд програмних засобів з різними функціональними можливостями. Вибір конкретного засобу залежить від потреб користувача, його рівня експертизи, типу даних та вимог до аналізу.

Якщо важливо мати потужні інструменти візуалізації та можливість редагування графічного представлення соціальної мережі, Gephi може бути відмінним вибором.

В аналізі соціальних мереж одним із важливих завдань є виявлення та аналіз структури спільноти. Зазвичай це завдання стосується функціональної організації мережі. Дані мережі можна зрозуміти за допомогою аналізу, а подальші результати можна ефективно передати. Наше дослідження повністю зосереджено на різних

алгоритмах, які використовуються для виявлення спільнот. Розглянули кілька мережевих метрик на основі графів, таких як складність алгоритму та мережеві показники разом із їх загальним порівнянням.

Висновки до розділу 1

Таким чином, перший розділ містить стислий огляд еволюції аналізу соціальних мереж. Розглянуто основні типи соціальних мереж та методологію їх аналізу, теоретичні засади аналітики та візуалізації соціальної мережі. Зроблено висновок, що найоптимальнішим інструментом моделювання соціальної мережі є традиційна теорія графів. Розглянуто основні метрики та методи аналізу соціальних мереж, порівняли інструменти для візуалізації графів, їх переваги та недоліки. На основі порівняльної характеристики зроблено висновок, що найефективнішим інструментом для візуалізації є Gephi.

Аналіз соціальної мережі відкриває додаткові можливості для дослідження соціальних об'єктів та зв'язків між ними. Це створює умови розвитку досліджень у багатьох сферах, де важливо враховувати соціальні настрої, вподобання, формувати суспільну думку соціальних груп, структур чи окремих об'єктів, реалізовувати маркетингові стратегії розвитку бізнесу та здійснювати управління поширенням даних в інформаційному просторі.

РОЗДІЛ 2 МОДЕЛЮВАННЯ ТА ВІЗУАЛІЗАЦІЯ СОЦІАЛЬНИХ МЕРЕЖ

2.1 Опис програмних засобів моделювання для аналізу та візуалізації

Як було зазначено раніше, найбільш зручним та оптимальним методом аналізу і візуалізації соціальної мережі є соціальний граф. У попередньому розділі нами досліджено, що найпопулярнішою програмою візуалізації і аналізу мереж та графів на сьогодні є Gephi [16]. Gephi дозволяє забезпечити швидку організацію та дієву фільтрацію, а також надає можливість для інтерактивного дослідження даних. Саме тому, Gephi вважається одним з найефективніших інструментів для візуалізації великомасштабних мереж.

Завантажити доступні версії Gephi для Mac OS X, Windows і Linux можливо за адресою <https://gephi.org/>. Варто зазначити, що для стабільної роботи програми необхідно мати Java v7 і вище. Gephi підтримує різноманітні формати файлів, які можуть зберігати атрибути вузлів і країв, а також макет і представлення інформації (наприклад, положення, розмір, колір тощо).

Власний формат – це файл .gephi, який може містити кілька «робочих областей» (тобто кілька різних графів).

Gephi також підтримує декілька відкритих форматів, які можна використовувати для обміну даними з іншими інструментами. В Gephi можливо завантажити дані в наступних форматах: таблиці Excel, GraphML, GML, Pajek (NET), GEXF, CSV, GDF, GraphViz (DOT), UCINET (DL), Netdraw (VNA) і Tulip (TPL). Таким чином, Gephi дозволяє розширити можливості користувача та дозволяє взаємодіяти з іншими зовнішніми програмами й системами аналізу і візуалізації графів. Gephi підтримує різноманітні формати файлів, які можуть зберігати атрибути вузлів і країв, а також макет і представлення інформації (наприклад, положення, розмір, колір тощо).

Також перевагою даного програмного забезпечення є наявність широкого спектру алгоритмів, які дозволяють обирати різну компоновку графу на площині, підлаштовувати налаштування кольорів та розмірів під побажання користувача, а також додавати мітки до графів.

Для того щоб збільшити функціональність, дозволити користувачу додавати власні макети, створювати нові алгоритми та використовувати різноманітні інструменти вимірювань, в Gephi є можливість завантажувати додаткові плагіни.

Gephi використовує мультипотокową архітектуру обробки даних, що дозволяє виконувати різні види аналізу одночасно.

Інтерфейс користувача програмної платформи Gephi включає три основні розділи (вікна):

Вікно «Лабораторія даних» містить всі вихідні дані про мережі, основні та додаткові значення для розрахунків.

Вікно «Обробка даних» включає величезний об'єм різноманітних операцій, доступних користувачеві. Наприклад, ручне редагування вигляду графу, можливість тестування макетів, встановлення різних значень фільтрів для ефективної фільтрації.

Вікно «Попередній перегляд» дозволяє уточнити форму виводу графу. Це можливо завдяки набору інструментів, які допомагають допрацювати граф та досягнути бажаного візуального результату. Це вікно також викликає експорт графу у формати PDF, PNG і SVG.

Існує три основні режими створення нового графа в Gephi: через інтерфейс «Граф» в режимі «Обробка»; через «Лабораторію даних»; через експорт даних графа із зовнішнього файлу (найпростіше з файлу в форматі CSV з роздільниками крапка з комою).

Режими дозволяють налаштувати індивідуальне зафарбовування або використати функцію зафарбовування всіх сусідів обраного вузла.

Для завантаження в Gephi спочатку треба підготувати файл для графа. Це можливо зробити, наприклад, у Excel, а потім зберегти його у форматі CSV. Процес обробки результатів представлений на рис. 2.1.

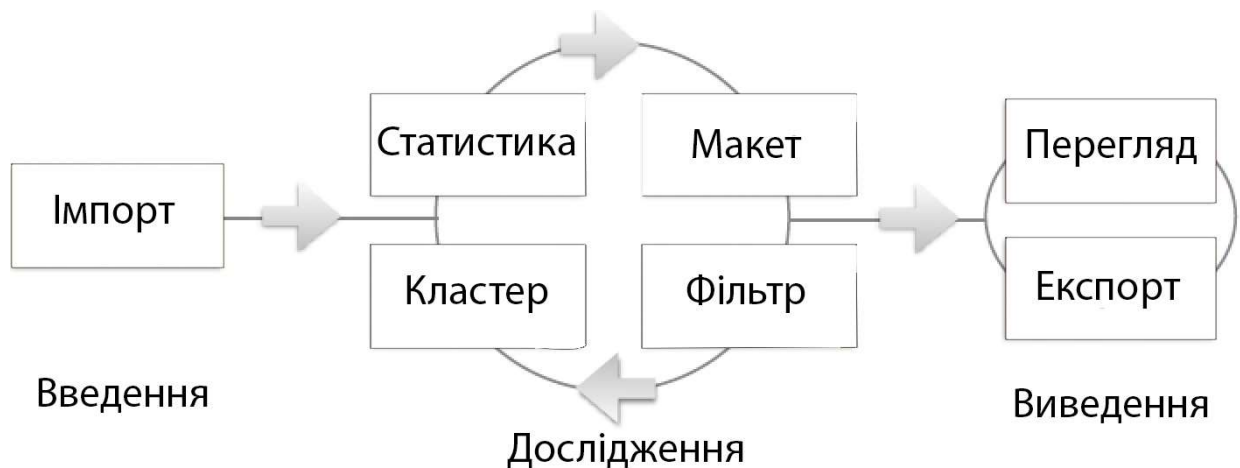


Рис. 2.1 Схема роботи Gephi у процесі візуалізації графу

2.2 Практична реалізація моделей соціальних мереж

Створимо дві моделі соціальної мережі:

1. Модель «Наукова спільнота студентів університету» – тренувальна, генерується із табличних файлів.

2. Модель «Наукометрична база Scopus» – для побудови графу засобами Python [27] із залученням бібліотеки gephi.

Обидві моделі представляють відомий тип соціальної мережі – мережу співавторства або граф співпраці (Collaboration graph). Вершини – це учасники мережі, ребро між двома вузлами вказує на наявність певного зв'язку між ними. Одним з найвідоміших прикладів такого графу співавторства математиків є граф Ердоша.

Візуалізація графу соціальної мережі «Наукова спільнота студентів університету»

Створюємо файл `students_nodes.csv`, який містить унікальні значення вершин (`id`). Додаткові значення з мітки стовпця `label` розбиваємо дані за допомогою фільтра на відповідні п'ять груп (проект, студент, кафедра, тема проекту) і записуємо у стовпець `Group` – ця процедура дозволить обирати кольори для забарвлення вузлів.

	A	B	C
1	id	Label	group
2	1	проект1	проект
3	2	проект2	проект
4	3	проект3	проект
5	4	проект4	проект
6	5	проект5	проект

Рис. 2.1 Формування файлу students_nodes.csv

Файл students_edges.csv містить інформацію про всі можливі переходи між вузлами. У цьому файлі присутні три поля: src, trgt і name_edg. Обов'язковими полями є лише src і trgt, а поле name_edg створюється автоматично, щоб додавати підписи на ребрах.

	A	B	C
1	src	trgt	name_edg
2	1	674	тема_проекта
3	2	675	тема_проекта
4	2	676	тема_проекта
5	3	677	тема_проекта
6	4	678	тема_проекта

Рис. 2.2 Формування файлу students_edges.csv

Під час завантаження даних необхідно обрати тип графу (орієнтований, неорієнтований або змішаний) та встановити прапорець «Додати до існуючої робочої області», щоб завантаження відбувалося до однієї робочої області. Кількість вузлів – 812 та ребер – 1625.

Після завантаження усіх даних переходимо до розділу «Укладка» (Layout) та обираємо укладку відповідно до топології, яку бажаємо виділити. Наприклад, алгоритми ForceAtlas добре підходять для відображення класифікації (групування), вони дозволяють налаштовувати сили відштовхування та притягання вершин.

Після завантаження даних отримуємо граф, який розділений на кластери. У нашому випадку вершини, які представляють студентів, згрупувалися навколо вершин – кафедр. Було обрано укладку ForceAtlas з силою відштовхування 1000.

Після запуску на виконання отримаємо граф із сірими вузлами (рис. 2.3) але бачимо, що студенти «згрупувалися» навколо кафедр.

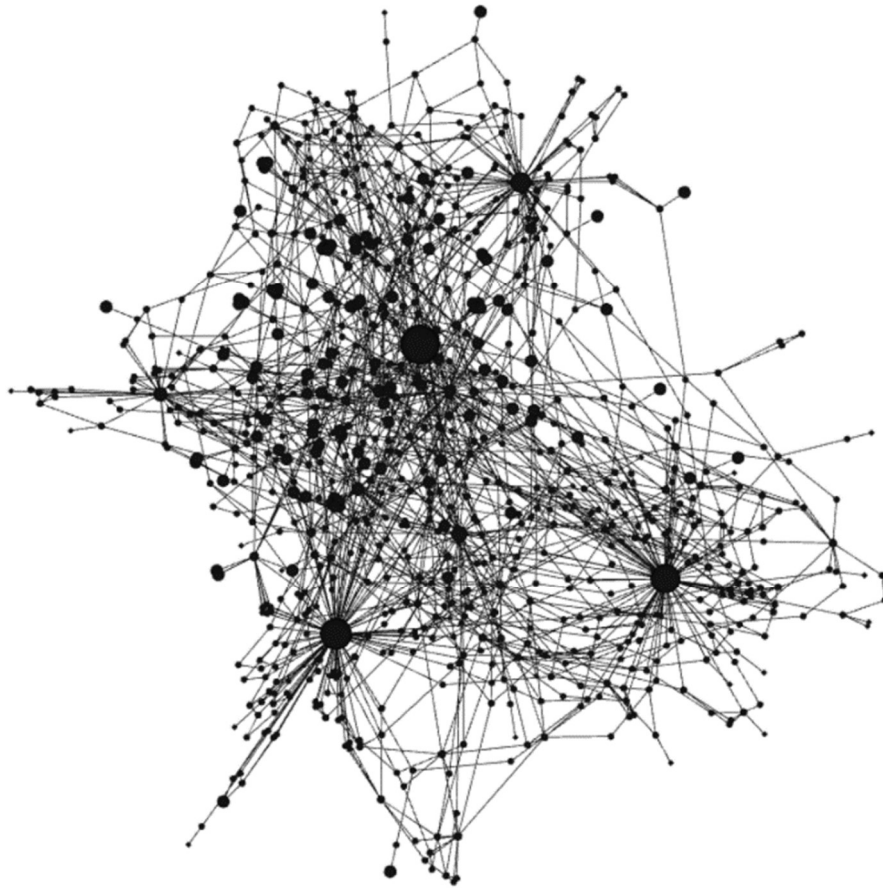


Рис. 2.3 Формування файлу `students_edges.csv`

Відкорегуємо розміри вершин. Для цього у блоці «Зовнішній вигляд» (вкладка «Вершини») обираємо розмір «Ранжування». Зі спадного списку обираємо «Ступінь» та встановлюємо розміри вершин від тієї, що зустрічається найменше (мінімальний розмір) до найпопулярнішої вершини (максимальний розмір) – встановлюємо значення від 10 до 50. Фрагмент графа, що отримали представлений на рис. 2.3.

Блок «Зовнішній вигляд» дозволяє налаштувати кольори вузлів та ребер. Розгорнувши палітру, обираємо тип забарвлення: унікальне (монохромне забарвлення), розподіл (може мати більше двох кольорів, все залежить від кількості груп/класів, на які поділили вузли), ранжування (плавний колірний перехід від менш популярних до популярних вузлів).

На зображенні рис. 2.4 представлено забарвлення «Partition» за полем «група» (це поле створено вручну).



Рис. 2.4 Використане забарвлення «Partition» за полем «група» (створено вручну).

Тепер залишилося підписати та відфільтрувати дані. Переходимо до підпису назв вузлів та ребер на графі. Розкриваємо панель, розташовану під графом.

Встановлюємо прапорці «Вузол» та «Ребро» і налаштовуємо шрифт та розмір. Розмір тексту відносно вузлів налаштовується за допомогою повзунка. Ця панель також дозволяє змінювати колір фону (вкладка «Загальне»), видаляти ребра та змінювати їх розмір (вкладка «Ребра»).

Якщо цікавлять усі проекти з темою під певним номером – перетягуємо «Ego Network» (з папки «Топологія») у вікно «Набори фільтрів» нижче. Встановлюємо значення «ID вузла» рівним «тема №» і застосовуємо фільтр. Якщо потрібно переглянути, які студенти брали участь у цих проектах, збільшуємо глибину (зазначаємо 2).

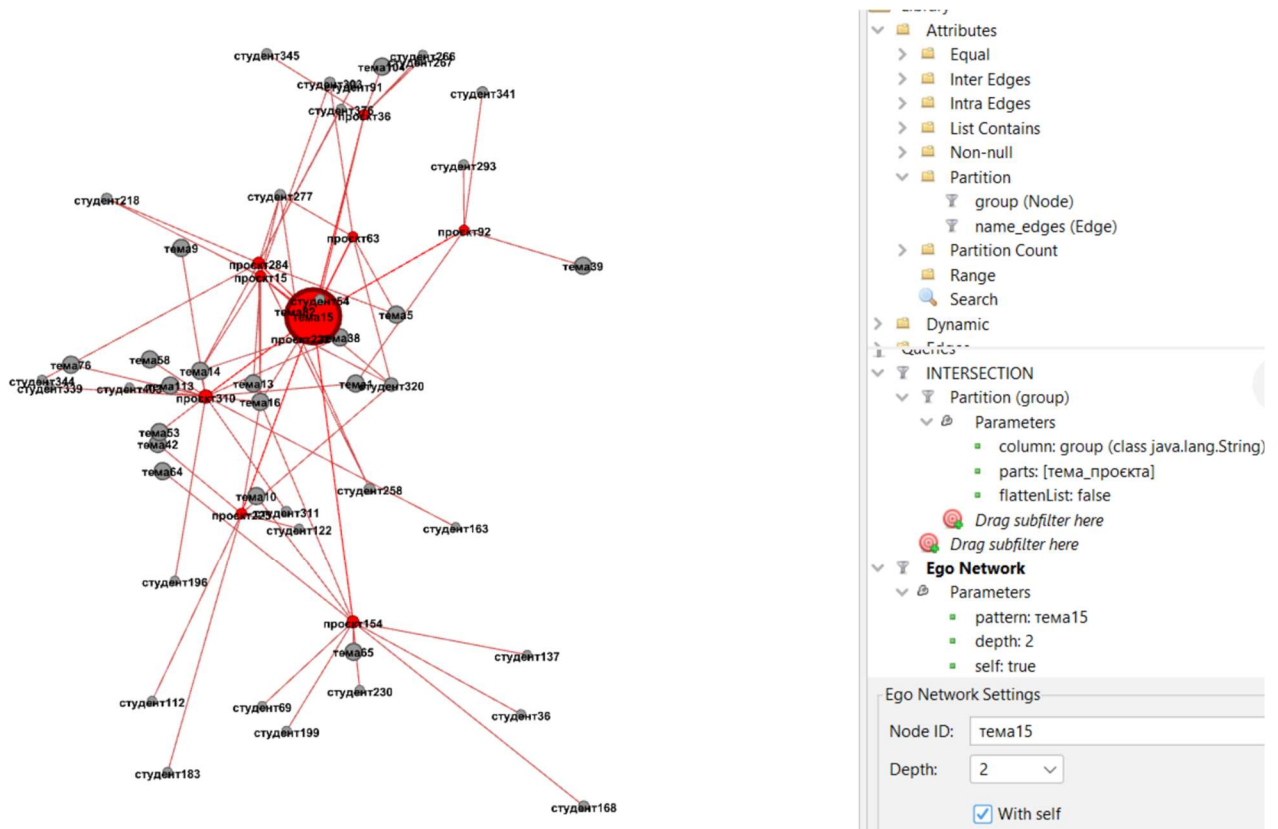


Рис. 2.5 Фільтрація даних на прикладі «Теми 15»

Щоб продемонструвати теми проєктів, якими займалися студенти кафедри №3. Спочатку відфільтруємо дані за кафедрою і розширимо глибину до 3, щоб включити назви тем (рівні глибини: 1 – студенти, 2 – проєкти, 3 – тема проєкту).

Оскільки нас цікавлять лише теми проєктів, залишаємо лише потрібні нам вузли. Додаємо у набір фільтрів оператор AND і атрибут group (створений з поля label) з вказаним значенням. У параметрах атрибута вказуємо «тема проєкту». Граф готовий. З масиву вузлів у Gephi отримали зрозумілий граф для аналізу з можливістю фільтрації та візуалізації.

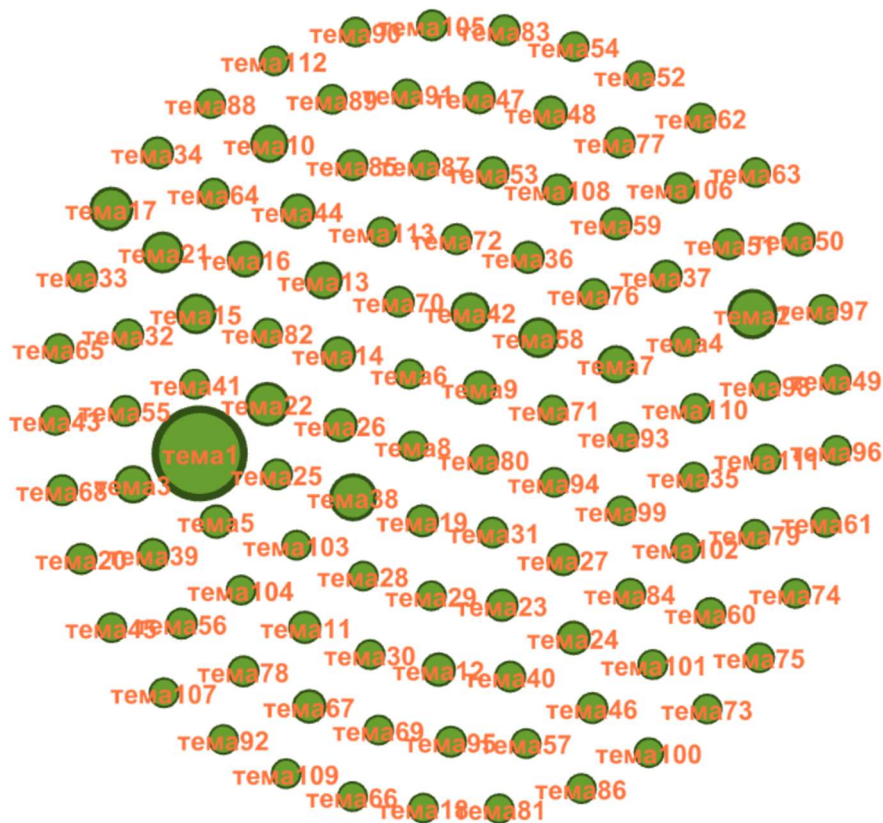


Рис. 2.6 Граф тем студентів «кафедри №3»

Моделювання та візуалізація графу соціальної мережі «Наукометрична база Scopus»

Виконаємо моделювання графу соціальної мережі за допомогою бібліотеки Python – `gephi`. Для цього використаємо базу даних Scopus. Можна також застосувати процедуру парсингу або скрапінгу для отримання необхідних для обчислення даних.

Щоб використати бібліотеку `gephi` у Python потрібно написати лише рядок імпорту (рис. 2.7).

```
from __future__ import print_function, division
from gephi import gephi
```

Рис. 2.7 Фрагмент коду для імпорту бібліотеки `Gephi`

Далі за допомогою методів побудови вузлів та ребер будується граф мережі Scopus (рис. 2.8). Утворення кластерів (співтовариств) на графі можна спостерігати візуально (виділені різними кольорами) але також можна застосувати і алгоритм знаходження кластерів (рис.2.9) або вбудований плагін алгоритму Лейдена.

```

import os
import gephi as gp
import pandas as pd

# каталог де знаходяться всі експортовані ключові слова
folder = "scopus exports/"
fnames = [folder+f for f in os.listdir(folder)]

# справжні суми для ключових слів із більшою кількістю публікацій, ніж
# максимальний ліміт 20 000 для завантаження
limited_node_sizes_fname = "True Amounts.xlsx"
limited_node_sizes = pd.read_excel(limited_node_sizes_fname,
                                  header=None, index_col=0, squeeze=True)

# розрахувати вузли
# include_internal_similarity=True для обчислення внутрішньої подібності
# False для швидкої роботи
nodes = gp.get_nodes(fnames, limited_node_sizes=limited_node_sizes,
                    includes_internal_similarity=True)
nodes.to_excel("out/nodes.xlsx")

# обчислити ребра
# if using Lens exports (patents or scholarly both work) instead of
# Scopus, set database="lens"
edges = gp.get_edges(fnames, limited_node_sizes=limited_node_sizes)
edges.to_excel("out/edges.xlsx")

```

Рис. 2.7 Фрагмент коду для візуалізації графу Scopus

```

import gephi as gp
import pandas as pd

nodes = pd.read_excel("nodes_to_analyze/Nodes_clustered.xlsx", index_col=0)
edges = pd.read_excel("nodes_to_analyze/edges.xlsx", index_col=0)

clusters = gp.get_cluster_info(nodes, edges)
clusters.to_excel("out/Clusters.xlsx")

nodes_out = gp.get_cluster_info_nodes(nodes, edges)
nodes_out.to_excel("out/Nodes_with_cluster_info.xlsx")

```

Рис. 2.8 Фрагмент коду для алгоритму кластеризації графу Scopus

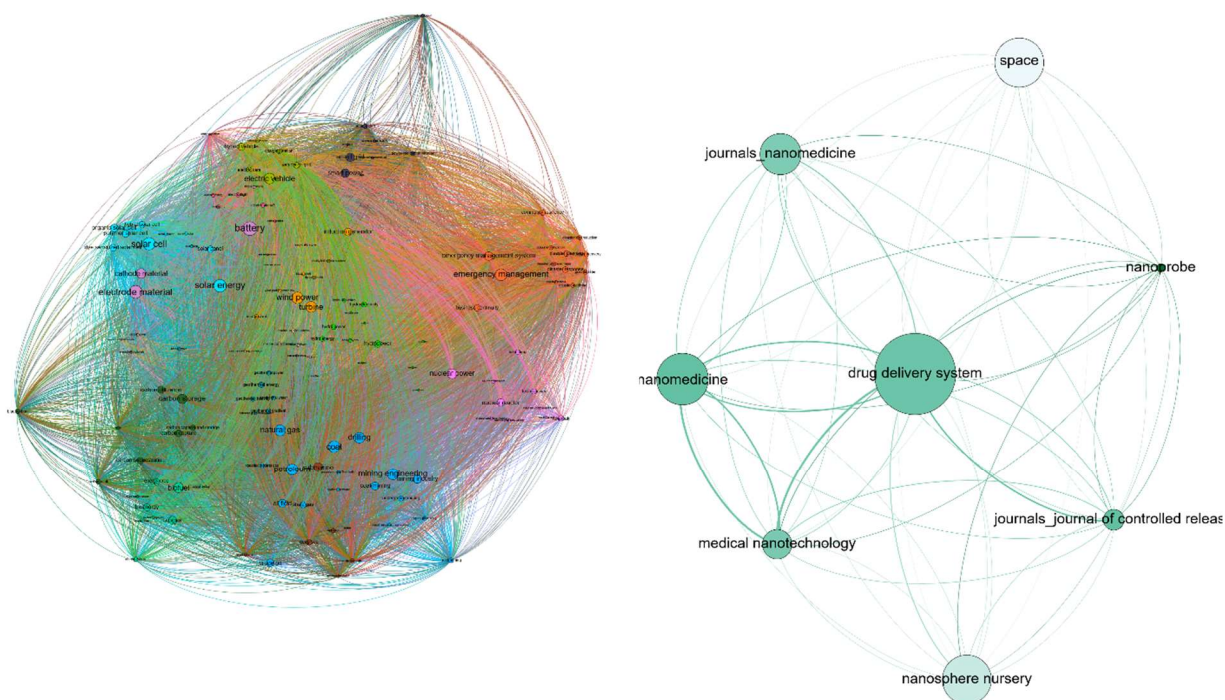


Рис. 2.9 Візуалізація графу наукометричної бази даних Scopus, побудованого у Python (зліва повний граф, кольором виділені співтовариства, справа – результат алгоритму кластеризації)

Аналіз метрики SNA допоможе виділити ключові вузли у мережі та зробити кількісний висновок щодо зміни характеру наукової мережі співавторства.

Тільки на основі візуальної оцінки графів уже можна зробити кілька висновків. Враховуючи, що розмір вузла на рисунку відображає вагу, тобто кількість робіт у одному проєкті.

Методи SNA (аналіз соціальних мереж) надають унікальний набір інструментів для обробки складних структурованих великих обсягів даних, що дозволяє оцінити характер процесів які протікають (причому як якісно, так і кількісно), зазирнути за межі звичайної статистики. Робота з соціальною мережею з точки зору теорії графів дозволяє встановити значущість окремих осіб у загальній картині соціальної взаємодії, виділити підгрупи пов'язаних вузлів та принципи поширення інформації у

мережі. У прикладній задачі вивчення структури соціальної мережі наукового співавторства метрики SNA можуть допомогти у виділенні найбільш продуктивних науковців, наприклад, для подальшої мотивації, а також прогнозу динаміки такої мережі для задач моніторингу та управління. Наступною задачею може бути дослідження такої соціальної мережі у динаміці для всієї організації, що вимагатиме обробки значно більших обсягів даних.

Висновки до розділу 2

У реалізації практичної частини створено візуалізації для двох моделей соціальних мереж. Аналіз та візуалізація моделей мереж здійснювалась програмним засобом Gephi та написанням коду у Python із використанням бібліотеки `gephi`.

Створити граф можна за допомогою файлів відповідних форматів (`csv`, `xls`). Для отримання даних для побудови можна скористатися парсингом (або скрапінгом) або заповнити матриці вручну. Граф мережі Scopus побудований за допомогою методів побудови вузлів та ребер. Утворення кластерів (співтовариств) на графі можна спостерігати візуально (виділені різними кольорами), але також можна застосувати і алгоритм знаходження кластерів або вбудований плагін алгоритму Лейдена.

ВИСНОВКИ

У роботі досліджені теоретичні, методологічні та програмні аспекти аналізу еволюції структури соціальної мережі з використанням візуалізації в Gephi.

Грунтовний огляд наукової та методичної літератури щодо аналізу соціальних мереж дозволив отримати методологію для подальшого дослідження.

Детально вивчено основні алгоритми та методи аналізу соціальних мереж, такі як виявлення важливих вузлів, спільнот, визначення центральності тощо. Це надало необхідні інструменти для подальшого експериментального аналізу.

За допомогою мови програмування Python та використання відповідних бібліотек здійснено аналіз структур двох моделей соціальних мереж. Зібрано та оброблено дані, змодельовані на Python та візуалізовані у програмному засобі Gephi.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Moreno J. L. Who Shall Survive? Foundations of Sociometry, Group Psychotherapy and Sociodrama. Sociometry monographs (V 29), V2, Beacon House, 1953. 763 p. URL: <https://archive.org/details/whoshallsurvivef00more>.
2. Сушко В. О. Історія виникнення та методологія мережного аналізу // Журнал соціології та соціальної антропології. 2018. Т. 21, № 1. С. 161–181. URL: <https://doi.org/10.31119/jssa.2018.21.1.7>.
3. Jones D. Sociometry and Social Network Analysis: Applications and Implications / ANZPA Journal, 15: 80. Jones D. Sociometry and Social Network Analysis: Applications and Implications // ANZPA Journal. 2006. P. 80.
4. A course on complex network analysis by Albert – László Barabási Network Science. URL: <http://barabasilab.neu.edu/courses/phys5116/>.
5. McFarland D. Social Network Analysis Labs in R / Daniel, McFarland, Solomon Messing, Michael Nowak, Sean J. Westwood. – Stanford University. 2010. URL: <http://sna.stanford.edu/rlabs.php>].
6. Coursera. URL: <https://www.coursera.org/>].
7. Google Trends. URL: <http://www.google.com/trends>.
8. Barnes J. A. Class and Committees in Norwegian Island Parish // Human Relations. 1954. V. 7. P. 39–58. URL: <http://pierremerckle.fr/wp-content/uploads/2012/03/Barnes.pdf>.
9. O'Reilly Tim What Is Web 2.0 2005. URL: <http://oreilly.com/web2/archive/what-is-web-20.html>.
10. Якими соцмережами користуються українці під час війни: статистика URL: <https://speka.media/yakimi-socmerezami-koristuyutsya-ukrayinci-pid-cas-viini-doslidzennya-p22nyp>.
11. The state of social media, 2022. [Electronic resource]. – Mode of access: URL: <https://www.meltwater.com/en/resources/state-of-social-2022>.
12. Мазуренко В. В., Штовба С. Д. Огляд моделей аналізу соціальних мереж // Вісник Вінницького політехнічного інституту. 2015. № 2. С. 62–74. URL: http://nbuv.gov.ua/UJRN/vvpi_2015_2_12.

13. Sala A., Cao L., Wilson C., Zablith R., Zheng H., Zhao B.Y. Measurement – calibrated Graph Models for Social Network Experiments // Proceedings of the 19 – th International conference on World wide web. 2010. P. 861–870.
14. 10 Metrics to Track for Social Media Success/ Aguis A. URL: [https://www.socialmediaexaminer.com/10 – metrics – to – track – for – social – media – success/](https://www.socialmediaexaminer.com/10-metrics-to-track-for-social-media-success/).
15. Соціальний граф. URL: https://uk.wikipedia.org/wiki/Соціальний_граф.
16. Newman M.E.J. The structure and function of complex networks // Society for Industrial and Applied Mathematics. 2003. Vol. 45. P. 167–256.
17. Динамічний мережевий аналіз. URL: https://hmn.wiki/uk/Dynamic_network_analysis.
18. Durlauf S.N., Peyton Y .H. Social dynamics – Cambridge, MA: MIT Press, 2004. 238 p.
19. Sen A., Smith T., Gravity Models of Spatial Interaction Behaviour – Heidelberg, Germany: Springer, 1995. 572 p.
20. Holland P. W., Leinhardt S. An Exponential Family of Probability Distributions for Directed Graphs // Journal of the American Statistical Association. 1981. Vol. 76, № 373. P. 33–50.
21. Robins G. Pattison P., Kalish Y., Lusher D. An Introduction to Exponential Random Graph (p^*) Models for Social Networks // Social Networks. 2007. Vol. 29, № 2. P. 173–191.
22. Snijders T. A.B., Bunt G. G., Steglich C. E.G. Introduction to Stochastic Actor – Based Models for Network Dynamics // Social Networks. Vol. 32, № 1. PP. 44–60.
23. Pajek. URL: [http://mrvar.fdv.uni – lj.si/pajek/](http://mrvar.fdv.uni-lj.si/pajek/).
24. Gephi. URL: <https://gephi.org/>.
25. IGraph, URL: <https://igraph.org/>.
26. NetMiner. URL: [http://www.netminer.com/main/main – read.do](http://www.netminer.com/main/main-read.do).
27. Python. URL: <https://python.org/>.

- 28.Amin F., Ahmad A., Choi G. S. Community Detection and Mining Using Complex Networks Tools in Social Internet of Things// In Proceedings of the IEEE TENCON, 2018. PP. 1–6.