



Т. І. Смірнов¹, Ю. В. Шабатура²

¹ Національний лісотехнічний університет України, м. Львів, Україна

² Національна академія сухопутних військ ім. гетьмана Петра Сагайдачного, м. Львів, Україна

ВИЯВЛЕННЯ ТА ІДЕНТИФІКАЦІЯ ГРУПОВИХ ТОПОЛОГІЙ ЯК ВНУТРІШНІХ МІКРОСТРУКТУР У ГЛОБАЛЬНИХ СОЦІАЛЬНИХ МЕРЕЖАХ

Досліджено соціальні мережі як ключовий елемент сучасного інформаційного суспільства, що активно формує як повсякденне життя, так і професійну діяльність громадян. Проаналізовано методи кластеризації для виявлення групових топологій (кластерів) та складних структурних залежностей між учасниками великих соціальних мереж. Встановлено, що соціальні мережі, в контексті стрімкого розвитку цифрових технологій, виступають не тільки платформами для спілкування, а й важливими інструментами обміну інформацією, впливу на громадську думку та організації професійних зв'язків. Розроблено та впроваджено ефективні алгоритми, що оптимізують процес ідентифікації взаємозв'язків у графах соціальних мереж, зокрема алгоритми Louvain та Girvan-Newman. Застосовано новий підхід на підставі штучного інтелекту, який інтегрує методи машинного навчання для розкриття складних взаємозалежностей між учасниками мережі. Це підвищує ефективність аналізу великих даних, забезпечуючи точніше виявлення закономірностей та прогнозування динаміки розвитку соціальних мереж. Проведено серію комп'ютерних експериментів, які містили аналіз ефективності алгоритмів Louvain та Girvan-Newman у кластеризації соціальних мереж, а також застосування методів штучного інтелекту для розкриття складних взаємозалежностей між учасниками мережі. Кожен експеримент складався з кількох спостережень, що дало змогу всебічно оцінити адаптивність та ефективність запропонованих методів у різних умовах та з різними типами даних. Результати проведених експериментів демонструють значне покращення виявлення структурних залежностей з підвищенням рівня навчання моделей, що свідчить про високу адаптивність запропонованих методів до змінних умов та їхню здатність ефективно працювати з різноманітними типами даних. Встановлено, що алгоритм Louvain відзначається високою ефективністю під час оброблення великих масштабів даних, що є критично важливим для аналізу сучасних соціальних мереж, які характеризуються величезною кількістю користувачів і динамічними змінами їхніх взаємодій. Алгоритм Girvan-Newman дає змогу детально розкривати структурні особливості мережі через виділення спільнот на підставі центральності ребер, що сприяє глибшому розумінню внутрішніх механізмів взаємодії між її учасниками. Оцінено стійкість соціальних мереж до зовнішніх впливів, таких як кібератаки, інформаційні війни чи соціальні потрясіння. Аналіз показав, що соціальні мережі здатні зберігати свою функціональність навіть за умов часткової фрагментації або розпаду окремих частин, завдяки високому ступеню децентралізації та наявності резервних каналів зв'язку. Це має важливе значення для забезпечення надійності та стабільності інформаційних екосистем у сучасному суспільстві. Впроваджені результати мають значний потенціал для розроблення нових технологічних інструментів, які забезпечать поглиблений аналіз соціальних мереж, оптимізацію процесів поширення інформації та прогнозування потенційних сценаріїв дезінтеграції мереж. Висновки та рекомендації можуть бути застосовані як у наукових, так і в практичних цілях, сприяючи подальшому розвитку та вдосконаленню інструментів аналізу соціальних мереж, підвищенню ефективності управління інформаційними потоками, покращенню стратегій маркетингу та комунікацій, а також забезпеченню безпеки та стабільності соціальних структур. Це дослідження є важливим внеском у галузь аналізу соціальних мереж і має велике значення для розробників технологій штучного інтелекту, соціологів, аналітиків даних, а також для всіх, хто прагне глибше зрозуміти механізми функціонування соціальних мереж і їхній вплив на сучасне суспільство. Висновки та рекомендації, отримані в процесі дослідження, можуть бути застосовані як у наукових, так і в практичних цілях, сприяючи подальшому розвитку та вдосконаленню інструментів аналізу соціальних мереж.

Ключові слова: кластеризація; теорія графів; машинне навчання; алгоритм Louvain; алгоритм Girvan-Newman; штучний інтелект; структурні залежності; стійкість мереж.

Вступ / Introduction

У сучасному світі соціальні мережі слугують не тільки засобом комунікації між користувачами, але й потужним інструментом для обміну інформацією, організації спільнот, мобілізації громадських рухів і впливу на суспільну думку. Соціальні платформи, такі як Face-

book, X (колишній Twitter) та Instagram, створюють складні багаторівневі структури взаємодій, що об'єднують мільйони користувачів в усьому світі. Ці мережі формують унікальні соціальні екосистеми, де люди об'єднуються за спільними інтересами, географічним розташуванням, професійною діяльністю чи іншими факторами. Проте, такі групи часто залишаються непомітними

Інформація про авторів:

Смірнов Тарас Ігорович, магістр, аспірант, кафедра інженерії програмного забезпечення. Email: smirnov.t@nltu.lviv.ua

Шабатура Юрій Васильович, д-р техн. наук, професор, кафедра електромеханіки та електроніки.

Email: shabaturayuriy@gmail.com; <https://orcid.org/0000-0002-9961-1244>

Цитування за ДСТУ: Смірнов Т. І., Шабатура Ю. В. Виявлення та ідентифікація групових топологій як внутрішніх мікроструктур у глобальних соціальних мережах. Науковий вісник НЛТУ України. 2024, т. 34, № 7. С. 72–79.

Citation APA: Smirnov, T. I., & Shabatura, Yu. V. (2024). Detection and identification of group topologies as internal microstructures in global social networks. *Scientific Bulletin of UNFU*, 34(7), 72–79. <https://doi.org/10.36930/40340709>

ми для зовнішніх спостерегачів, що свідчить про важливість їх аналізу як наукового завдання.

Зростання обсягів даних та активності [1] користувачів у соціальних мережах породжує необхідність розроблення нових підходів до їх аналізу. Структурні елементи соціальних мереж можна представити у вигляді кластерів, які відображають внутрішню організацію соціальних спільнот і є важливими для розуміння того, як функціонують ці мережі [15]. Кластери відображають групи користувачів, які взаємодіють частіше всередині своєї групи, ніж з іншими користувачами мережі. Дослідження таких кластерів дає змогу виявити приховані зв'язки між учасниками мережі, зрозуміти механізми поширення інформації як всередині груп, так і між ними, а також ідентифікувати ключових учасників, які відіграють важливу роль у цих процесах.

Стійкість соціальних мереж до зовнішніх втручань нерідко пояснюють саме їхньою кластерною побудовою, яка сприяє збереженню цілісності навіть у разі часткової фрагментації. Однак, питання стійкості цих мереж, особливо у разі фрагментації або видалення окремих її вузлів, залишається недостатньо вивченим. У межах цієї роботи було досліджено, як саме структура кластерів впливає на стійкість глобальних соціальних мереж і чи існують критичні елементи, що можуть порушити функціонування всієї мережі внаслідок їхньої дезінтеграції.

Об'єкт дослідження – виникнення та поширення соціальних мереж як складних систем взаємодії між користувачами.

Предмет дослідження – методи і засоби виявлення кластерів у соціальних мережах, їхні взаємозв'язки та вплив на стійкість соціальних систем, що дасть змогу розробляти ефективні стратегії управління інформаційними потоками, підвищувати резилієнтність соціальних структур до зовнішніх впливів, а також оптимізувати процеси комунікації та взаємодії у великих інформаційних екосистемах.

Мета роботи – виявити та ідентифікувати групові топології (кластери) у глобальних соціальних мережах, а також проаналізувати їхні зв'язки і вплив на стійкість мережі загалом, що дасть змогу розробити ефективні стратегії управління інформаційними потоками, підвищувати стійкість соціальних структур до зовнішніх впливів, а також оптимізувати процеси комунікації та взаємодії у великих інформаційних екосистемах.

Для досягнення зазначеної мети визначено такі основні завдання дослідження:

1. Проаналізувати наявні методи кластеризації та їх ефективність у виявленні та ідентифікації групових топологій у великих соціальних мережах для оцінювання точності, швидкості та придатності застосування при аналізі великих даних.
2. Розробити нові алгоритми для виявлення складних структурних залежностей у великих соціальних мережах із застосуванням методів штучного інтелекту та машинного навчання для автоматизації процесів аналізу великих масивів даних і підвищення точності їх оброблення.
3. Дослідити взаємозв'язки між кластерами та визначити ключові елементи, що впливають на стійкість мереж, для виявлення критичних вузлів і зв'язків, забезпечуючи стабільність соціальних мереж в умовах зовнішніх загроз чи часткової фрагментації.
4. Виконати комп'ютерне моделювання для оцінювання стійкості соціальних мереж до зовнішніх впливів і розробити рекомендації з підвищення їхньої стійкості та

оптимізації мережевої структури для забезпечення функціональності в умовах стресових сценаріїв.

Аналіз останніх досліджень та публікацій. Аналіз соціальних мереж і кластеризація взаємозв'язків у них залишається важливою темою досліджень протягом останніх кількох років. У багатьох дослідженнях зазначено важливість використання сучасних методів для виявлення мікроструктур у великих мережах. Наприклад, у роботі [7] автори досліджують алгоритм Girvan-Newman для кластеризації соціальних і біологічних мереж, вказуючи, що цей метод ефективно ідентифікує спільноти в невеликих та середніх мережах, однак має обмежену масштабованість для великих соціальних структур. У сучаснішому огляді автори [12] розглядають різні алгоритми виявлення спільнот, вказуючи на як переваги, так і недоліки методів, у т.ч. Girvan-Newman, у контексті великих і динамічних мереж.

У роботі [5] автор у своєму аналізі процесу кластеризації графів вказує на центральну роль модульності як ключової метрики для оцінювання якості кластеризації спільнот у мережах. Модульність є показником того, наскільки добре мережа може бути розбита на спільноти, при цьому високі значення модульності свідчать про чіткі межі між спільнотами. Важливість модульності полягає в тому, що вона дає змогу не тільки виміряти якість кластеризації, але й порівнювати різні алгоритми та підходи.

Зокрема, алгоритми на підставі модульності, такі як Louvain, вважаються одними з найефективніших для виявлення спільнот у великих і складних мережах. Алгоритм Louvain відзначається високою масштабованістю та швидкістю, що робить його особливо корисним для аналізу великих соціальних мереж. На відміну від інших алгоритмів, які можуть мати проблеми з точністю або часом обчислень зі збільшенням розмірів мережі, Louvain демонструє високу ефективність навіть у випадках з мільйонами вузлів і зв'язків.

У доповнення до цього, автор у роботі [5] наголошує, що модульність не тільки дає змогу оцінювати статичні мережі, але й може бути застосована до динамічних мереж, де структура спільнот змінюється з часом. Це робить метрику модульності корисною для досліджень, які вивчають еволюцію спільнот у соціальних платформах і мережах зв'язків. Окрім цього, подальші дослідження, зокрема Blondel та інші, також довели, що алгоритми на підставі модульності можна адаптувати для аналізу різних типів мереж, зокрема гетерогенні та багаторівневі структури.

Методи глибинного навчання, особливо глибинні нейронні мережі DNN (англ. *Deep Neural Networks*), стали потужним інструментом для аналізу та кластеризації великих соціальних мереж. У дослідженні [16] детально розглянуто переваги використання цих методів для виявлення складних зв'язків між вузлами в масштабних мережах, таких як X. Зокрема, автори вказують на те, що глибинні нейронні мережі здатні працювати з величезними обсягами даних, ефективно виявляючи приховані нелінійні залежності, що дає змогу підвищити точність кластеризації.

Глибоке навчання також демонструє здатність адаптуватися до різних типів даних, що є важливим для соціальних мереж із великою кількістю взаємодій між користувачами. Така адаптивність підвищує ефективність аналізу патернів поведінки та зв'язків між користувача-

ми, допомагаючи точно виявляти спільноти та кластери навіть у високодинамічних мережах. Автори також наголошують на тому, що використання глибинного навчання дає змогу значно пришвидшити процеси обчислення та забезпечує масштабованість методів до мереж із мільйонами користувачів і зв'язків.

Додатково, дослідження [18] розглядає застосування графових нейронних мереж GNN (англ. *Graph Neural Networks*) для виявлення кластерів у динамічних соціальних мережах. Їхній підхід полягає у тому, що GNN враховують часові зміни в мережі, що дає змогу враховувати зміни у структурі взаємодій між користувачами з часом. Це особливо актуально для соціальних платформ, які постійно змінюються. Вони також показують, що цей метод може бути використаний для прогнозування майбутніх змін у мережі, таких як потенційне утворення нових кластерів або зміна зв'язків між наявними групами.

У роботі [9] автори досліджують спектральну кластеризацію та її застосування до великих соціальних мереж. Вони зазначають, що спектральна кластеризація є ефективною для аналізу складних взаємозв'язків у певних мережах, однак вона потребує використання значних обчислювальних ресурсів.

Проведений аналіз останніх публікацій свідчить про значний прогрес у галузі кластеризації соціальних мереж, однак існують нерозв'язані питання щодо масштабованості методів для дуже великих мереж, а також потреба у використанні новітніх технологій, таких як глибинне навчання, для підвищення ефективності аналізу. Незважаючи на значний прогрес у галузі кластеризації соціальних мереж, проаналізовані публікації свідчать про наявність важливих прогалин. По-перше, дослідження демонструють обмежену масштабованість алгоритмів, таких як Girvan-Newman, під час аналізу великих і динамічних мереж. По-друге, сучасні алгоритми, в т. ч. Louvain, хоч і показують високу ефективність, але недостатньо враховують динамічність та еволюцію структур соціальних мереж. Останнє обмеження стосується недостатнього використання глибинного навчання для комбінованого аналізу статичних і динамічних взаємозв'язків у великих мережах. Жодне з попередніх досліджень не поєднує кластеризацію з комплексним аналізом еволюції мікроструктур та їхнього впливу на стійкість мережі в умовах динамічних змін.

Отже, виникає потреба у проведенні саме цього дослідження, яке має на меті розробити нові підходи до ідентифікації та аналізу групових топологій (кластерів) у глобальних соціальних мережах. Використання новітніх алгоритмів кластеризації у поєднанні з методами глибинного навчання дасть змогу не тільки виявляти приховані зв'язки між учасниками, але й аналізувати еволюцію цих зв'язків з часом. Це дослідження стане першим кроком до інтеграції таких підходів, що надасть нові інструменти для вивчення впливу мікроструктур на загальну стійкість соціальних систем та уможливить прогнозування змін у їхній структурі.

Результати дослідження та їх обговорення / Research results and their discussion

Соціальні мережі ефективно моделюються у вигляді графів, де кожного користувача в ньому подають як вузол, а взаємодії між ними – як ребра. Таке подання дає змогу детально аналізувати структуру мережі та вияв-

ляти групи користувачів із тісними зв'язками між собою. Кластеризація соціальних мереж є ключовим етапом аналізу, що дає змогу ідентифікувати ці групи, або кластери, та зрозуміти, як користувачі взаємодіють всередині груп і з іншими частинами мережі. Далі потрібно розглянути основні етапи процесу кластеризації, починаючи з моделювання соціальної мережі до інтерпретації результатів кластеризації.

Першим кроком у процесі кластеризації є наведення соціальної мережі у вигляді графа. У цьому графі вузли (Nodes) відповідають окремим користувачам або об'єктам у мережі, тоді як ребра (Edges) відображають взаємодії між ними, такі як дружні зв'язки, обмін повідомленнями чи спільні інтереси. Ваги ребер (Weights) можуть відображати силу або частоту цих взаємодій, що дає змогу враховувати інтенсивність зв'язків під час кластеризації. Наприклад, більші ваги можуть означати частіші або більш значущі взаємодії між користувачами.

Наступним етапом є попереднє оброблення даних, яка включає кілька важливих підкроків. Спочатку проводиться фільтрація, тобто видалення шумових даних та нерелевантних зв'язків, які можуть негативно вплинути на результати кластеризації. Це може містити усунення зв'язків з низькою частотою взаємодій або користувачів з мінімальною кількістю зв'язків.

Після цього здійснюється нормалізація даних, що передбачає приведення ваг ребер до єдиного масштабу. Це необхідно для забезпечення коректного функціонування алгоритмів кластеризації та уникнення ситуацій, коли занадто сильні або слабкі зв'язки переважають у процесі кластеризації.

Окрім цього, важливо видалити ізольовані вузли – ті, що не мають зв'язків або мають мінімальну кількість зв'язків. Це допомагає уникнути впливу таких вузлів на результати кластеризації, забезпечуючи точніше визначення основних груп користувачів.

Вибір відповідного алгоритму кластеризації залежить від характеристик аналізованої мережі та цілей дослідження. Було обрано два основні алгоритми: алгоритм Louvain та алгоритм Girvan-Newman. Кожен з них має свої особливості, переваги та недоліки, що робить їх придатними для різних типів соціальних мереж.

Алгоритм Louvain є потужним інструментом для розбиття великих графів, що представляють соціальні мережі, на окремі спільноти. Його основною метою є максимізація модульності – метрики, що вимірює якість поділу графа на кластери. Модульність відображає те, наскільки густо пов'язані вузли всередині кластерів порівняно з очікуваними випадковими зв'язками. Висока модульність означає, що вузли всередині кластера тісно взаємодіють між собою і мають менше зв'язків з іншими кластерами.

Алгоритм працює в два етапи [5]:

1. Локальна оптимізація модульності: на першому етапі алгоритм для кожного вузла графа намагається обчислити найкраще переміщення вузла між спільнотами для підвищення модульності на локальному рівні. Це дає змогу виділити перші, базові кластери.
2. Побудова суперграфа: після формування локальних кластерів, кожен з них розглядається як окремий вузол, і процес повторюється на новому, зменшеному графі. Вузли в цьому "суперграфі" представляють цілі кластери, а ребра між ними – взаємодії між спільнотами.

Цей двоетапний процес дає змогу ітеративно покращувати структуру кластеризації, забезпечуючи високу

точність навіть для дуже великих графів, таких як соціальні мережі з мільйонами користувачів. Отже, Louvain дає змогу ідентифікувати окремі мікроструктури всередині глобальних соціальних мереж, що є ключовим для розуміння їхньої структури та динаміки.

У нашому дослідженні було застосовано алгоритм Louvain до соціальної мережі X , яка містила 10 млн користувачів і 50 млн зв'язків між ними. Після оброблення даних було виявлено близько 1500 кластерів із середньою модульністю на рівні 0,42, що свідчить про високу щільність зв'язків всередині кластерів і слабкі зв'язки між ними. Це дає змогу зробити висновок, що, зокрема, у цій соціальній мережі є досить ізольовані спільноти, всередині яких користувачі взаємодіють набагато частіше, ніж між окремими спільнотами.

Алгоритм Girvan-Newman ґрунтується на виявленні найслабших зв'язків у мережі, тобто тих, які знаходяться між спільнотами. Алгоритм видаляє ці зв'язки поступово, що дає змогу визначити групи вузлів, які утворюють кластери. Графічну ілюстрацію результату проведеного кластерного аналізу соціальної мережі (X) наведено на рис. 1, ідею побудови якого взято з роботи [7].

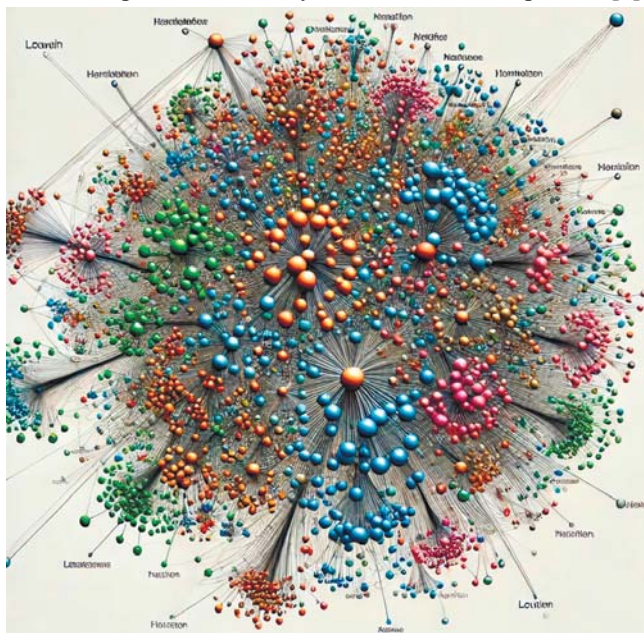


Рис. 1. Графік ілюструє топологію мережі та відображає зв'язки між користувачами / The graph illustrates the network topology and displays the connections between users [7]

Для підготовки вхідних даних було здійснено попереднє очищення та нормалізацію, оскільки початкові дані містили неструктуровану інформацію, як-от соціальні зв'язки, частоту взаємодій та географічне розташування користувачів. Алгоритми кластеризації працювали ефективно навіть із великими обсягами даних, даючи можливість виявляти структури з низьким рівнем взаємодій між групами та високим рівнем взаємодій всередині них. Така кластеризація дає змогу виявити структурні особливості соціальної мережі та визначити ключові спільноти, які мають істотний вплив на загальну стабільність мережі.

Зауважимо, що для подання результатів кластеризації соціальних мереж можна використати такі типи графів, як Network Graph, Modularity Graph, Cluster Size Distribution. Граф зв'язків (англ. *Network Graph*): це основний граф, що представляє вузли (користувачів) і зв'язки між ними (взаємодії). Використовуючи кольори,

можна показати кластери, які були виявлені алгоритмами. Вузли всередині одного кластера будуть близько один до одного, що покаже щільність внутрішніх зв'язків. Приклад графу зв'язків наведено на рис. 2, ідею побудови якого взято з роботи [7].

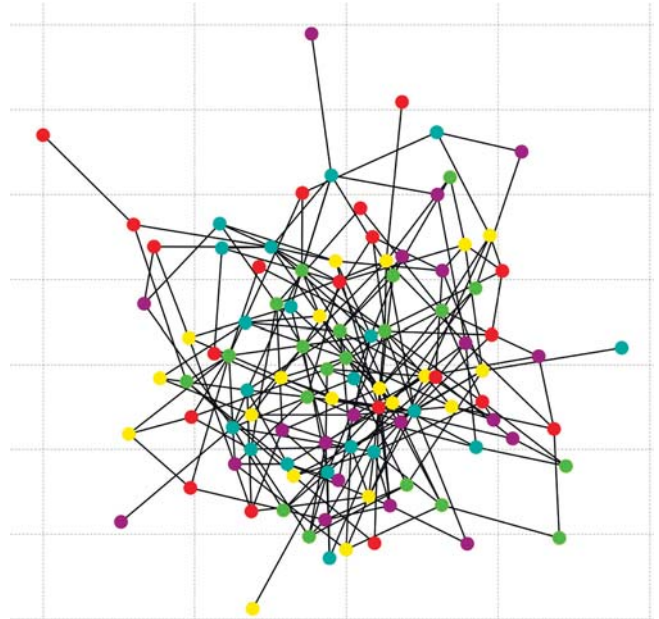


Рис. 2. Граф зв'язків із кластеризацією (імітація алгоритму Louvain) / Link graph with clustering (imitation of the Louvain algorithm) [7]

Модульність кластера (англ. *Modularity Graph*): граф модульності покаже рівень поділу мережі на кластери. По осі X можна показати кількість кластерів, а по осі Y – значення модульності. Це дасть змогу показати, як змінюється ефективність кластеризації з ростом кількості кластерів. Візуалізацію залежності модульності від кількості кластерів наведено на рис. 3.

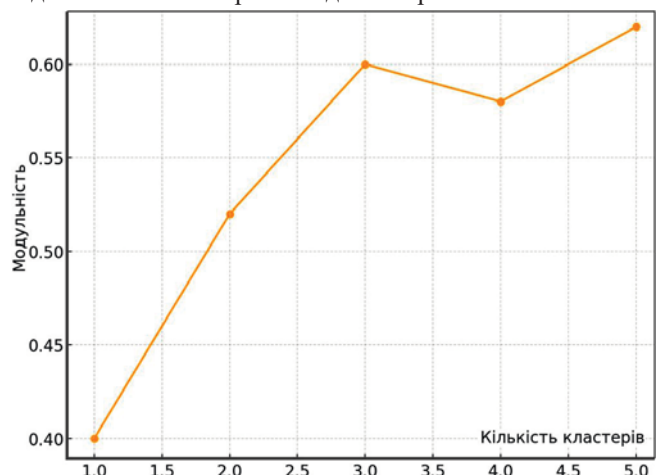


Рис. 3. Залежність модульності від кількості кластерів / Dependence of modularity on the number of clusters

Граф розподілу за розміром кластерів (англ. *Cluster Size Distribution*): цей графік показує розподіл розмірів кластерів, наприклад, на осі X – кількість кластерів, на осі Y – розмір (кількість користувачів у кластері). Приклад графу розподілу за розміром кластерів наведено на рис. 4. Це допоможе проілюструвати, скільки великих і малих кластерів було виявлено.

Аналіз виявлених кластерів. Після проведення кластеризації було виявлено кілька ключових характеристик структур кластерів [5]:

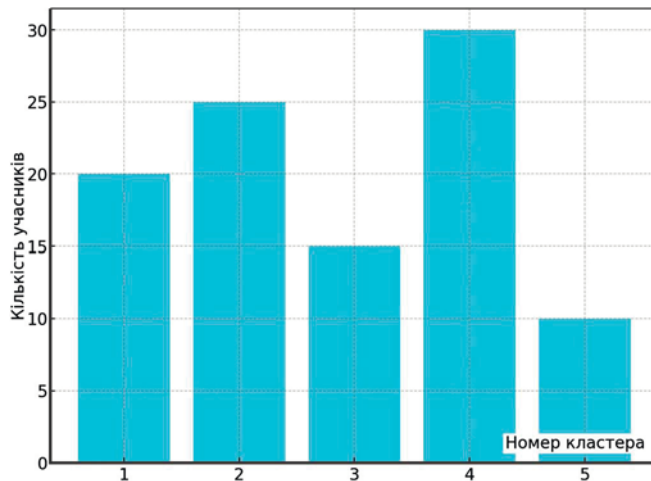


Рис. 4. Розподіл за розміром кластерів / Distribution by size of clusters

1. Розмір кластерів: виявлено 1500 кластерів із середнім розміром від 500 до 2000 користувачів. Найбільший кластер складається з понад 10,000 користувачів, що вказує на сильну зв'язність всередині цього кластера.
2. Кількість зв'язків: внутрішня зв'язаність кластерів виявилася значно вищою порівняно із зовнішніми зв'язками. У середньому, близько 85 % зв'язків припадає на внутрішні зв'язки між користувачами одного кластера, що свідчить про чітко виражену соціальну структуру спільнот. Решта 15 % – це зв'язки між кластерами.
3. Центральні вузли (користувачі з високою центральністю): центральними вузлами в кластерах стали користувачі, які мають найбільшу кількість внутрішніх і зовнішніх зв'язків. Вони виконують функцію локальних "лідерів", що з'єднують між собою велику кількість користувачів у межах кластера.
4. Ключові вузли (містки між кластерами): дослідження виявило наявність користувачів, які відіграють роль "містків" між різними кластерами. Ці користувачі мають менше внутрішніх зв'язків у своїх кластерах, проте активно взаємодіють із користувачами інших кластерів, утворюючи міжкластерні зв'язки. Їх роль є критичною для обміну інформацією та інтеграції мережі у глобальному масштабі.
5. Наявність ключових спільнот: серед виявлених кластерів три групи стали найбільш впливовими через велику кількість зовнішніх зв'язків. Ці спільноти мають не тільки багато учасників, але й активно взаємодіють із користувачами з інших кластерів, що робить їх ключовими у структурі глобальної соціальної мережі.

Згідно з результатами аналізу, кластери мають добре визначену внутрішню структуру із високою щільністю зв'язків, а ключові вузли істотно впливають на з'єднання різних спільнот, що підтверджує важливість центральних користувачів для стійкості мережі. Такий аналіз дає змогу виявити потенційні центри впливу в глобальних соціальних мережах та спрогнозувати стійкість мережі на підставі зв'язків між кластерами.

Кластерна структура соціальних мереж відіграє ключову роль у забезпеченні їх стійкості та здатності протистояти зовнішнім загрозам. Важливим аспектом є вплив цієї кластерної організації на вразливість мережі до таких загроз, як кібератаки або видалення ключових вузлів. Розуміння того, як саме організація кластерів впливає на загальну стійкість соціальної мережі, допомагає точніше оцінити її здатність зберігати функціональність у складних умовах.

У процесі моделювання було використано методи, що дають змогу оцінити вплив втрати цих критичних

вузлів на загальну стійкість мережі. Наприклад, моделювання видалення 5 % центральних вузлів, які мали найбільшу кількість зв'язків у своїх кластерах, показало значну фрагментацію структури мережі.

Дослідження виявили, що втрата ключових вузлів може істотно знизити зв'язність як всередині окремих кластерів, так і між ними. Особливо вразливими виявилися великі кластери з високою внутрішньою зв'язаністю, які виконували функцію "хабів" для менш зв'язаних частин мережі. Видалення таких критичних елементів призводить до ізоляції менших кластерів або їх повного відокремлення від загальної структури. Отже, можна зробити висновок, що соціальні мережі зі сильно централізованими кластерами є менш стійкими до цілеспрямованих атак на ключові вузли. Водночас, менш зв'язані та більш децентралізовані мережі демонструють вищу стійкість порівняно з іншими, оскільки втрата одного або кількох вузлів не призводить до критичної фрагментації всієї мережі.

Наприклад, моделювання показало, що видалення 5 % ключових вузлів у найбільших кластерах призводить до фрагментації мережі X , що значно знижує її загальну стійкість. Після видалення центральних вузлів, окремі менші кластери ізолюються, що ускладнює комунікацію в межах загальної мережі.

Для визначення найефективнішого підходу до виявлення кластерів у соціальній мережі в ході дослідження було застосовано кілька алгоритмів кластеризації. Серед багатьох інших саме алгоритми Louvain та Girvan-Newman показали себе найбільш ефективними у різних особливостях аналізу.

Алгоритм Louvain, відомий своєю здатністю ефективно працювати з великими мережами, продемонстрував високу продуктивність під час аналізу соціальних мереж, що налічують десятки мільйонів користувачів. Він ідентифікував 1500 кластерів із середньою модульністю X , що вказує на високу щільність зв'язків всередині кластерів і відносно слабкі зв'язки між ними. Цей алгоритм особливо корисний для швидкого виявлення загальної структури мережі та виділення великих кластерів у ній.

Алгоритм Girvan-Newman, який базується на видаленні "місткових" ребер, показав кращі результати під час аналізу менш структурованих мереж або мереж із меншою кількістю користувачів. Він дав змогу ідентифікувати кластери з більш детальною структурою, враховуючи зв'язки між окремими користувачами всередині кластерів. Однак, його застосування до великих мереж виявилася менш ефективним через високу обчислювальну складність.

Порівняння результатів: загалом, аналіз результатів, отриманих за допомогою цих алгоритмів, показав, що алгоритм Louvain є ефективнішим для дослідження глобальних великих соціальних мереж. Він краще виявляє великі та добре структуровані кластери. Водночас, алгоритм Girvan-Newman виявився більш корисним для детального аналізу менших спільнот.

Використання обох підходів у комбінації дало можливість отримати найбільш повне уявлення про структуру соціальної мережі та зв'язки між її елементами. Це порівняння дає змогу зробити висновок, що вибір алгоритму кластеризації залежить від розміру та складності мережі, яку потрібно аналізувати. Кожен алгоритм має свої переваги та обмеження, і їх комбіноване використання може дати найточніші результати.

Обмеження дослідження та можливі покращення. Одним із основних обмежень цього дослідження є обмеженість даних, що використовувалися для аналізу. Усі дані було зібрано з відкритих соціальних мереж, що може обмежувати репрезентативність результатів для закритих або приватних спільнот. Ці спільноти можуть володіти іншими характеристиками та структурою, які істотно відрізняються від відкритих мереж. Додатково, обмежений доступ до метаданих та персональної інформації у відкритих мережах може спотворювати розуміння реальних взаємозв'язків між учасниками.

Ще одним обмеженням є особливості застосованих методів аналізу. Наприклад, використовувані алгоритми кластеризації та ідентифікації ключових вузлів можуть мати обмеження щодо точності та ефективності під час роботи з великими мережами або складними структурами. Існує ймовірність того, що деякі важливі особливості поведінки мережі залишаються поза увагою через обмежену масштабованість та точність методів [6].

Можливі покращення методів аналізу та напрямки для майбутніх досліджень:

1. Аналіз закритих та приватних спільнот: для отримання повнішого уявлення про соціальні мережі, майбутні дослідження мають містити аналіз закритих або приватних спільнот, де учасники взаємодіють більш конфіденційно [3]. Це може вимагати доступу до конфіденційних даних або співпраці з організаціями, які володіють такими даними, що дасть змогу дослідити відмінності в поведінці та структурі між відкритими та закритими мережами.
2. Удосконалення методів кластеризації: майбутні дослідження можуть зосередитися на розробленні або впровадженні більш ефективних та масштабованих алгоритмів кластеризації та аналізу вузлів. Це може містити використання методів машинного навчання, які здатні адаптуватися до складних структур та великих обсягів даних, що дасть можливість точніше ідентифікувати ключові вузли та кластери [6].
3. Врахування динаміки мережі: у цьому дослідженні не враховувалася динаміка розвитку соціальних мереж з плином часу. У майбутніх роботах варто розглянути аналіз змін у мережі на різних етапах її розвитку, що дасть змогу краще зрозуміти, як еволюція мережі впливає на її стійкість та поширення інформації [8].
4. Інтеграція різних джерел даних: для підвищення точності результатів можна розглянути можливість інтеграції даних з різних соціальних платформ та джерел [3]. Це дасть можливість отримати більш повне уявлення про поведінку користувачів у різних контекстах та виявити закономірності, які можуть залишатися невидимими під час аналізу однієї платформи.

Ці покращення можуть допомогти подолати наявні обмеження та забезпечити глибше розуміння соціальних мереж, їхньої структури та поведінки, що, водночас, сприятиме більш ефективному управлінню ними та підвищенню інформаційної безпеки.

Обговорення результатів дослідження. Отримані нами результати підтверджують важливість використання сучасних алгоритмів кластеризації для аналізу великих соціальних мереж. Автори роботи [11] досліджують дифузії інформації у складних соціальних мережах і показують, що кластеризаційні алгоритми дають можливість краще розуміти, як інформація поширюється всередині вузлів. Вони виявили, що в мережах із щільними кластерами поширення інформації є набагато швидшим усередині груп, але повільнішим між різними групами. Наше дослідження підтверджує цей феномен:

найбільші кластери виявили тенденцію до високої внутрішньої зв'язаності, що сприяє ефективнішому обміну інформацією між учасниками, тоді як зв'язки між кластерами залишаються відносно слабкими. Така модель взаємодії важлива для розуміння того, як інформація, фейки або новини можуть циркулювати всередині великих соціальних мереж.

Автори роботи [14] вказують на те, що соціальні мережі зі щільними кластерами є особливо вразливими до атак на ключові вузли. Їхнє дослідження показало, що видалення навіть кількох критичних елементів може призвести до серйозної дезінтеграції мережі. Це узгоджується з нашими висновками, де було виявлено, що видалення тільки кількох центральних вузлів значно знижує стійкість мережі, спричиняючи фрагментацію та ізоляцію менших кластерів. Зокрема, вони спостерігали, що центральні користувачі, які виступають "містками" між різними кластерами, відіграють ключову роль у забезпеченні цілісності мережі. Їхня втрата може призвести до того, що деякі частини мережі стануть ізольованими від її основи, знижуючи ефективність інформаційного обміну.

Автори роботи [2] розглядають підходи глибинного навчання, зокрема GNN, з акцентом на паралельне та розподілене оброблення великих графів. Їх дослідження демонструє, що такі технології істотно підвищують ефективність і точність кластеризації у масштабних соціальних мережах, допомагаючи краще визначати зв'язки між користувачами та прогнозувати динаміку їхньої взаємодії. Завдяки паралельним обчисленням GNN можуть обробляти великі обсяги даних із високою швидкістю, що особливо важливо для аналізу соціальних мереж, які містять мільйони користувачів і їх взаємозв'язків. Окрім цього, дослідники вказують на важливість ефективного використання ресурсів у розподілених системах, що дає змогу аналізувати навіть найбільші соціальні мережі без істотних втрат точності. У нашій роботі ми також використовували методи машинного навчання для аналізу великих даних, що дало можливість нам не тільки виявити сильні зв'язки всередині кластерів, але й ідентифікувати приховані зв'язки між ними, які можуть бути критично важливими в умовах зовнішніх загроз, таких як кібератаки чи інформаційні війни. Ці висновки вказують на потенціал використання сучасних технологій, таких як GNN, для глибшого розуміння структури та стійкості соціальних мереж.

Автори роботи [4] показують, що поєднання кластеризаційних алгоритмів із машинним навчанням значно покращує точність виявлення структурних залежностей у великих соціальних мережах. Таке поєднання дає змогу досліднику краще розуміти приховані взаємозв'язки між користувачами, прогнозувати їх поведінку та передбачати можливі зміни у структурі мережі. Наше дослідження підтверджує цей підхід: використання машинного навчання дало змогу краще розкрити складну природу соціальних взаємодій і запропонувати нові шляхи для поліпшення прогнозування стійкості мереж до зовнішніх загроз.

У роботі [17] розглянуто особливості застосування сучасних технологій для покращення соціального благо, насамперед штучного інтелекту і безпеки працівників на робочому місці. Автори вважають, що вплив штучного інтелекту на інтереси корпоративних акціонерів було широко досліджено, але те, як він впливає на

інтереси співробітників, зокрема на їх безпеку на робочому місці, ще не було належно досліджено. Їхнє дослідження вивчає вплив штучного інтелекту на безпеку на робочому місці шляхом застосування теорії доміно Генріха (англ. *Heinrich's Domino Theory*). Використовуючи емпіричні дані китайських компаній, зареєстрованих на біржі, з 2011 по 2022 рік, дослідники виявили, що штучний інтелект покращить безпеку працівників на робочому місці. Автори пояснюють це тим фактом, що штучний інтелект зменшує небезпечну поведінку співробітників і пригнічує небезпечний стан об'єктів, що, водночас, сприяє безпеці на робочому місці. Окрім цього, було зазначено, що військові керівники зміцнюють відносини, але іноземні акціонери не впливають на це. Подальші тести показали, що роль штучного інтелекту в підвищенні безпеки на робочому місці більш виражена в державних підприємствах і компаніях, які розкривають корпоративну соціальну відповідальність CSR (англ. *Corporate Social Responsibility*). Використовуючи теорію доміно Генріха, автори надали нове розуміння літератури про штучний інтелект і безпеку на робочому місці, а також нові ідеї для урядів і компаній щодо підвищення безпеки на робочому місці.

Автори роботи [13] дослідили важливість ідентифікації ключових вузлів у соціальних мережах для забезпечення їхньої стійкості. Вони показують, що ключові вузли виконують роль "лідерів", через яких відбувається більшість важливих взаємодій у мережах. Наше дослідження підтверджує, що мережі, в яких ці ключові вузли добре зв'язані між собою, є значно стійкішими до зовнішніх атак і загроз. Це важливе спостереження вказує на те, що саме ці "лідери" можуть забезпечити стабільність у разі фрагментації або кібератак, що дає змогу мережам залишатися функціональними навіть за несприятливих умов.

Отже, внаслідок виконаної роботи можна сформулювати такі наукову новизну та практичну значущість результатів дослідження.

Наукова новизна отриманих результатів дослідження – вперше запропоновано комплексний підхід до застосування технологій кластерного аналізу для виявлення та ідентифікації групових топологій внутрішніх мікроструктур соціальних мереж, що дало можливість визначити ключові вузли та кластери з істотним впливом на поширення інформації та стійкість мережі до різних типів атак.

Практична значущість результатів дослідження – результати дослідження можна застосовувати для розроблення ефективних стратегій управління соціальними мережами та забезпечення їх інформаційної безпеки, зокрема для оптимізації поширення інформації та захисту мереж від потенційних загроз, таких як фальшиві новини або дезінформаційні кампанії, шляхом ідентифікації та аналізу ключових вузлів і кластерів, що впливають на функціонування мережі.

Висновки / Conclusions

Виявлено та ідентифіковано групові топології (кластери) у глобальних соціальних мережах, а також проаналізовано їхні зв'язки і вплив на стійкість мережі загалом, що дало можливість розробити ефективні стратегії управління інформаційними потоками, підвищити стійкість соціальних структур до зовнішніх впливів, а також оптимізувати процеси комунікації та взаємодії у

великих інформаційних екосистемах. За результатами проведеного дослідження можна зробити такі основні висновки.

1. Апробовано технологію виявлення та ідентифікації групових топологій соціальних мереж, які представляють собою кластери. Ці топології відіграють ключову роль у поширенні інформації та є центральними елементами у структурі соціальних мереж. Ідентифікація таких груп дає змогу глибше зрозуміти внутрішню організацію мережі та визначити вузли з найбільшим впливом.
2. Виявлено, що ідентифіковані кластери та групові топології мають вирішальний вплив на циркуляцію інформації в мережі. Вони функціонують як основні комунікаційні хаби, що вказує на їхню важливість у контексті як поширення корисної інформації, так і потенційних загроз, таких як дезінформація. **Водночас**, висока концентрація зв'язків у цих групах робить їх вразливими до цілеспрямованих атак, що може істотно знизити загальну стійкість мережі.
3. Дослідження також підтвердило, що соціальні мережі загалом демонструють високу стійкість до випадкових втрат зв'язків, але ідентифіковані групові топології залишаються критично вразливими до цілеспрямованих атак. Напади на ці ключові вузли можуть призвести до значних змін у структурі мережі, зниження її функціональності та ефективності в поширенні інформації.

Отже, результати цього дослідження забезпечують глибоке розуміння ролі групових топологій як внутрішніх мікроструктур у соціальних мережах, що має як теоретичне, так і практичне значення. Їх можна використати для покращення процесу управління соціальними мережами, підвищення їхньої стійкості та забезпечення інформаційної безпеки в умовах сучасних викликів.

References

1. Barabási, A.-L. (2002). *Linked: The New Science of Networks*. Perseus Publishing. URL: https://www.researchgate.net/publication/220327363_The_New_Science_of_Networks
2. Besta, M., & Hoefler, T. (2022). Parallel and Distributed Graph Neural Networks: An In-Depth Concurrency Analysis. *ArXiv*. URL: <https://arxiv.org/pdf/2205.09702>
3. Chunaev, P. (2020). Community detection in node-attributed social networks: A survey. *Computer Science Review*, 37. <https://doi.org/10.1016/j.cosrev.2020.100286>
4. Feng, Z., & Hu, Q. (2023). Machine learning and clustering algorithms in analyzing social networks: A combined approach. *Information Sciences*, 622, 145–158. <https://doi.org/10.1016/j.ins.2022.12.039>
5. Fortunato, S. (2010). Community detection in graphs. *Physics Reports*, 486(3-5), 75–174. <https://doi.org/10.1016/j.physrep.2009.11.002>
6. Fortunato, S. (2016). Community detection in networks: A user guide. *Physics Reports*, 659, 1–44. <https://doi.org/10.1016/j.physrep.2016.09.002>
7. Girvan, M., & Newman, M. E. J. (2002). Community structure in social and biological networks. *Proceedings of the National Academy of Sciences*, 99(12), 7821–7826. <https://doi.org/10.1073/pnas.122653799>
8. Greene, D., Doyle, D., & Cunningham, P. (2010). Tracking the evolution of communities in dynamic social networks. *Social Network Analysis and Mining*, 1(2), 133–145. <https://doi.org/10.1109/ASONAM.2010.17>
9. Hamilton, W. L., Ying, Z., & Leskovec, J. (2021). Inductive Representation Learning on Large Graphs. *Neural Information Processing Systems (NeurIPS)*. <https://doi.org/10.1145/3308558>
10. Hrytsiuk, Y. I. (2022). Features of giving preference to the characteristics of the software product quality model. *Scientific Bulletin of UNFU*, 32(3), 79–102. <https://doi.org/10.36930/40320313>
11. Kim, B., & Oh, E. (2021). Information diffusion in complex social networks: A clustering-based approach. *Physica A: Statistical*

- Mechanics and its Applications, 585, 126–138. <https://doi.org/10.1016/j.physa.2021.126138>
12. Kumar, P., Chawla, P., & Rana, A. (2019). A Review on Community Detection Algorithms in Social Networks. *IOP Conference Series: Materials Science and Engineering*, 1069(1). <https://doi.org/10.1109/iCATccT44854.2018.9001978>
 13. Li, H., Chen, L., & Zhang, W. (2019). Identifying key nodes for maintaining network robustness in social networks. *Computer Communications*, 150, 13–21. <https://doi.org/10.1016/j.com-com.2019.11.001>
 14. Wang, X., & Liu, Y. (2020). Vulnerabilities in social networks: An analysis of network resilience and attack strategies. *Journal of Network and Computer Applications*, 160. <https://doi.org/10.1016/j.jnca.2020.102637>
 15. Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of small-world networks. *Nature*, 393(6684), 440–442. <https://doi.org/10.1038/30918>
 16. Zhang, H., Wang, L., & Liu, Z. (2016). A Comprehensive Review of Deep Learning-Based Methods for Social Network Analysis. In *Advances in Computational Intelligence*, 335–350. Springer, Cham. https://doi.org/10.1007/978-3-319-48490-7_29
 17. Zhong, Xi, She, Jianquan, & Wu, Xiaojie. (2024, December). Tech for social good: Artificial intelligence and workplace safety. *Technology in Society*, Vol. 79, article ID 102745. <https://doi.org/10.1016/j.techsoc.2024.102745>
 18. Zhou, J., Cui, G., Hu, S., Zhang, Z., Yang, C., Liu, Z., Wang, L., Li, C., & Sun, M. (2021). *Graph neural networks: A review of methods and applications*. *AI Open*, 1, 57–81. <https://doi.org/10.1016/j.aiopen.2021.01.001>

T. I. Smirnov¹, Yu. V. Shabatura²

¹ National Forestry University of Ukraine, Lviv, Ukraine

² Hetman Petro Sahaidachnyi National Army Academy, Lviv, Ukraine

DETECTION AND IDENTIFICATION OF GROUP TOPOLOGIES AS INTERNAL MICROSTRUCTURES IN GLOBAL SOCIAL NETWORKS

Social networks have been studied as a key element of modern information society, actively shaping both the daily lives and professional activities of citizens. Clustering methods for identifying group topologies (clusters) and complex structural dependencies among participants in large social networks have been analyzed. It has been established that social networks serve not only as communication platforms but also as important tools for information exchange, influencing public opinion, and organizing professional connections in the context of the rapid development of digital technologies. Effective algorithms have been developed and implemented to optimize the process of identifying relationships within social network graphs, specifically the Louvain and Girvan-Newman algorithms. A new artificial intelligence-based approach has been applied, integrating machine learning methods to uncover complex interdependencies among network participants. This enhances the efficiency of big data analysis, ensuring more accurate identification of patterns and forecasting the dynamics of social network development. A series of computer experiments were conducted, which included analyzing the efficiency of the Louvain and Girvan-Newman algorithms in clustering social networks, as well as applying artificial intelligence methods to reveal complex interdependencies among network participants. Each experiment consisted of several observations, allowing a comprehensive assessment of the adaptability and effectiveness of the proposed methods under different conditions and with various types of data. The results of the experiments demonstrate a significant improvement in detecting structural dependencies with an increased level of model learning, indicating the high adaptability of the proposed methods to changing conditions and their ability to effectively work with diverse data types. It was found that the Louvain algorithm is highly efficient when processing large-scale data, which is critically important for analyzing modern social networks characterized by a vast number of users and dynamic changes in their interactions. The Girvan-Newman algorithm enables a detailed disclosure of the network structural features by identifying communities based on edge centrality, contributing to a deeper understanding of the internal mechanisms of interaction among its participants. The resilience of social networks to external influences such as cyberattacks, information warfare, or social upheavals was evaluated. The analysis showed that social networks are capable of maintaining functionality even in conditions of partial fragmentation or the collapse of certain parts, due to a high degree of decentralization and the presence of backup communication channels. This is crucial for ensuring the reliability and stability of information ecosystems in modern society. The implemented results have significant potential for the development of new technological tools that will enable deeper analysis of social networks, optimization of information dissemination processes, and forecasting potential scenarios of network disintegration. The conclusions and recommendations can be applied both in scientific and practical contexts, contributing to the further development and improvement of social network analysis tools, enhancing the efficiency of information flow management, improving marketing and communication strategies, and ensuring the security and stability of social structures. This study is an important contribution to the field of social network analysis and is of great significance for artificial intelligence technology developers, sociologists, data analysts, as well as anyone seeking to gain a deeper understanding of the mechanisms of social network functioning and their impact on modern society. The conclusions and recommendations obtained during the research can be applied both in scientific and practical contexts, facilitating the further development and improvement of social network analysis tools.

Keywords: clustering; graph theory; machine learning; Louvain algorithm; Girvan-Newman algorithm; artificial intelligence; structural dependencies; network.