

Міністерство освіти і науки України
Харківський національний університет радіоелектроніки

Факультет _____ Комп'ютерних наук _____
(повна назва)
Кафедра _____ Штучного інтелекту _____
(повна назва)

АТЕСТАЦІЙНА РОБОТА
Пояснювальна записка

рівень вищої освіти _____ другий (магістерський) _____
_____ Моделювання асортативних безмасштабних мереж _____
_____ (тема)

Виконав:

студент 2 курсу, групи _____ СШМ-18-1 _____
_____ Лукієнко І.В. _____
(прізвище, ініціали)

Спеціальність _____ 122 – Комп'ютерні науки _____
(код і повна назва спеціальності)

Тип програми _____ освітньо-професійна _____
(освітньо-професійна або освітньо -наукова)

Освітня програма _____ Системи штучного _____
_____ інтелекту (СШІ) _____
(повна назва освітньої програми)

Керівник _____ доц. Шергін В.Л. _____
(посада, прізвище, ініціали)

Допускається до захисту

Зав. кафедри _____ В.О. Філатов _____
(підпис) (прізвище, ініціали)

2019 р.

Харківський національний університет радіоелектроніки

Факультет _____ Комп'ютерних наук
Кафедра _____ Штучного інтелекту
Рівень вищої освіти _____ другий (магістерський)

Спеціальність _____ 122 – Комп'ютерні науки
(код і повна назва)
Тип програми _____ Освітньо-професійна
(освітньо-професійна або освітньо-наукова)

Освітня програма _____ Системи штучного інтелекту (СШІ)
(повна назва)

ЗАТВЕРДЖУЮ:

Зав. кафедри _____
(підпис)

« _____ » 20 ____ р.

ЗАВДАННЯ
НА АТЕСТАЦІЙНУ РОБОТУ

студентові _____ Лукієнку Іллі Володимировичу
(прізвище, ім'я, по батькові)

1. Тема роботи _____ Моделювання асортативних безмасштабних мереж

затверджена наказом по університету від _____ 04.11.2019 р. _____ № 1623Ст

2. Термін подання студентом роботи _____ 19.12.2019 р.

3. Вихідні дані до роботи _____ Науково-технічні публікації, дані Інтернет-джерел щодо розробки існуючих методів моделювання та дослідження асортативності мереж

4. Зміст пояснювальної записки (перелік питань, що потрібно розробити) _____

Вступ, 1 Аналіз предметної області, 1.1 Поняття складних мереж, 1.2 Моделі складних мереж, 1.3 Асортативність мереж, 1.4 Постановка задачі дослідження, 2 Аналіз властивостей і характеристик складних мереж, 2.1 Властивості і характеристики мереж, 2.2 Безмасштабні мережі, 2.3 Властивості мереж Барабаші-Альберт, 2.4 Фрактальні властивості мереж Барабаші-Альберт, 2.5 Оцінювання меж коефіцієнта асортативності мереж Барабаші-Альберт, 3 Експериментальне дослідження меж показника асортативності мереж Барабаші-Альберт, 3.1 Вибір програмних засобів реалізації, 3.2 Генерація мереж Барабаші-Альберт, 3.3 Оцінювання фрактальної розмірності мереж Барабаші-Альберт, 3.4 Експериментальне оцінювання граничних значень показника асортативності мереж Барабаші-Альберт, Висновки, Перелік посилань

5. Перелік графічного матеріалу із зазначенням креслеників, схем, плакатів, комп'ютерних ілюстрацій (слайдів) Слайд 1 – Титульний. Слайд 2 – Об'єкт, предмет, мета дослідження. Слайд 3– Основні задачі. Слайд 4 – Характерні властивості складних мереж. Слайд 5 – Типові розподіли вузлів мереж за кількістю зв'язків. Слайд 6 – Основні моделі мереж. Слайд 7–Модель Барабаші-Альберт. Слайд 8 – Асортативність мереж. Слайд 9 – Показникасортативності. Слайд 10 – Показники асортативності деяких мереж. Слайд 11– Структураекстремальних мереж. Слайд 12 – Постановка задачі пошуку структури екстремально асортативної мережі. Слайд 13 – Алгоритм побудови мереж, екстремальних за асортативністю. Слайд 14 – Структура мереж, екстремальних за асортативністю. Слайд 15– Матрицясуміжності екстремально дизасортативної БА-мережі з $n=32$ вузлів. Слайд 16 – Матриця суміжності екстремально асортативної БА-мережі з $n=32$ вузлів. Слайд 17 – Оцінки меж показника асортативності мереж Барабаші-Альберт. Слайд 18 – Точні та асимптотичні межі коефіцієнту асортативності в залежності від розміру мережі Барабаші-Альберт. Слайд 19 – Висновки

6. Консультанти розділів роботи

Найменування розділу	Консультант (посада, прізвище, ім'я, по батькові)	Позначка консультанта про виконання розділу	
		підпис	дата
Основний розділ	доцент Шергін В.Л.		

КАЛЕНДАРНИЙ ПЛАН

№	Назва етапів роботи (проекту)	Термін виконання етапів проекту (роботи)	Примітка
1	Отримання завдання на атестаційну роботу	04.11.2019	виконано
2	Аналіз предметної області	05.11.2019 – 10.11.2019	виконано
3	Постановка завдання та узгодження з керівником	11.11.2019 – 12.11.2019	виконано
4	Аналіз моделей графів та мереж, числових характеристик	13.11.2019 – 27.11.2019	виконано
5	Дослідження моделі Барабаші-Альберт	18.11.2019 – 22.11.2019	виконано
6	Дослідження асортативності безмасштабних мереж	23.11.2019 – 27.11.2019	виконано
7	Експериментальні дослідження	28.11.2019 – 05.12.2019	виконано
8	Написання пояснювальної записки	06.12.2019 – 10.12.2019	виконано
9	Попередній захист	13.12.2019	виконано
10	Захист перед ЕК	19.12.2019	

Дата видачі завдання _____

Студент _____ Лукієнко І.В.
(підпис)

Керівник роботи (проекту) _____ доц. Шергін В.Л.
(підпис) (посада, прізвище, ініціали)

РЕФЕРАТ

Записка пояснювальна: 74 с., 24 рис., 3 табл., 2 дод., 21 джерело.

АСОРТАТИВНІСТЬ, БЕЗМАСШТАБНА МЕРЕЖА, МЕРЕЖА
БАРАБАШІ-АЛЬБЕРТ, МЕРЕЖА ЕРДЕША-РЕНІ, МЕРЕЖА КАЛЛАВЕЯ,
ПЕРЕНАПРАВЛЕННЯ ЗВ'ЯЗКІВ

Об'єктом досліджень є мережі Барабаші-Альберт.

Предметом досліджень атестаційної роботи магістра є асортативність.

Метою роботи є аналіз моделей складних мереж, дослідження структури мереж та їх впливу на показник асортативності, проведення експериментальних досліджень з моделювання безмасштабних мереж, розрахунку та модифікації асортативності мереж, які моделюються.

РЕФЕРАТ

Записка пояснительная: 74 с., 24 рис., 3 табл., 2 прил., 21 источник.

АССОРТАТИВНОСТЬ, БЕЗМАСШТАБНАЯ СЕТЬ,
ПЕРЕНАПРАВЛЕНИЕ СВЯЗЕЙ, СЕТЬ БАРАБАШИ-АЛЬБЕРТ, СЕТЬ
КАЛЛАВЕЯ, СЕТЬ ЭРДЕША-РЕНЬИ

Объектом исследований являются сети Барабаши-Альберт.

Предметом исследований аттестационной работы магистра является ассортативность.

Целью работы является анализ моделей сложных сетей, исследование структуры сетей и её влияния на показатель ассортативности, проведение экспериментальных исследований по моделированию безмасштабных сетей, расчёту и модификации ассортативности моделируемых сетей.

ABSTRACT

Explanatory note: 74 pages, 24 figures, 3 tables, 21 sources, 2 appendixes.

ASSORTATIVITY, BARABASI-ALBERT NETWORK, CALLAWAY NETWORK, ERDES-RENI NETWORK, REWIRING, SCALE-FREE NETWORK

The object of research is the Barabasi-Albert network.

The subject of research is assortativity.

Целью работы является анализ моделей сложных сетей, исследование структуры сетей и её влияния на показатель ассортативности, проведение экспериментальных исследований по моделированию безмасштабных сетей, расчёту и модификации ассортативности моделируемых сетей.

The aim of the work is analysis models of complex networks, studying the structure of networks and its influence on assortativity, modelling the scale-free networks, evaluation its assortivity indexes, rewiring.

ЗМІСТ

Перелік умовних позначень, одиниць, скорочень та термінів	7
Вступ.....	8
1 Аналіз предметної області	10
1.1 Визначення і напрямки складних мереж	11
1.2 Моделі графів складних мереж	15
1.3 Постановка задачі дослідження.....	22
2 Асортативність мереж	23
2.1 Поняття асортативності мереж.....	23
2.2 Вплив асортативності на властивості мереж	30
2.3 Асортативність у неорієнтованих мережах.....	36
2.4 Процедура перенаправлення зв'язків	44
2.5 Моделювання мереж з переважним приєднанням	46
3 Обчислення показника асортативності в безмасштабних мережах	53
3.1 Вибір програмних засобів реалізації.....	53
3.2 Генерація мереж Барабаші-Альберт	59
3.3 Розрахунок та модифікація асортативності мережі Барабаші-Альберт.	62
Висновки	64
Перелік джерел посилання	65
Додаток А.....	67
Додаток Б	73

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, ОДИНИЦЬ, СКОРОЧЕНЬ ТА ТЕРМІНІВ

БА – мережа Барабаші-Альберт;

MIM – масштабно-інваріантна мережа;

ПП – переважне приєднання (переважне зв'язування);

box-counting – підрахунок числа шаблонів;

ER – Erdos-Renyi network – мережа Ердеша-Рен'ї;

SF-мережа – scale-free network – безмасштабна мережа;

SW – small world – сеть малого світу.

ВСТУП

Складні мережі – це відносно нова міждисциплінарна область знань, що бурхливо розвивається. Зараз закладаються її основні поняття і отримані тільки перші результати. Працюючи в цій області дослідники прийшли з математики, комп'ютерних наук, фізики, біології, соціології, економіки. Відповідно результати досліджень мають як теоретичне значення, так і практичне використання в цих науках.

У теорії складних мереж виділяють три основних напрямки: дослідження статистичних властивостей, які характеризують поведінку мереж; створення моделі мереж; передбачення поведінки мереж при зміні структурних властивостей. У прикладних дослідженнях зазвичай застосовують такі типові для мережевого аналізу характеристики, як розмір мережі, мережева щільність, ступінь центральності і т.п.

При аналізі складних мереж, як і в теорії графів, досліджуються параметри окремих вузлів; параметри мережі в цілому; мережеві підструктури.

Формально, будь-яка мережа це граф. І, знову ж таки, формально, теорія, яка вивчає властивості складних мереж, повинна була б бути названа теорією графів. Як основою класичної статистичної фізики є класична механіка, так основою теорії складних мереж є теорія графів. При цьому, теорія графів може отримувати змістовні твердження як правило для графів (мереж) невеликого розміру або спеціальної структури. У той час як теорія складних мереж має справу з великою кількістю, як правило, випадковим чином з'єднаних вузлів. У зв'язку з цим, з одного боку, багато поширених питань теорії графів не представляють інтересу для теорії складних мереж. Наприклад, таке питання – планарний даний граф чи ні? Для випадкового графа з великим числом вузлів і зв'язків відповідь очевидна (і не представляє інтересу) – немає. З іншого боку багато дуже важливі поняття в теорії складних мереж для графа з невеликим числом вузлів і

зв'язків або не представляють інтересу, або їх просто важко змістовно сформулювати. Наприклад, така важлива характеристика в теорії складних мереж як функція розподілу за ступенем вузлів може бути підрахована для будь-якого, в тому числі і малої кількості вузлів. Однак в силу свого ймовірного змісту це поняття виявляється корисним тільки в разі великого числа вузлів і зв'язків.

Чим же займається теорія складних мереж? Перерахуємо деякі проблеми і завдання. По-перше, дослідженням стандартних характеристик графів для складних мереж різної природи – випадкових графів, безмасштабних мереж, мереж малого світу і т.п.

По-друге, визначенням і вивченням нових характеристик складних мереж, таких, наприклад, як середній мінімальний шлях, посередництво, коефіцієнт кластеризації.

По-третє, вивченням різних «фізичних» процесів на складних мережах – дифузії, епідемічних процесів, різних потоків (інформації, електричного струму, і т.д.). Зрештою, знаменитий алгоритм PageRank розглядає блукання по зв'язках (гіперпосиланням) в складній мережі WWW.

По-четверте, є дуже важливий в прикладному відношенні напрямок – методи відновлення, захисту та знищення мереж. Таке питання – скільки вузлів (зв'язків) потрібно «вбити» для того, що б, наприклад, зруйнувався «гігантський кластер» або що б значно збільшився мінімальний середній шлях? Тобто, як кажуть сисадміни, що б «мережа лягла». Сюди ж примикають і питання оптимізації мереж.

По-п'яте, пошук латентних, тобто неявних зв'язків.

1 АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ

Складні мережі виникають повсюдно: це і Інтернет, і соціальні мережі, і транспортні мережі, і мережі міжбанківських взаємодій, і біологічні мережі, і багато інших подібних структур. За останні десятиліття стало ясно, що всі ці мережі мають ряд стійких властивостей, дослідження і моделювання яких дозволяє вирішувати різні практичні завдання. Зараз моделювання складних мереж знаходиться на стику багатьох дисциплін – математики, інформатики, фізики, біології, соціології, економіки [3].

Історично вивченням складних мереж займається теорія графів. У той час як теорія графів спочатку описувала регулярні графи, починаючи з 1950-го року, складні мережі, які не мали очевидних принципів побудови, стали описувати за допомогою теорії випадкових графів, запропонованих в якості найбільш відповідної моделі складних мереж. Вперше випадкові графи були вивчені угорськими математиками Полом Ердосом і Альфредом Рен'ї [5].

Такі графи вивчаються теорією графів з ймовірнісної точки зору. Цей напрямок з кожним роком стає все більш актуальним.

Теорія графів відіграє величезну роль у фундаментальній і прикладній математиці. Їй присвячено сотні монографій і тисячі – якщо не десятки тисяч – статей.

Типова постановка питання така: чи велика ймовірність того, що граф має дані властивості? Питання виключно важливе, і ми в цьому не раз переконуємося. Правда, в ньому ні слова не сказано про те, як саме ми розуміємо термін «ймовірність». Кожна людина, яка має уявлення про аксіоматику Колмогорова, добре знає, що можна вкласти безліч різних смислів в цей термін. І його можна дійсно визначати по-різному. Залежно від визначення вийде та чи інша модель випадкового графа.

Суто з математичних позицій будь-яка така модель має право на існування. Однак для додатків – в тому числі додатків до транспортної проблематики – деякі з цих моделей більш цікаві, деякі – менш.

Відповідно, в цій роботі розглянуто два класи моделей, кожен з яких за десятиліття, що минули з моменту своєї появи, зарекомендував себе плідним як в рамках «чистої» математики, так і в рамках її різноманітних додатків, серед яких надійність транспортної мережі, зростання інтернету і інших соціальних і біологічних мереж, теорія алгоритмів та інше. У даній роботі я намагатимуся виділити лише найосновніші та принципові моменти теорії випадкових графів.

1.1 Визначення і напрямки складних мереж

З кінця 1990-х років для вивчення складних систем розвивається новий ефективний інструмент дослідження – теорія складних мереж. Вузли в таких мережах є елементами цих складних систем, а зв'язки між вузлами – взаємодії між елементами. Такі мережі утворюють своєрідний каркас відповідних складних систем, а дослідження властивостей цих мереж дає змістовну інформацію про властивості складних систем в цілому [14].

Мережевий підхід вивчення оточуючого нас світу бере свій початок з 1736 року, коли Леонард Ейлер, що жив тоді в Петербурзі, вирішив знамениту задачу о семи мостах в Кенігсберзі, і заклав тим самим основи нової математичної дисципліни – теорії графів [4].

Проблема семи мостів Кенігсберга – старинна математична задача. З давніх-давен серед жителів Кенігсберга була поширена така загадка: як пройти по всіх мостах через річку Преголя (рис. 1.1), не проходячи за жодним з них двічі.

Багато кенігсбержців намагалися вирішити цю задачу як теоретично, так і практично, під час прогулянок. Втім, довести або спростувати можливість існування такого маршруту ніхто не міг.



Рисунок 1.1 – Схема мостів Кенігсберга

В 1736 задача о семи мостах зацікавила видатного математика – Леонарда Ейлера. Він був членом Петербурзької академії. Про свою зацікавленість він написав у листі італійському математику і інженеру Джованні Джакомо Маринони від 13 березня 1736 року. Там написано, що він зміг знайти правило, користуючись яким, легко визначити, чи можна пройти по всіх мостах, не проходячи двічі по жодному з них. Відповідь була «не можна».

На спрощеній схемі частини міста (графі) мости – це лінії (дуги графа), а частини міста – точки з'єднання ліній (вершини графа). В ході міркувань Ейлер прийшов до таких висновків:

- число непарних вершин (вершин, до яких веде непарне число ребер) графа має бути парне;
- не може існувати граф, який мав би непарне число непарних вершин;
- якщо всі вершини графа парні, то можна, не відриваючи олівця від паперу, накреслити граф, при цьому можна починати з будь-якої вершини графа і завершити його в тій же вершині;
- граф з більш ніж двома непарними вершинами неможливо накреслити одним розчерком.

Граф Кенігсбергських мостів мав чотири (синім) непарні вершини (тобто всі), отже, неможливо пройти по всіх мостах, не проходячи за жодним з них двічі [1].

Незважаючи на те, що в розгляд теорії складних мереж потрапляють різні мережі – електричні, транспортні, інформаційні, найбільший внесок в розвиток цієї теорії внесли дослідження соціальних мереж.

Термін «соціальна мережа» визначає зосередження соціальних об'єктів, які можна розглядати як мережу (або граф), вузли якої – об'єкти, а зв'язки – соціальні відносини.

Цей термін був введений в 1954 році соціологом з «Манчестерської школи» Дж. Барнсом (J. Barnes) в роботі «Класи і збори в норвезькому острівному приході». В XX столітті (другій половині) поняття «соціальна мережа» стало популярним у західних дослідників. У теорії соціальних мереж отримало розвиток такий напрямок, як аналіз соціальних мереж (Social Network Analysis, SNA). Сьогодні термін «соціальна мережа» позначає поняття, яке виявилось ширше свого соціального аспекту, воно включає, наприклад, багато інформаційні мережі, в тому числі і WWW.

У теорії складних мереж виділяють три основних напрямки:

- дослідження статистичних властивостей, які характеризують поведінку мереж;
- створення моделі мереж;
- передбачення поведінки мереж при зміні структурних властивостей.

У прикладних дослідженнях зазвичай застосовують такі типові для мережевого аналізу характеристики, як розмір мережі, мережева щільність, ступінь центральності і т.п [1].

При аналізі складних мереж, як і в теорії графів, досліджуються параметри окремих вузлів; параметри мережі в цілому; мережеві підструктури.

Нова парадигма – «складні мережі» охоплює мережі, що мають властивості:

- великі розміри;
- елементи випадковості при формуванні;
- зростання (зміна) в часі;

– деякі вузли можуть утворювати компактні групи – ансамблі [1].

Важливою віхою в розвитку теорії графів стала середина минулого століття, коли угорськими математиками Ердемем і Ренї була створена теорія випадкових графів, що стала основним напрямком вивчення мережевих структур аж до кінця двадцятого століття. Іноді ці мережі називають пуассоновськими випадковими мережами, оскільки розподіл вузлів по числу зв'язків в таких мережах підкоряється закону Пуассона. Дуже активно мережеві методи використовувалися в соціології практично з моменту зародження цієї наукової дисципліни. Тут були отримані важливі результати і сформульовані важливі поняття, затребувані і сучасною теорією мереж. Найбільш значні результати останніх років у вивченні мережевих структур були отримані фізиками. Виявилося, що методи теоретичної фізики, перш за все теорії критичних явищ, добре підходять для вивчення проблем в цій галузі.

На відміну від математиків, фізики в своїх дослідженнях спираються на емпіричні дані про реальні мережі, такі як Інтернет, Всесвітня павутина, мережі друзів і знайомих або біологічні метаболічні мережі. На відміну від соціологів, фізики вивчають статистичні властивості мереж, наприклад, закони розподілу вузлів по числу зв'язків. При вивченні цих питань були виявлені універсальні властивості реальних мереж, на які не звернули уваги математики і соціологи. Ці властивості послужили стимулом для розробки нових теорій і моделей, виявлення нових фундаментальних властивостей мереж. Саме фізичні журнали в даний час публікують переважну кількість найбільш важливих робіт в цій галузі.

Важливим фактором інтенсифікації вивчення складних мереж в останні роки став високий рівень розвитку алгоритмів і комп'ютерних програм дослідження мереж. Останнім часом стали доступні числені бази даних реальних мережевих структур: інституційну співпрацю акторів кіно і вчених в різних областях науки, мереж білкових взаємодій і метаболічних реакцій в живих клітинах, нейронна мережа хробака нематода *C.elegans*, а також

деяких технологічних мереж, наприклад, мережа приблизно п'яти тисяч електростанцій США. Ці набори даних дозволяють в досить повному вигляді відобразити структуру і функції реальних складних систем [4].

1.2 Моделі графів складних мереж

1.2.1 Мережі Ердеша-Рен'ї

Мережа (граф) Ердеша-Рен'ї – це така мережа, коли кожна пара вузлів з'єднана з ймовірністю p . У межі великого числа вузлів N функція розподілу ступенів вузлів має вигляд:

$$P_k = e^{-\langle k \rangle} \frac{\langle k \rangle^k}{k!}. \quad (2.1)$$

По-іншому можна ще сказати так: в моделі Ердеша – Рен'ї кожне ребро незалежно від всіх інших ребер входить у випадковий граф з ймовірністю p .

Наука про випадковий графі вивчає ймовірності тих чи інших властивостей графів. Найбільший інтерес представляє «динаміка» цих ймовірностей, тобто їх зміна з ростом n . А числу вершин і справді ніщо не заважає рости. При цьому ймовірність виникнення ребра графа теж цілком може еволюціонувати [2].

Це перша модель випадкових графів. На рубежі 50-х і 60-х років XX століття цю модель запропонували класики сучасної комбінаторики і теорії ймовірностей П. Ердеш і А. Рен'ї. Варто відзначити, що Ердеш – це, мабуть, одна з найяскравіших постатей в математиці XX століття. Йому належать сотні статей і завдань, які справили величезний вплив на розвиток багатьох математичних дисциплін. Рен'ї також відіграв значну роль у формуванні угорської ймовірнісної школи, і його ім'ям названо математичний інститут в

Будапешті [7].

В теорії графів модель Ердеша – Ренї є модель для генерації випадкового графа з постійною ймовірністю появи ребра між двома вершинами незалежно від інших ребер. Відповідно до цієї моделі, граф конструюється за допомогою з'єднання пар вершин випадковим чином. Кожне ребро включається в граф з ймовірністю p незалежно для кожної пари вершин. З ростом p від 0 до 1 граф стає все більш насиченим ребрами. При $p = 0$ маємо порожній граф (граф, який не має ребер), при $p = 1$ – повнозв'язний граф. При фіксованій кількості вершин у графі ймовірність p – єдиний вхідний параметр моделі Ердеша – Ренї.

Простота моделі Ердеша – Ренї випадкового графа дозволяє отримувати різні аналітичні оцінки. Але варто відзначити, що модель в середньому не відтворює такі типові властивості реальних мереж, як ступінь вершин, величина коефіцієнта кластеризації і довжини шляху [8].

1.2.2 Масштабно-інваріантні мережі

Масштабно-інваріантна мережа (ще її називають моделлю Барабаші-Альберта) – дуже популярна модель в Complexity Science, а також в новій модній дисципліні – аналізі соціальних мереж.

Масштабно-інваріантна мережа – це зростаюча структура, що складається з особливим чином зв'язаних вузлів-елементів. Її побудова починається з єдиного самотнього вузла. Потім на кожному кроці еволюції в структуру додається по одному новому елементу, який зв'язується з одним з вже наявних в мережі елементів. Так мережа постійно збільшується, зростає.

Ключова особливість масштабно-інваріантної мережі полягає в алгоритмі з'єднання вузлів. Новий вузол тим імовірніше приєднається до якогось старого вузла, чим більше інших вузлів до нього вже приєдналося. Нехай, наприклад, у нас є мережа, в якій вже три вузла (рис. 1.2).

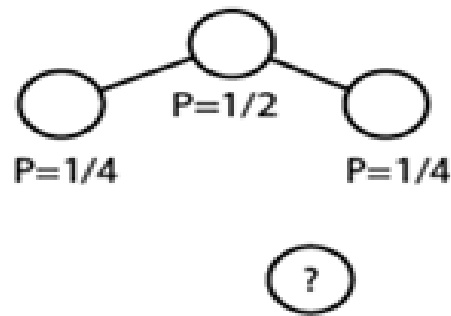


Рисунок 1.2 – Приклад мережі з різною ймовірністю приєднання
нового вузла

Ймовірність приєднання нового вузла до одного з них визначається пропорцією $1/4:1/2:1/4$, тому що відповідні вузли мають зв'язками в пропорції $1:2:1$. В даному випадку з ймовірністю $1/2$ новий вузол приєднається до другого вузла – і це найбільш вірогідний розвиток подій. Взагалі, вузол тим швидше накопичує нові зв'язки, чим їх у нього більше, тому що найімовірніше, що нові вузли будуть приєднуватися саме до нього. Тобто, «багаті» вузли стають ще «багатшими». Навпаки, «бідні» вузли схильні так і залишатися «бідними» (цей алгоритм часто називають правилом «багатий стає багатшим»).

В результаті розвивається мережа характерного вигляду, яку і називають масштабно-інваріантною (рис. 1.3).

Слід зауважити, що це живописне зображення масштабно-інваріантною мережі умовне, тому що насправді у вузлів моделі немає якогось просторового розташування, воно для вузлів не задано. Ми б могли зобразити мережу яким завгодно заплутаним чином.

Очевидна особливість цієї структури – є дуже багато вузлів, у яких мало зв'язків, і дуже мало вузлів у яких багато зв'язків.

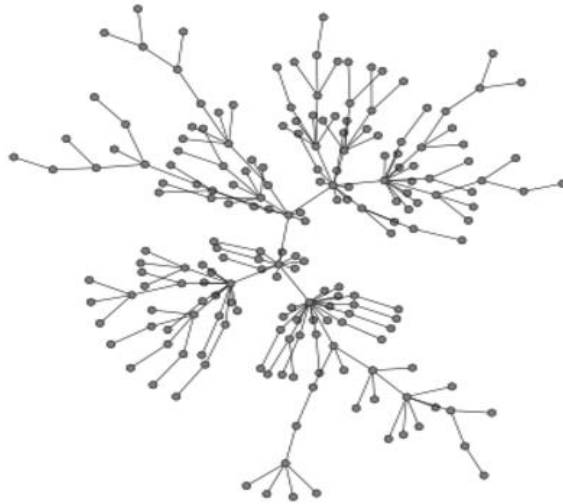


Рисунок 1.3 – Масштабно-інваріантна мережа

Важливо, що це степеневий розподіл виникає вже на ранніх стадіях еволюції моделі (коли вузлів ще не більше 100) і зберігається майже незмінним на всьому доступному для огляду за допомогою звичайного комп'ютера горизонті часу.

Чому утворену структуру називають масштабно-інваріантною мережею? Як говорилося раніше, ступенному розподілу характерні для параметрів, які не мають характерного масштабу, тобто, для масштабно-інваріантних. В даному випадку таким параметром є кількість зв'язків у вузлів. Степеневий розподіл їх кількості означає, що у цього параметра немає якогось «правильного» або характерного значення, так що він є масштабно-інваріантним.

Отже, як ми бачимо, модель масштабно-інваріантною мережі дуже проста і елегантна, і це робить її безумовно привабливою в якості пояснення походження багатьох натуральних степеневих розподілів. Зокрема, як правило, з її допомогою пояснюють степеневий розподіл популярності веб-сайтів, степеневу статистику завантаження інтернет-вузлів і т.д. Далі, правило «багатий стає багатшим» природно наводить на думку пояснювати з її допомогою степеневий розподіл багатства серед представників соціуму – це цілком співзвучне розхожою максими «гроші притягують гроші».

Нарешті, ця модель здається привабливою можливістю пояснити походження закону Зіпфа в розподілах населення по містах.

Дійсно, уявімо, що кожен вузол являє собою конкретне географічне місце, а зв'язки між ними – це люди. Тоді появу нового вузла можна зіставити з ситуацією народження нової людини в певній географічній точці, а зв'язок, який новий вузол протягне до одного з уже наявних вузлів – це відображення вибору місця проживання, який в кінцевому підсумку зробить ця людина. Правило «багатий стає багатшим» в даному випадку відображає перевагу людей вибирати для життя ті місця, де вже мешкає багато людей, тому що в цьому випадку відкриваються максимальні можливості для реалізації своїх здібностей, збагачення і т.д.

Модель масштабно-інваріантної мережі призводить до степеневого розподілу вузлів за кількістю зв'язків, і це, начебто, виглядає як просте і красиве пояснення степеневого розподілу міст за кількістю жителів [9].

1.2.3 Мережі малого світу Ваттса – Строгатца

У 70-і роки минулого століття американський психолог Милграм (Milgram) провів цікаве дослідження. Він задався питанням, яка «відстань» між двома випадково вибраними людьми. Під відстанню розуміється кількість знайомств, необхідне для встановлення зв'язку між даними людьми. Милграм поступив таким чином – оскільки він жив в Бостоні, то було обране далеке від Бостона місто – Небраска, і випадково обраним людям було роздано конверти, які потрібно було передати в Бостон. Конверти можна було передавати тільки через своїх знайомих і родичів. Милграм отримав вельми несподіваний результат: в середньому кожен конверт пройшов через шість чоловік. Так і народилася теорія «шестиоюрукостискань». Тобто кожна людина пов'язана з будь-яким іншою ланцюгом не більш ніж шість особистих знайомств. У цьому сенсі про наш світ говорять як про малий світі – «small world».

Модель переходу від великого (регулярного) світу до малого запропонували Ваттс (Watts) і Строгац (Strogatz). Ця модель являє собою одновимірну регулярну решітку, що складається з N вузлів, де кожен вузол з'єднаний тільки зі своїми k найближчими сусідами і накладені періодичні граничні умови, тобто ґрати згорнули в кільце. Після чого кожен зв'язок з імовірністю перекидали на інший випадково обраний вузол. Правда при такій процедурі є ймовірність появи ізольованих вузлів (рис 1.4).

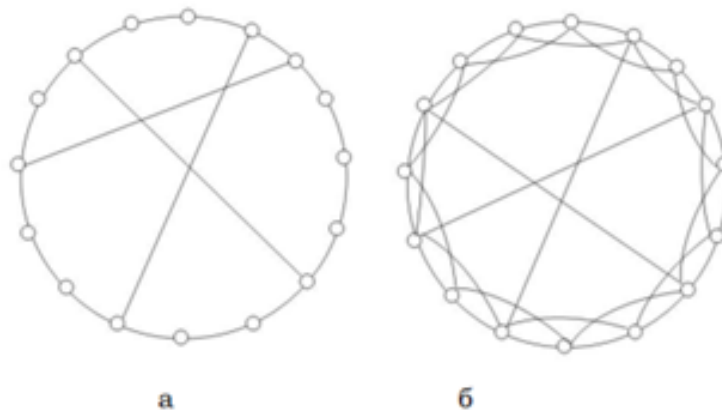


Рисунок 1.4 – Приклад малого світу з трьома перебрсами: а) кожен вузол з'єднаний зі своїми найближчими сусідами; б) кожен вузол з'єднаний з чотирма сусідами

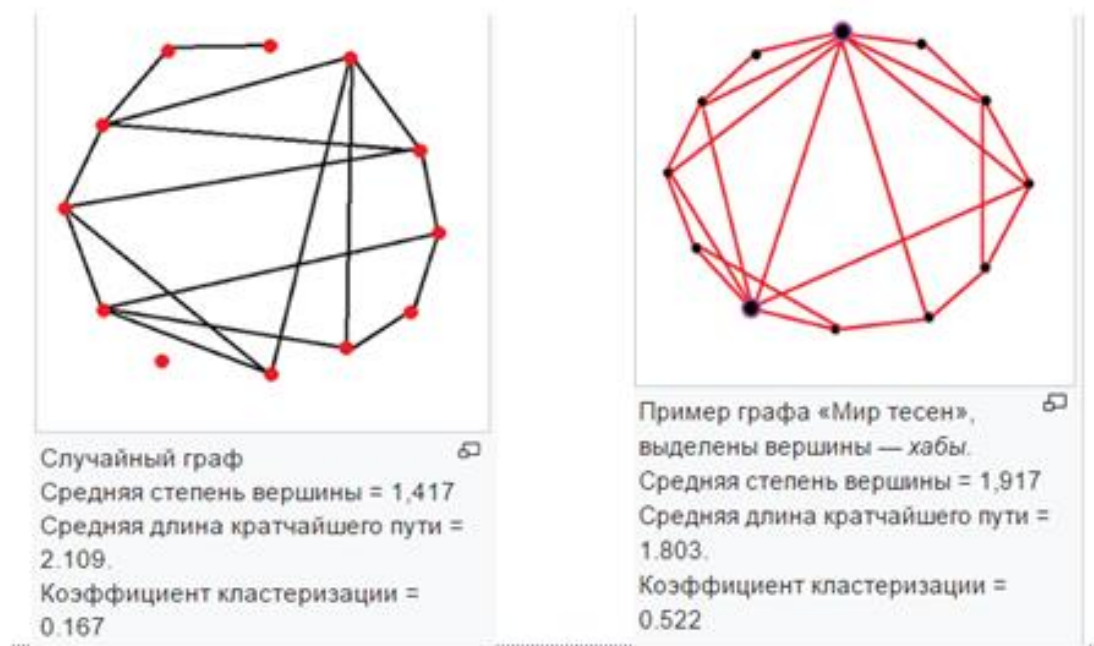
Формально відстань до них від будь-якого вузла буде нескінченною. Щоб уникнути цього, Ньюман (Newman) і Ваттс запропонували зв'язки не перекидати, а просто додавати [1].

Модель Ватца-Строгатца є основним механізмом конструювання графів «Світ тісний». Граф «Світ тісний» (маленький світ) – різновид графа, який має наступну властивість: якщо взяти дві довільні вершини a і b , то вони з великою ймовірністю не є суміжними, проте одна досяжна з іншої за допомогою невеликої кількості переходів через інші вершини. А саме: мережа, в якій типова відстань L між двома довільно обраними вершинами росте пропорційно логарифму від числа вершин N , таким чином: $L \propto \log N$.

В контексті соціальної мережі це призводить до феномену «Світ тісний», тобто незнайомих людей пов'язує невелика кількість проміжних знайомих. Багато реально існуючих графів добре моделюються через графи «Світ тісний». Соціальні мережі, зв'язність мережі Інтернет, вікі-сайти, такі, як Вікіпедія, і генні мережі проявляють властивості графа «Світ тісний». Дункан Ватц і Стівен Строгац в 1998 році ідентифікували певну категорію графів «Світ тісний» як клас випадкових графів. Вони відзначили, що такі графи можуть бути класифіковані відповідно до двох незалежних структурних особливостей, а саме коефіцієнт кластеризації і відстань від однієї вершини до іншої в середньому (також відоме як довжина найкоротшого шляху в середньому). Абсолютно випадкові графи, побудовані відповідно до моделі Ердеша – Рен'ї, мають малу довжину найкоротшого шляху в середньому (вона росте як логарифм від кількості вершин у графі) і низький коефіцієнт кластеризації. Ватц і Строгац з'ясували, що більшість реально існуючих мереж мають малу довжину найкоротшого шляху в середньому, але коефіцієнт кластеризації в них істотно вище, ніж очікується при випадковому виборі. Після цього Ватц і Строгац запропонували нову модель графа, в даний час звану модель Ватца-Строгатца, для якої характерні мала довжина найкоротшого шляху в середньому, і великий коефіцієнт кластеризації (рис. 1.5).

Властивості графа «світ тісний» були виявлені в багатьох об'єктах реального світу, включаючи карти доріг, харчові ланцюжки, електричні мережі, мережі протікання метаболізму (metabolite processing networks), нейронні мережі, мережі виборців, граф телефонних дзвінків і мережі соціального впливу.

Білок-білкові взаємодії мають таку властивість графа «Світ тісний», як відповідність степеневим законом розподілу. Аналогічно, властивості графа «Світ тісний» є у регуляції транскрипції, в якій вершинам відповідають гени, причому вони пов'язані, якщо один ген має вгору- або вниз- регулюючий генетичний вплив на інший [6].



а)

б)

Рисунок 1.5 – Приклади графів: а) випадковий граф; б) граф «Світ тісний»

1.3 Постановка задачі дослідження

Метою даної роботи є моделювання та дослідження асортативності випадкових безмасштабних мереж.

Для досягнення мети роботи необхідно вирішити наступні завдання:

- розглянути існуючі моделі мереж та їх характеристики;
- провести аналіз моделей безмасштабних мереж;
- провести дослідження асортативності, її числових мір, впливу асортативності на структуру мережі;
- провести дослідження методів зміни асортативності шляхом перенаправлення зв'язків;
- здійснити програмну реалізацію генерації безмасштабних мереж, вимірювання асортативності мереж, перенаправлення зв'язків.

2 АСОРТАТИВНІСТЬ МЕРЕЖ

2.1 Поняття асортативності мереж

Поняття асортативності було введено Ньюменом [32] у 2002 році і з того часу широко вивчається. Асортативність – це одна з метрік графів. Вона представляє, якою мірою вузли в мережі асоціюються з іншими вузлами в мережі, будучи подібними або протилежними. Взагалі, асортативність мережі визначається за ступенем (тобто кількістю прямих сусідів) вузлів у мережі. Однак концепція асортативності може бути застосована і до інших характеристик вузла, таких як вага вузла, ядровість, посередництво вузла, ступінь вузла k -го рівня (кількість вузлів, яких можна досягти за не більше, ніж k кроків. Крім того, асортативність може бути застосована до характеристик вузлів, які не пов'язані безпосередньо з топологією.

Асортативність виражається у вигляді скалярного значення, ρ , в діапазоні $-1 \leq \rho \leq 1$. Ступенева асортативність позначається як ρ_D . Мережа називається асортативною, коли вузли з великою кількістю зв'язків (з великим ступенем) в середньому з'єднані з вузлами, які також мають велику ступінь, а вузли з малою кількістю зв'язків (з низьким ступенем) в середньому з'єднані з вузлами, які також мають низьку ступінь.

Мережа називається дизасортативною, якщо в середньому вузли високого ступеня підключаються до вузлів з низьким ступенем, а вузли низького ступеня підключаються (в середньому) до вузлів з високим ступенем.

Асортативність надає інформацію про структуру мережі, а також про динамічну поведінку мережі та надійність мережі на випадок, наприклад, випадкової або цільової атаки та розповсюдження вірусів. Головна характеристика мережі – розподіл вузлів за ступенем – як правило, не дає достатньої в цьому сенсі інформації про мережу.

Оригінальне визначення асортативності (Ньюмен [32]) для незважених,

неорієнтованих мереж базується на понятті кореляції між випадковими змінними. Коефіцієнт лінійної кореляції між двома випадковими змінними X і Y має вигляд:

$$\rho(X, Y) = \frac{E\{XY\} - \mu_X \mu_Y}{\sigma_X \sigma_Y} = \frac{E\{X - E\{X\}\} \cdot E\{Y - E\{Y\}\}}{\sqrt{E\{(X - E\{X\})^2\} \cdot E\{(Y - E\{Y\})^2\}}}, \quad (2.2)$$

де μ_X і μ_Y – середні значення X і Y відповідно;

$E\{XY\} - \mu_X \mu_Y$ – коваріація між X та Y ;

σ_X та σ_Y – середньоквадратичні відхилення.

Для застосування загального виразу (2.2) до розрахунку коефіцієнта асортативності застосовується наступний підхід: Випадково обираються зв'язки l у графі. Зв'язки з'єднують два вузли: початковий вузол, позначений l_- , та кінцевий вузол, позначений l_+ . Ступіні цих вузлів позначаються відповідно як D_{l_-} та D_{l_+} . Тоді згідно з Ньюмен [32] похідний від рівняння (2.2) коефіцієнт асортативності має вигляд:

$$\rho_D = \frac{1}{\sigma_q^2} \sum_{j,k} jk(e_{jk} - q_j q_k), \quad (2.3)$$

де e_{jk} – ймовірність спільного надлишкового ступеню для вершин з надлишковими ступенями j та k (надлишкова ступінь вузла визначається як ступінь цього вузла мінус одиниця);

q_k – нормалізований розподіл надлишкового ступеня для випадково вибраного вузла (2.4);

σ_q – стандартне відхилення величини q_k .

$$q_k = \frac{(k+1)p_{k+1}}{\sum_j jp_j}. \quad (2.4)$$

Асортативність кількісно визначається коефіцієнтом кореляції Пірсона між X і Y . Поділ на σ_q^2 (тобто на дисперсію q) служить для нормалізації асортативності в діапазоні $[-1;1]$.

Обґрунтуванням застосування надлишкових ступенів вузлів

$$p_i = d_i - 1 \quad (2.5)$$

є те, що тенденція до з'єднання двох вузлів i та j визначається з моменту, коли ці вузли ще не пов'язані, тобто мають ступінь $d_i - 1$ та $d_j - 1$.

Визначення асортативності (2.3) не є зручним для практичних обчислень. Більш зручна форма показника асортативності було отримана та має вигляд

$$\rho = \frac{S_3 - S_2^2 / S_1}{N_3 - S_2^2 / S_1}, \quad (2.6)$$

$$S_k = \sum_{i=1}^N d_i^k, \quad N_3 = d^T A d = \sum_{i=1}^N \sum_{j=1}^N a_{ij} d_i d_j. \quad (2.7)$$

Можна побачити, що $S_0 = N$, $S_1 = 2L$ (подвоєна кількість зв'язків).

Рівняння (2.6) може бути модифіковано далі до

$$\rho_D(G) = 1 - \frac{\sum_{i \neq j} (d_i - d_j)^2}{\sum_{i=1}^N d_i^3 - \frac{1}{2L} \left(\sum_{i=1}^N d_i^2 \right)^2}. \quad (2.8)$$

Мережа являє собою граф $G(N, L)$ з набором вузлів N та зв'язків L . Граф може бути представлений через матрицю суміжності A розміром $N \times N$, з елементами $a_{ij} = 1$, якщо зв'язок між вузлом i та вузлом j існує, та $a_{ij} = 0$, якщо між вузлом i та вузлом j не існує зв'язку. Вважається, що петлі відсутні (отже, $a_{ii} = 0$) і немає ланок, що перекриваються, тобто не може бути більше однієї ланки між i та j . Такий граф відомий як простий граф.

Для неорієнтованих графів зв'язки (які в цьому разі називають ребрами) не мають напрямку. Для таких графів матриця суміжності симетрична, тобто $a_{ij} = a_{ji}$ та $A = A^T$. Для орієнтованих графів посилення мають напрямки, тоді $a_{ij} \neq a_{ji}$.

Графи можуть містити зважені зв'язки, і в цьому випадку зв'язок має вагу, пов'язану з нею. Зазвичай ця вага w_{ij} виражається невід'ємним числом. Відповідна зважена матриця суміжності W містить ваги w_{ij} замість a_{ij} .

Головною характеристикою мережі є розподіл вузлів за ступенями, втім ця характеристика не може розглядатися як вичерпна, що й було однією з причин, що спонукали Ньюмена впровадити поняття асортативності.

Існує багато типових топологій мереж (класів): модель Ердеша-Рені (ER), модель Барабаші-Альберт (BA), модель Калавея (C), модель «малого світу» Ваттса-Строгатца (WS), решітка та інші.

Граф класу ER містить набір з N вузлів. Вузли в мережі з'єднані посиленням з ймовірністю p . Наявність зв'язку між парою вузлів стохастично не залежить від наявності зв'язку між будь-якою іншою парою вузлів. Мережа ER має біноміальний розподіл ступенів вузлів (d_k):

$$\Pr\{D = k\} = C_{N-1}^k p^k (1-p)^{N-1-k}. \quad (2.9)$$

Для великих N у разі незалежності величини $\lambda = p(N-1)$ від N , розподіл ступенів (2.9) розвивається в Пуассонівський:

$$\Pr\{D = k\} = e^{-\lambda} \frac{\lambda^k}{k!}. \quad (2.10)$$

Граф класу ВА – це зростаюча мережна модель. Мережа побудована з початкового графа, як правило, повного графа з N_0 вузлами, до якого ітеративно додається вузол. Кожен щойно доданий вузол з'єднаний з m існуючими вузлами. Імовірність приєднання до випадково вибраного вузла пропорційна ступеню цього існуючого вузла, що пояснює термін переважне приєднання. Для великих N граф ВА має безмасштабний (ступеневий) розподіл ступенів вузлів d_k :

$$\Pr\{D = k\} = ck^{-3}. \quad (2.11)$$

Узагальненням моделі ВА є безмасштабний граф (SF, scale-free), для якого (5) відповідно узагальнюється до

$$\Pr\{D = k\} = ck^{-\gamma}, \quad \gamma > 1. \quad (2.12)$$

Для більшості мереж реального світу $1.2 \leq \gamma \leq 3.5$. Параметр c є нормалізуючим фактором, який дорівнює неповній функції дзета-Рімана $c = \zeta(k, N)$, так, що забезпечується тотожність

$$\sum_{k=1}^{N-1} \Pr\{D=k\} = c \cdot \sum_{k=1}^{N-1} k^{-\gamma} = 1. \quad (2.13)$$

Граф WS будується із звичайного графа, такого як (зазвичай) кільцева решітка. Для повторного з'єднання кожної ланки, з ймовірністю p_r до випадково вибраного вузла застосовується ітераційний процес перенаправлення зв'язків (rewiring procedure). Термін «малий світ» використовується тому, що мережі цього класу мають найкоротші шляхи між вузлами. У той же час коефіцієнт кластеризації графа C залишається високим. Коефіцієнт кластеризації графа C є мірою кількості трикутників у графі Δ_i , відносно кількості триплетів («виделок») V_i у графі:

$$C = 3 \frac{\sum \Delta_i}{\sum V_i}. \quad (2.14)$$

Коли p_r збільшується до одиниці, мережа WS буде близькою до мережі ER. Ступені вершин графа WS відцентровані відносно ступенів вузлів, не задіяних у процедурі перенаправлення зв'язків. Якщо розглянути граф WS, побудований з кільцевої решітки без перенаправлення зв'язків ($p_r = 0$), то ступінь кожного вузла однакова, оскільки кожен вузол з'єднаний з однаковою кількістю сусідів. У міру того, як ймовірність перенаправлення зв'язків p_r збільшується або як збільшується кількість циклів перенаправлення, граф стає більш випадковим і розподіл ступеня еволюціонує до біноміального розподілу (3), або розподілу Пуассона (4), залежно від N та L . При $p_r = 0$ середня довжина найкоротших шляхів зростає як $O(N)$. Коли p_r стає достатньо великим, середня довжина найкоротших шляхів стає $O(\log(N))$.

Модель графів Callaway, так саме, як і BA, є класом, який описує графи, що зростають. Граф будується з початкового повного графу з малим N_0 , напр. 5. Потім вузли додаються ітераційно. Для кожного доданого вузла

додається посилення між двома вузлами. Два вузли обираються випадковим чином із набору існуючих вузлів та нещодавно доданого вузла. Отже, модель мережі Callaway може вважатися варіантом моделі мережі ВА. Розподіл ступенів вузлів у цій мережі має експоненціальну форму.

Решітка може бути кільцевою, сферичною, квадратною, кубічною тощо. Решітка – це звичайна структура, за допомогою якої кожен вузол підключається до визначеної кількості сусідів. Однак для квадратної решітки, наприклад, вузли на краю мережевої структури мають менший ступінь.

Дводольна мережа (узагальнена до k -дольної мережі). Дводольна мережа включає одну підмережу з n вузлами та іншу підмережу з m вузлами. Зв'язки існують між вузлами різних підмереж, але не між вузлами однієї з підмереж.

Зоряна мережа є особливим випадком (n, m) -дводольної мережі, для якої $n = 1$.

Деревовидна мережа – це зв'язаний граф з $L = n - 1$ зв'язками.

Складні мережі певного класу мають певні ознакові характеристики, одна з яких – розподіл вузлів за ступенями. Перенаправлення зв'язків змінює топологію мережі з метою зміни певної характеристики мережі. Перенаправлення зв'язків може призвести до того, що мережа перетвориться з графу одного класу в граф іншого класу. Якщо розглянути, наприклад, граф ВА та застосувати випадкове перенаправлення зв'язків, то мережа поступово стане мережею, що має характеристики випадкової мережі ER.

Ступінь D випадково вибраного вузла в графі представляє собою кількість прямих сусідів. Вектор ступенів позначається як $(d_1, d_2, \dots, d_n)^T$. Функція щільності розподілу ступенів (PDF – probability density function), визначається як $f_D(k) = \Pr(D = k)$. Функція щільності сумісного розподілу ступенів описує ймовірність того, що два випадково обрані вузли мають дві конкретні значення ступеня: $f_{D_1 D_2}(k, l) = \Pr(D_1 = k, D_2 = l)$. Якщо випадковий вибір першого вузла стохастично не залежить від випадкового вибору другого вузла, то $f_{D_1 D_2}(k, l) = f_D(k) \cdot f_D(l)$. Асортативність стосується

спільного розподілу ступенів зв'язків у мережі. При випадковому виборі зв'язка l у графу ступінь вузла на одному його кінці, d_{l-} , не є стохастично незалежною від ступеня вузла на іншому кінці зв'язка, d_{l+} . Асортативність – це міра ступеня співвіднесеності d_{l-} та d_{l+} для мережі.

2.2 Вплив асортативності на властивості мереж

Коли ми розглядаємо граф ER, існування зв'язку між двома вузлами не залежить від ступенів цих вузлів. Розглядаючи граф BA, ми зауважуємо, що переважне приєднання зв'язку між щойно доданим вузлом та існуючим вузлом є функцією ступеня цього існуючого вузла. Для графа BA з досить великим розміром мережі ми спостерігаємо, що в середньому для кожного вузла i розподіл ступеня вузлів, підключених до вузла i , не залежить від ступеня вузла i ; тобто кореляція ступеня не існує.

Ньюмен зауважив, що у багатьох реальних мережах ступінь вузлів, підключених до випадково вибраного вузла i залежить від ступеня цього вузла та від ступеня самого вузла i . Дві мережі з однаковим векторним ступенем можуть мати різну асортативність. Деякі значення емпірично знайденої асортативності, а також результати моделювання штучних моделей мереж наведені в таблиці (рис. 2.1).

Network	assortativity
Physics coauthorship [4]	0.363
Mathematics coauthorship [17]	0.120
Company directors (see [37] for reference)	0.276
Connections between autonomous systems on the Internet [7]	-0.189
World-Wide Web (see [33] for reference)	-0.067
Undirected hyperlinks between Web pages in a single domain [2]	-0.065
Neural network (see [37] for reference)	-0.163
Experimental Erdős-Rényi (ER) graph (for sufficiently large network size)	~ 0
Experimental Barabási-Albert (BA) graph (for sufficiently large network size)	~ 0

Рисунок 2.1 – Таблиця асортативності різних мереж

Автори фізики та автори математики, хоча останні в меншій мірі, прагнуть публікувати статті з співавторами, які мають однаково високу або однаково низьку кількість публікацій, тобто утворюють асортативну мережу. Що стосується з'єднань в Інтернеті, то можна помітити, що високоз'єднані автономні системи (хаби) підключені до інших хабів, які самі мають мало з'єднань, тобто утворюють дизасортативну мережу. Це ж спостереження стосується і непрямих гіперпосилань між веб-сторінками.

Для штучних (тобто генерованих) мереж ER та BA асортативність близька до нуля.

Асортативність має пряме відношення до надійності мережі, з точки зору сполученості мережі. Полومка або націлена атака на вузол високого ступеня в асортативній мережі залишить інші вузли високого ступеня, з'єднані один з одним. Це мінімізує шанс відключення мережі в цілому. У дизасортативній мережі вузли високого ступеня менш пов'язані один з одним. Багато шляхів між вузлами в мережі залежать від вузлів високого ступеня. Полумка вузла високого ступеня в дизасортативній мережі, отже, матиме більший вплив на зв'язок мережі.

Ньюмен [33] проводить загальне дослідження асортативності, використовуючи різні характеристики вузла, на основі яких може відбуватися асортативне змішування. Піравенен [42] провів широкомасштабне дослідження поняття асортативності. Він також дослідив структуру мереж, які є ідеально асортативними або ідеально дизасортативними. У ідеально асортативній мережі всі вузли підключені до інших вузлів одного ступеню. Одним із прикладів є повна мережа, де всі вузли підключені до всіх інших вузлів і всі вузли мають ступінь $d_i = n - 1$. Для такої мережі асортативність максимальна: $\rho_D = 1$. Якщо мережа складається з вузлів з різними ступенями, то ідеальна (тобто одинична) асортативність мережі все ж можлива. Ідеальної асортативності було б досягнуто, коли мережа подріблена на підмережі, кожна з яких є повною мережею (рис. 2.2).

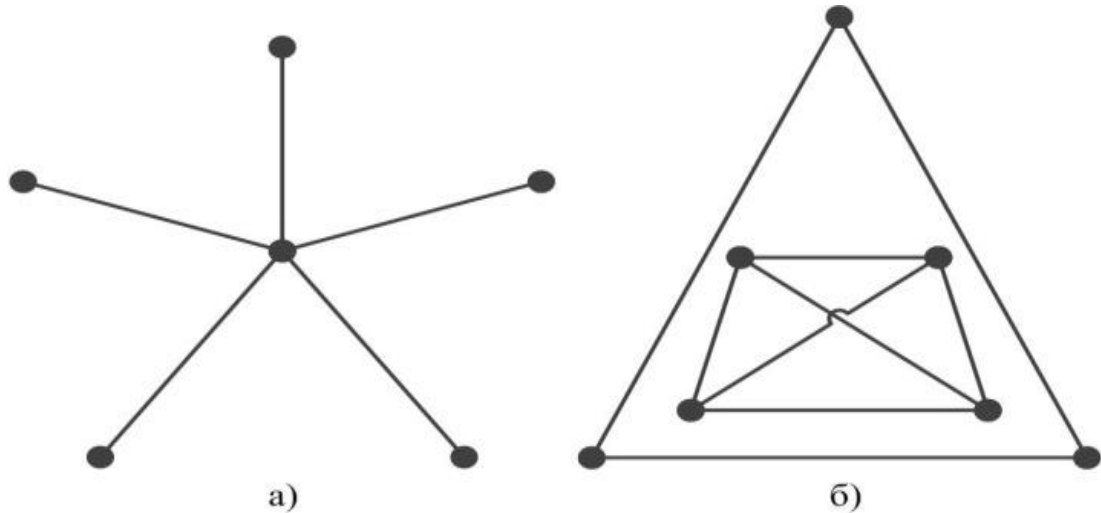


Рисунок 2.2 – Структури максимально дизасортативної мережі (а) та максимально асортативної мережі (б)

Ідеальної дизасортативності досягти важче. Одним з класів мереж, який є ідеально дизасортативним, – це повний дводольний граф (m, n) , з $m \neq n$. Зіркоподібний граф – приклад повного дводольного графа $(m, 1)$ (рис. 2.1).

Piraveenen та ін. [39] пов'язують асортативність також із інформаційним змістом мережі. Мережі, які є асортативними, або дизасортативними, мають більш високий інформаційний вміст, ніж мережі з нейтральною (нульовою) асортативністю. Розглядаючи випадковий вузол у мережі, яка є асортативною, або дизасортативною, ми знаємо, яку ступінь можна очікувати для вузлів, підключених до вузла i . Розглядаючи випадковий вузол у нейтральній за асортативністю мережі міри, ми не можемо робити ніяких висновків щодо очікуваного ступеня вузлів, підключених до вузла i . Цей зв'язок між асортативністю та мережевою інформацією також спостерігав Соле та Вальверде [45]. Вміст інформації, пов'язаний з набором зв'язків графа, визначено наступним чином:

$$I(q) = \sum_{j=0}^{N-1} \sum_{k=0}^{N-1} e_{j,k} \ln \frac{e_{j,k}}{q_j q_k}, \quad (2.15)$$

де q_j, q_k – це ймовірності вузлів мати надлишкові ступені j та k відповідно (надлишкова ступінь вузла на одиницю менше кількості зв'язків цього вузла);

$e_{j,k}$ – ймовірність того, що випадково обране ребро з'єднує вершини з надлишковими ступенями j та k .

Згідно з (2.15) на інформаційний вміст мережі впливає співвідношення між ймовірністю спільного ступеня $e_{j,k}$ та добутком відповідних ймовірностей q_j та q_k . Змінюючи топологію мережі через перенаправлення зв'язків, можна збільшувати або зменшувати інформаційний вміст цієї мережі. Píraveenan [40] спостерігали прямий зв'язок між інформаційним вмістом мережі та її асортативністю / дизасортативністю. Асортативність не виявляє інформації про окремі вузли. Мережа з заданою асортативністю містить вузли, які сприяють цій асортативності. Не всі вузли в однаковій мірі сприяють асортативності мережі. Мережа, яка в цілому не є асортативною, може містити вузли, які самі по собі є високоасортативними. Мережі з однаковою асортативністю можуть мати різний розподіл внеску асортативності на вузол. Внесок вузла у асортативність мережі називається локальною асортативністю [42] або асортативністю вузла.

Асортативність для мереж із зваженими зв'язками вивчав Чанг [6]. Обґрунтування полягає в тому, що в багатьох мережах зв'язок між двома вузлами може бути слабким або сильним, тобто може мати вагу, що відображає такі аспекти, як можливість передачі даних, вартість передачі даних або довжина. IP-мережі на базі OSPF – це один практичний приклад, коли посилання мають певну вагу, а саме ємність передачі даних між двома підключеними маршрутизаторами OSPF4. Ступінь вузла – це вже не дискретна кількість, а скоріше набір реальних змінних. Чанг та ін. [6] пропонують, що сила міцності вузла – це сума ваг зв'язків, з'єднаних з цим вузлом. Отже, асортативність цих мереж стосується тенденції вузлів однакової або протилежної сили до з'єднання один з одним. Це визначення сили вузла також застосовується Wang et al. [63], який вивчає, серед іншого,

розподіл ваги зв'язку для одного вузла. Асортативність для мереж із зваженими зв'язками також вивчає Леун та Чау [21]; зважена асортативність ρ_w запропонована в [21]. Розглядаючи мережу із зваженими посиленнями, ми також можемо розрахувати звичайну (незважену) асортативність, ігноруючи вагу посилення. Потім спостерігається (наприклад, у [21]), що ρ_w та ρ_D можуть істотно відрізнятися для мережі. Однак можна поставити під сумнів справедливність ігнорування ваг посилення у зваженій мережі. У цьому випадку посилення з невеликою вагою має те саме значення з'єднання, що і зв'язок з високою вагою, тоді як ці посилення можуть значно сприяти надійності мережі та іншим характеристикам мережі.

У роботі König та ін. [20] представлена модель мережевої трансформації, за якою стаціонарна, неасортативна або дизасортативна мережа мігрує до асортативної мережі. Мережеве перетворення включає комбінацію додавання та видалення посилення (розпад існуючих посилення). Сенс цієї моделі полягає в тому, що вузли в мережі (наприклад для мережі взаємодії людини або технологічної мережі), обмежені кількістю зв'язків, які вони можуть підтримувати. Ця модель виразно відрізняється від моделей зростання мережі для генерації мережі із заданою асортативністю. Для цих моделей мережа зростає за рахунок додавання вузлів і посилення, за деяким правилом. Ці моделі не визначають обмеження в кількості посилення, що трапляються на вузол. Вони також не розглядають питання про видалення посилення. Літтак та Ван дер Хофстад [25] зауважують, що мережі, які за своєю суттю є дизасортативними, наприклад, Інтернет, демонструють дизасортативність ступеня зменшення, оскільки з ростом розміру мережі n показник асортативності ρ_D рухається в бік 0.

Широкий розподіл ступенів (діапазон значень ступенів) має зменшуючий вплив на значення асортативності. Для великих мереж розподіл ступеня в середньому буде ширшим, ніж для малих мереж. Щоб пом'якшити цей очевидний недолік визначення асортативності пропонується [25]

використовувати кореляцію рангів як альтернативний метод розрахунку асортативності.

Кореляція рангів визначається наступним чином:

$$\rho_n^{rank} = \frac{\sum_{i=1}^n \left(r_i^X - \frac{n+1}{2}\right) \left(r_i^Y - \frac{n+1}{2}\right)}{\sqrt{\sum_{i=1}^n \left(r_i^X - \frac{n+1}{2}\right)^2 \cdot \sum_{i=1}^n \left(r_i^Y - \frac{n+1}{2}\right)^2}}, \quad (2.16)$$

де X і Y – випадкові величини, що представляють ступінь на кінцях випадково обраного зв'язка мережі;

змінні r_i^X та r_i^Y – ранги спостережень X_i та Y_i для випадку, коли значення вибірок впорядковані у порядку зменшення.

Ранкова кореляція ґрунтується на класичній мірі ρ Спірмена [46] (на відміну від асортативності, яка базується на коефіцієнті кореляції Пірсона). Застосування (2.16) дозволяє виявити дизасортативність в мережах навіть для великих n , тому пропонується, щоб кореляція рангів була поряд із асортативністю стандартним інструментом для аналізу складних мереж, зокрема мереж, які за своєю природою не є безмасштабними.

Слід зазначити, що для традиційного визначення асортативності (2.3) існують аналітичні її позначення для графів (2.6), (2.8). Для асортативності у формі (2.16), яка заснована на ранговій кореляції, відповідної аналітичної форми для графів не існує. Під час аналізу графів з аналітичним визначенням асортативності (2.6), (2.8) працювати набагато простіше, ніж з оригінальним (2.3), (2.16). Це робить кореляційну асортативність (2.3) більш привабливою метрикою, ніж рангова (2.16).

2.3 Асортативність у неорієнтованих мережах

Мережі можуть бути класифіковані за розподілом ступенем їх вузлів. Цей розподіл характеризується функцією щільності (PDF), $f(k) = \Pr\{D=k\}$ цієї мережі. На рисунку 2.3 показано ряд прикладів PDF для графів різних класів. В обох випадках кількість вузлів складала $N=10000$. Для ER-мережі використовувався параметр $p=0.1$, для ВА-мережі – $m=2$, $N_0=4$.

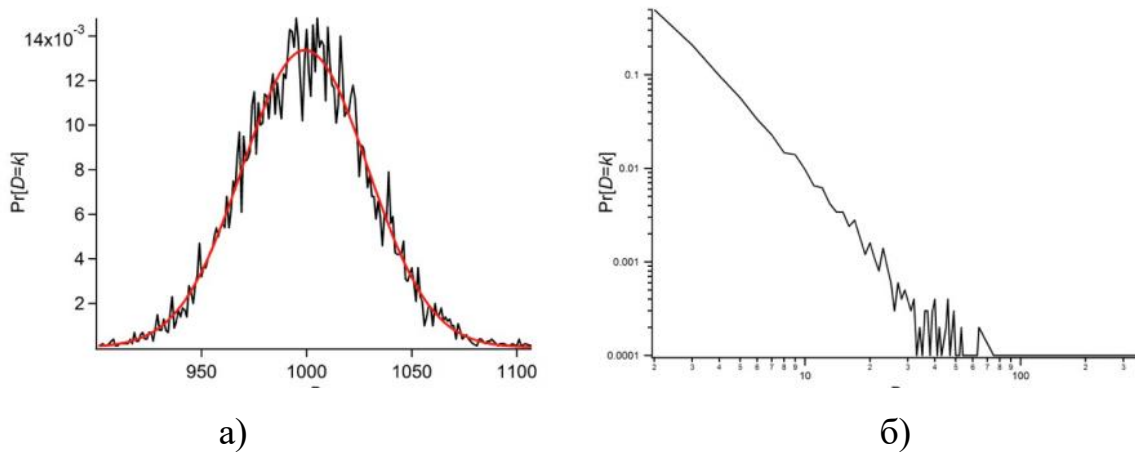


Рисунок 2.3 – Щільності розподілу вершин за ступенями: а) граф ER, б) граф ВА

Коли ми розглядаємо один вузол у графі, ймовірність того, що той вузол має певний ступінь, впливає із розподілу ступеня цього графа. Конкретний вузол i зі ступенем $d_i = k$ з'єднаний з k іншими вузлами. Кожен з цих k інших вузлів має свій ступінь $d_{j|i} = l$. Усереднюючи середній ступінь вузлів, що поєднані з даним $\bar{d}_i = \frac{1}{k} \sum_k d_{j|i}$ можна визначити співвідношення між d_i та d_j . Враховуючи, що може бути декілька вузлів із певним ступенем k , ми оцінюємо це співвідношення для всіх вузлів із конкретним k . Візуалізуючи це відношення, виявляємо кореляцію ступенів в графі. Зокрема, це виявляє, що для певних класів графів середній ступінь вузлів, прилеглих до i , залежить або не залежить від ступеня самого вузла i .

Це показано на рисунку 2.4 для графу ER з $n = 10000$ та $p = 0.1$.

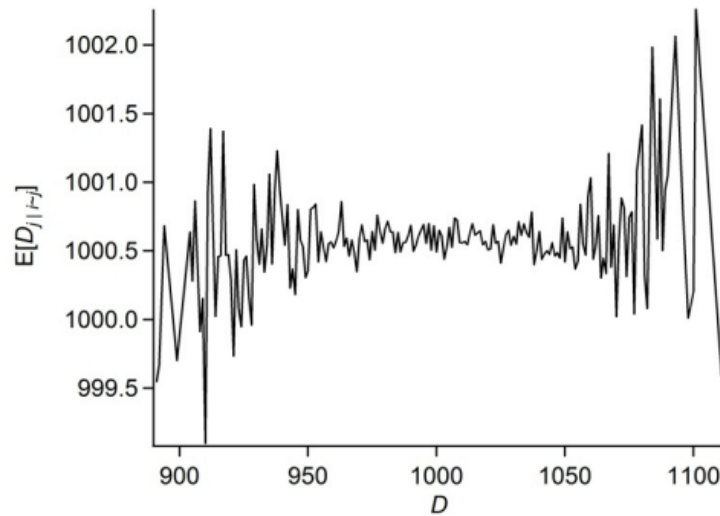


Рисунок 2.4 – Кореляція поєднаних вузлів у ER-мережі

У цьому прикладі середній ступінь вузлів j , що прилягають до вузла i , стабільний близько 1000, для різних k . Є сенс розглянути такі розподіли ступенів для графа:

- розподіл ступенів, $\Pr\{D = k\}$, визначаючи ймовірність випадково вибраного вузла мати ступінь k ;
- комбінований розподіл ступенів:

$$\begin{aligned} \Pr\{D_1 = j, D_2 = k\} &= \Pr\{D_1 = j | D_2 = k\} \cdot \Pr\{D_2 = k\} = \\ &= \Pr\{D_2 = k | D_1 = j\} \cdot \Pr\{D_1 = j\} \end{aligned} \quad (2.17)$$

визначаючи ймовірність випадково вибраної пари поєднаних вузлів мати ступені j та k відповідно. Припускаючи, що вибір другого вузла стохастично не залежить від вибору першого, комбінований розподіл ступенів (2.17) може бути записаний як $\Pr\{D_1 = j, D_2 = k\} = \Pr\{D_1 = j\} \cdot \Pr\{D_2 = k\}$;

- спільний розподіл ступенів, $\Pr\{D_{l-} = j, D_{l+} = k\}$, що визначає

ймовірність випадково вибраної ланки l мати ступінь j на одному кінці (позначається l_-) та ступінь k на іншому кінці (позначається l_+). Це відношення являє собою асортативність.

Можна виділити три типи мереж:

- нейтральна (неасортативна) мережа. Розподіл ступеня вузлів j , з'єднаних з випадково вибраним вузлом i , відбувається за тим же законом, що і розподіл ступенів у мережі в цілому;

- асортативна мережа. Для випадково вибраного вузла i буде концентрація високої спільної ступеня ймовірності для підключених вузлів j , що мають той самий або схожий ступінь, як i ;

- дизасортативна мережа. Для випадково вибраного вузла i буде концентрація високої ймовірності спільного ступеня для з'єднаних вузлів j , що мають ступінь, відмінну від ступеня i .

Оскільки ступінь розподілу мережі вважається метрикою першого порядку для характеристики мережі, то асортативність розглядається як метрика другого порядку.

Релевантність асортативності сильно пов'язана з діапазоном асортативності, $\rho_{\max} - \rho_{\min}$ мережі, зберігаючи при цьому розподіл ступенів цієї мережі. Діапазон ρ_D для розрідженої мережі виявляється більшим, ніж діапазон ρ_D для щільної мережі. Великий діапазон ρ_D передбачає більшу актуальність асортативності як метрики другого порядку для мережі, порівняно з малим діапазоном ρ_D .

Проаналізуємо асортативність мереж різних класів. Чотири класи графів, що розглядаються, це: випадковий граф Ердеша-Рені (ER), граф безмасштабної мережі Барабаші-Альберт (BA), граф "малого світа" Ваттса-Строгаца (WS) та модель зростання Каллавея (C).

2.3.1 Асортативність безмасштабних мереж Барабаші-Альберт

Мережа Барабаші-Альберт (ВА) будується з невеликої (наприклад, $N_0 = 3$) повної мережі. Вузли в початковому графу, а також вузли, які були додані в попередній частині процесу зростання, мають більш високу очікувану ступінь. Під час процесу зростання ці вузли мають більшу привабливість для нових вузлів, які потрібно додати, через більш високий ступінь цих вузлів.

Отже, ступінь цих вузлів, як правило, збільшується більше, ніж ступінь вузлів, які були додані пізніше в процесі росту. Інакше кажучи, коли один вузол через вибір довільних вузлів моделі зростання має вищий ступінь, ніж інші вузли в мережі, цей вузол, як правило, продовжує залучати більше нових з'єднань, ніж інші вузли.

Інтуїтивно зрозуміла думка [32] може полягати в тому, що графи ВА демонструватимуть дизасортативність. Причина полягає в тому, що кожного разу, коли додається новий зв'язок, він поєднує щойно доданий вузол n , який має ступінь $0 \leq d_n \leq m$ (m – кількість зв'язків, що додаються для кожного щойно доданого вузла), тобто низьку, та існуючий вузол, який, більш ймовірно є вузлом з відносно високим ступенем, що є результатом переважного поєднання.

Вочевидь, створення зв'язку між вузлом низького ступеня і вузлом високого ступеня має зробити граф у цілому більш дизасортативним. У таблиці 2.1 наведено діапазон асортативності для різних графів ВА.

Мережа побудована до розміру $n = 1000$, починаючи з повного графа з $3 \leq N_0 \leq 10$ та $2 \leq m \leq 5$. Асортативність обчислюється для $M = 10000$ примірників графа.

Таблиця 2.1 – Асортативність графа ВА з $n = 1000$ вузлів

Graph definition		Assortativity		
N_0	m	r_{min}	r_{max}	$r_{average}$
3	2	-0.147	-0.038	-0.092
4	2	-0.158	-0.038	-0.089
5	2	-0.135	-0.038	-0.084
10	2	-0.116	-0.006	-0.064
10	3	-0.093	-0.018	-0.055
10	5	-0.078	-0.008	-0.046

На рисунку 2.5 показана функція щільності розподілу показника асортативності для графів ВА.

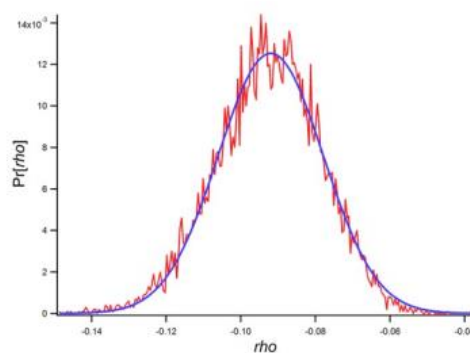
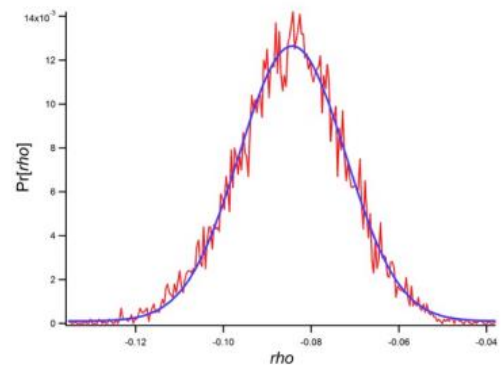
(a) $N=1000, N_0=3, m=2$ (b) $N=1000, N_0=5, m=2$

Рисунок 2.5 – Щільність розподілу показника асортативності мережі ВА

Згідно з рисунком 2.4 та таблицею 2.1, в цілому модельована мережа ВА є трішки дизасортативною, ρ_D розподілена приблизно між -0,10 та -0,08. Втім, подальші експерименти показують, що для великих значень розміру мережі (n) показник асортативності буде прагнути до нуля.

2.3.2 Асортативність мереж «малого світу»

Граф «малого світу» отримують шляхом застосування n кроків випадкового перенаправлення зв'язків на ланках кільцевої решітки. Однак принцип побудови графа малого світу може бути застосований і до інших регулярних класів графів, таких як квадратні решітки, кубічна решітка або сферична решітка. Граф WS, отриманий з квадратної решітки, може не імітувати малий світ (тому ми можемо не хотіти називати його «графом WS»). Граф WS, можливо, може бути сформований шляхом додавання довільних зв'язків, на відміну від випадкового перенаправлення зв'язків.

На рисунку 2.6 показано моделі графів WS, побудовані на базі решіток з $N = 100$ вузлів.

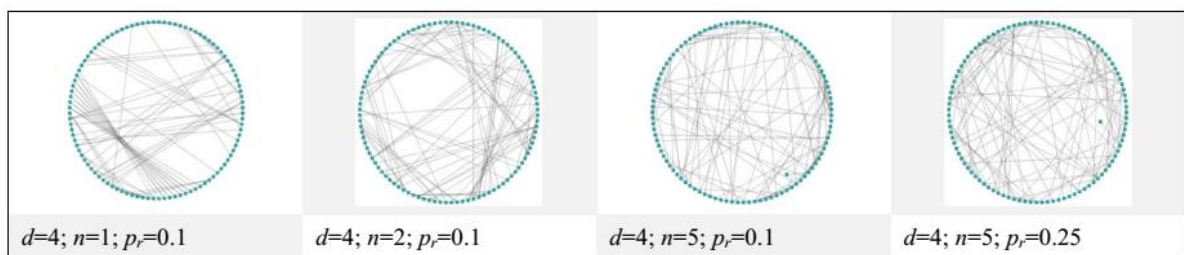


Рисунок 2.6 – Мережі «малого світу»

На рисунку 2.7 показано розподіл асортативності для двох моделей графів WS, показаних на рисунку 2.6, з тією різницею, що $N = 1000$. Розподіл асортативності отримують шляхом генерації кожної моделі графа в $M = 10000$ разів з однаковими характеристичними параметрами (N, d, n, p_r). Слід зазначити, що асортативність початкової кільцевої решітки дорівнює одиниці, оскільки всі вузли мають однаковий ступінь.

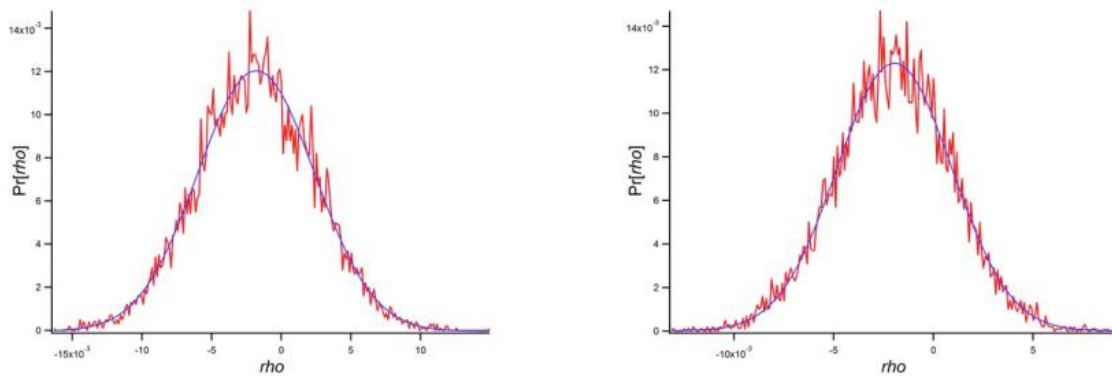


Рисунок 2.7 – Асортативність мереж "малого світу"

Рисунок 2.7 ілюструє, що ітеративний процес перенаправлення робить мережу все більш дизасортативною. Цього можна очікувати, оскільки кожен крок перемотування «розбиває» ланцюг вузлів, однаково з'єднаних із визначеною кількістю відповідних сусідніх вузлів. Етапи перенаправлення зв'язків, очевидно, порушують це. Зі збільшенням кількості кроків повторного перенаправлення зв'язків мережа стає випадковою, тому ρ_D наближається до нуля.

2.3.3 Асортативність моделі випадкового зростання Каллавея

Модель Каллавея [3] входить до класу «моделей зростання». Для кожного щойно доданого вузла додається посилення з ймовірністю δ між двома випадковим чином обраними вузлами (вибраними рівномірно з усіх вузлів, включаючи щойно доданий вузол), за умови, що між двома вибраними вузлами ще немає зв'язку. Модель Callaway відрізняється від моделі ER тим, що вузли, які додаються пізніше, мають менший шанс бути вибраними один раз або кілька разів для створення зв'язку з цього вузла, до іншого вузла.

Отже, ці вузли в середньому матимуть нижчий ступінь. Модель росту Callaway також відрізняється від моделі зростання BA, оскільки для моделі

росту ВА завжди створюються нові зв'язки між щойно доданим вузлом та існуючим вузлом. Однак модель росту Callaway може створити мережу з ізольованими вузлами. Подальша відмінність між моделлю росту Callaway і моделлю росту ВА полягає в тому, що для моделі Callaway ступінь вузлів не впливає на їх вірогідність залучення більшої кількості посилок.

Отже, в додатку немає переваг. Обґрунтування моделі росту Callaway при порівнянні з моделлю Ердеша-Рені полягає в тому, що ступінь вузла стає функцією віку цього вузла.

Callaway та ін. [3] зауважте, що модель їх росту в мережі проявляє виразну ступінчасту кореляцію як функцію ступеня вузла. В [3] показано, що існує позитивний коефіцієнт кореляції для з'єднання вузлів з іншими вузлами аналогічного ступеня. Це суттєво відрізняється від ER випадкового графа, згідно з яким ступінь кореляції в середньому 0.

Позитивна ступінчаста кореляція є в графах Callaway, особливо переважно для вузлів вищого ступеня. Вузли високого ступеня, як правило, підключаються до інших вузлів високого ступеня. Для вузлів нижчого ступеня ступінь з'єднаних вузлів більш рівномірно розподіляється між вузлами низького ступеня та високого. Показник асортативності цієї мережі асимптотично дорівнює

$$\rho = \frac{\delta}{1 + 2\delta}. \quad (2.18)$$

Результати чисельного експерименту з обчислення коефіцієнта асортативності мережі Каллавея показані на рисунку 2.8.

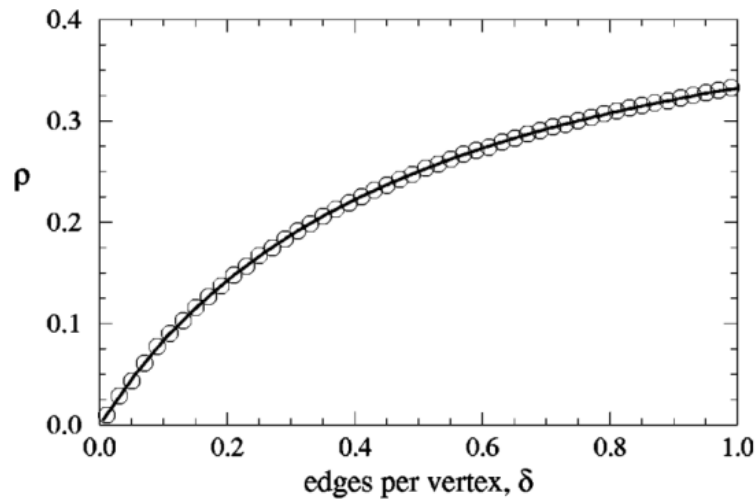


Рисунок 2.8 – Асортативність мережі Коллоєвої як функція щільності

2.4 Процедура перенаправлення зв'язків

У попередніх розділах показано, що асортативність мережі являє собою специфічний структурний аспект цієї мережі і що є бажання впливати на цей аспект мережі. Наприклад, ми можемо захотіти збільшити спектральний радіус або збільшити алгебраїчну зв'язність графу. У цьому розділі буде показано, що ці спектральні значення можна змінити, змінивши асортативність графа. В [61] показано, що збільшення асортативності мережі робить її більш надійною проти усунення вузлів, але в той же час робить мережу менш стабільною. Однак далі спостерігається, що для взаємозалежних мереж збільшення асортативності зменшує стійкість. Як правило, змінюючи асортативність мережі через топологічні зміни в цій мережі, ми змінюємо й інші показники цієї мережі, наприклад, ефективний опір графа.

Незважаючи на те, що асортативні мережі більш схильні до поширення збоїв, для цих мереж потрібен більший час інтервенції для запобігання подальших збоїв (поширення епідемії). Асортативність мережі стосується як стійкості мережі до відмови вузла, так і динамічної поведінки мережі щодо

розповсюдження відмов. Існує поняття оболонки стійкості; залежність між асортативністю та стійкістю до відмови вузла залежить від типу відмови. Помірне збільшення асортативності збільшує стійкість мережі до націлених нападів вузлів, тоді як помірне зниження асортативності збільшує стійкість мережі проти випадкових рівномірних атак вузлів (або поломки вузла). Зміна топології мережі може приймати різні форми. Найпростішою та найпоширенішою з них є процедура перенаправлення зв'язків (рис. 2.9).

Обґрунтування цієї процедури полягає в тому, що послідовність ступенів (метрика графа 1-го порядку) не впливає, в той час як атрибутивність або інша метрика графа 2-го порядку змінюється. Вплив на асортативність за допомогою модифікації мережі сильно залежить від виду модифікації, від ланок та вузлів, що беруть участь у модифікації, та від класу мережі. Модифікація топології вплине також на інші показники мережі, ніж асортативність. Наприклад, видалення зв'язку збільшить середній найкоротший шлях у мережі або може відключити граф; це також збільшить ефективний опір графа. Видалення посилення, додавання посилення та перенаправлення посилення (якщо не пара) впливатимуть на ступінь вектора графа.

Отже, зміни топології мережі на практиці будуть обмежені певними обмеженнями. Перенаправлення зв'язків тягне за собою те, що мережу перемотують таким чином, щоб вектор ступеня $d = (d_1, d_2, \dots, d_N)$ зберігався. Перенаправлення зв'язків має практичну характеристику, що вузол зберігає кількість транспортних з'єднань. Один видатний приклад складається з IP-маршрутизаторів, які мають певну кількість інтерфейсів. IP-підмережа може бути підключена, зберігаючи кількість кабелів на маршрутизаторі постійною. Залишаючи векторний ступінь незмінним, ми можемо впливати на асоціативність мережі через перенаправлення зв'язків.

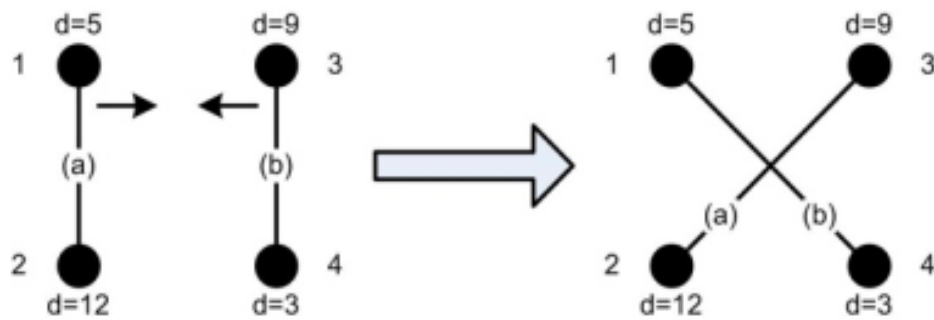


Рисунок 2.9 – Приклад перенаправлення зв'язків

Перенаправлення зв'язків, зображене на рисунку 2.9, відноситься до графів з непрямыми, незваженими зв'язками. Шляхом повторного підключення вузла 2 до вузла 3 та вузла 4 до вузла 1 обидва вузли 2 та 3 стають більш асортативними, оскільки тепер вони з'єднані з вузлами (3 та 2 відповідно), які мають ступінь, наближений до власного. Вузли 1 і 4 також стають більш асортативними, оскільки вони поєднуються з подібними собі. Основа зміни асортативності завдяки перенаправленню зв'язків ґрунтується на зміні показника N_3 з (2.7):

$$\Delta N_3 = d_1 d_3 + d_2 d_4 - d_1 d_2 - d_3 d_4 = (d_1 - d_4)(d_3 - d_2). \quad (2.19)$$

Можуть бути розроблені різні стратегії зміни асортативності мережі до бажаного стану за якомога меншу кількість кроків перенаправлення зв'язків. Крім того, стратегія перенаправлення зв'язків повинна бути такою, щоб пара зв'язків, яка підходить для повторного з'єднання, була знайдена за якнайменшу кількість спроб.

2.5 Моделювання мереж з переважним приєднанням

Найбільш відома з моделей мережі з переважним приєднанням була створена в 1999р. Барабаші та Альберт (BA-tree).

Мережа ВА задається наступними правилами побудови:

- мережа є неорієнтованою;
- спочатку мережа складається з $m_0 \geq m$ вузлів і деякої кількості e_0 зв'язків (ребер);
- кожен новий вузол з'єднується з наявними фіксованим числом зв'язків (m);
- ймовірність p_i того, що новий вузол приєднується до існуючого (i) пропорційна кількості зв'язків у цього вузла (k_i):

–

$$p_i = \frac{k_i}{\sum k_i}. \quad (2.20)$$

Таким чином, мережа ВА відрізняється від розглянутої раніше мережі з рівноймовірним приєднанням (ER) тільки правилом приєднання вузлів: замість рівноймовірного приєднання використовується переважне приєднання (preferential attachment) (2.20). До речі кажучи, правило (2.20) теж можна назвати «рівноймовірним приєднанням», але в іншому сенсі: рівноймовірно обираються не вузли, а зв'язки. Різниця між правилами приєднання в рівноймовірній мережі і в мережі ВА показана на рисунку 2.10.

У 1999 році А.-Л. Барабаш (Albert-Laszlo Barabasi) разом зі своєю аспіранткою Р. Альберт (Reka Albert) вивчали властивості реальних мереж і виявили, для багатьох реальних мереж розподіл вузлів по числу зв'язків підпорядковується степеневим законом Парето (точніше кажучи, закону Юла-Саймона).

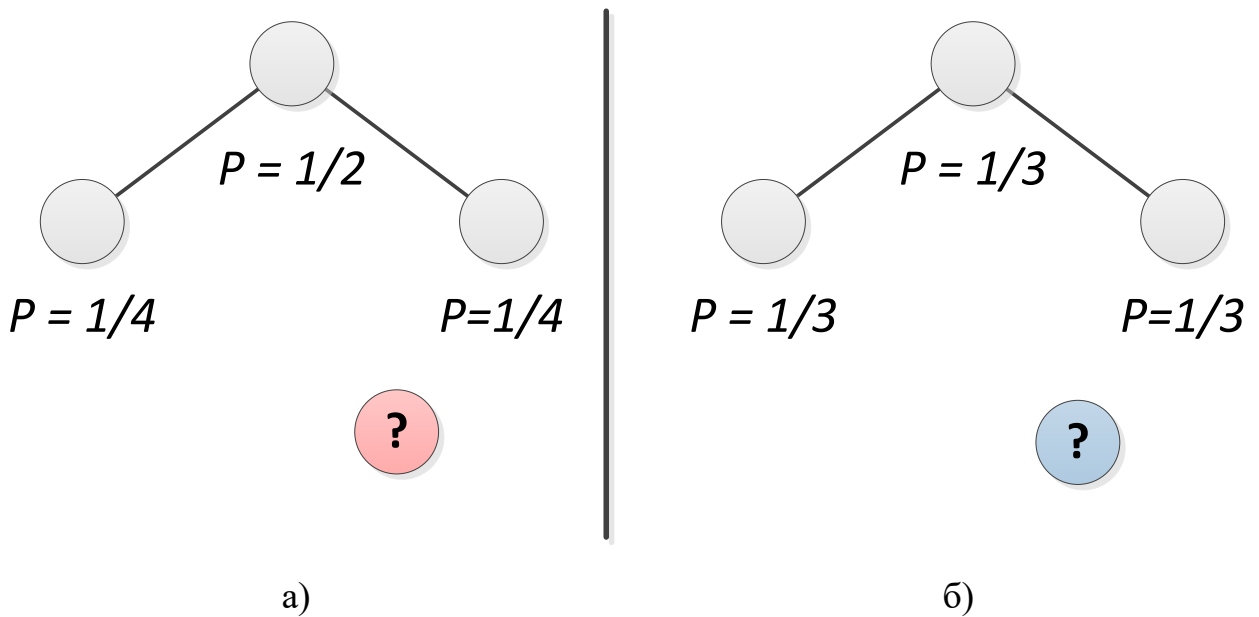


Рисунок 2.10 – Правила приєднання вузлів в мережі ВА (а) і в мережі рівноймовірного приєднання (б)

Іншими словами, у багатьох реальних мережах невелике число вузлів (хабів) мають дуже багато зв'язків, а величезна кількість вузлів містить лише кілька зв'язків. Такі мережі отримали назву безмасштабних мереж (scale free networks). Поняття безмасштабності найпростіше проілюструвати знаменитим правилом Парето «20/80»: нехай в деякій мережі 80% всіх зв'язків припадає на 20% вузлів. Якщо з мережі виключити ці 20% вузлів (і інцидентні їм зв'язку), то в що залишилася мережі буде дотримуватися та ж пропорція: 80%, що залишилися зв'язків будуть контролюватися 20%, що залишилися вузлів. Для правила «20/80» показник розподілу (показник скейлінга) становить $\theta \approx 2.16$.

Для знаходження розподілу вузлів по числу зв'язків в мережі ВА застосовується метод балансових рівнянь. Вони мають вигляд:

$$N_k(n+1) - N_k(n) = m \left(\frac{(k-1)N_{k-1}(n) - kN_k(n)}{n \cdot l(n)} \right), \quad k > m, \quad (2.21)$$

$$N_m(n+1) - N_m(n) = 1 - m \frac{m N_m(n)}{n \cdot l(n)}, \quad (2.22)$$

де $l(n)$ – середня кількість зв'язків в мережі, що складається з n вузлів, і оскільки кожен новий вузол приносить в мережу рівно $2m$ зв'язків, то незалежно від конфігурації початкового графа- «затравки»,

$$l(n) \underset{n \rightarrow \infty}{=} 2m. \quad (2.23)$$

Як і раніше, припускаємо, що при $n \rightarrow \infty$ ймовірності розподілу вузлів за ступенями сходяться до стаціонарних значень. Тоді з (2.21)-(2.23) випливає, що

$$p_k = \frac{1}{2}((k-1)p_{k-1} - k p_k), \quad k > m, \quad (2.24)$$

$$p_m = 1 - \frac{1}{2}m p_m. \quad (2.25)$$

Рішення системи (2.24)-(2.25) має вигляд

$$p_k = \frac{2m(m+1)}{k(k+1)(k+2)}, \quad (k \geq m). \quad (2.26)$$

Можна відзначити, що (2.26) є окремим випадком закону розподілу Юла-Саймона $YS(\theta, k_{\min})$ при значеннях його параметрів

$$\theta = 3, \quad k_{\min} = m. \quad (2.27)$$

Середнє число зв'язків у вузла мережі ВА таке ж, як і у ER-мережі; власне кажучи, воно визначається процедурою побудови мережі і також становить $\langle k \rangle = 2m$.

У той же час, концентрація зв'язків у мережі БА набагато вище: математичне очікування максимального числа зв'язків у вузла становить

$$N_{\max} \approx m \cdot \sqrt{n}, \quad (2.28)$$

що набагато вище, ніж для ER-мережі.

Приклад безмасштабної мережі ВА для $m=1$ (тобто дерева) показано на рисунку 2.11.

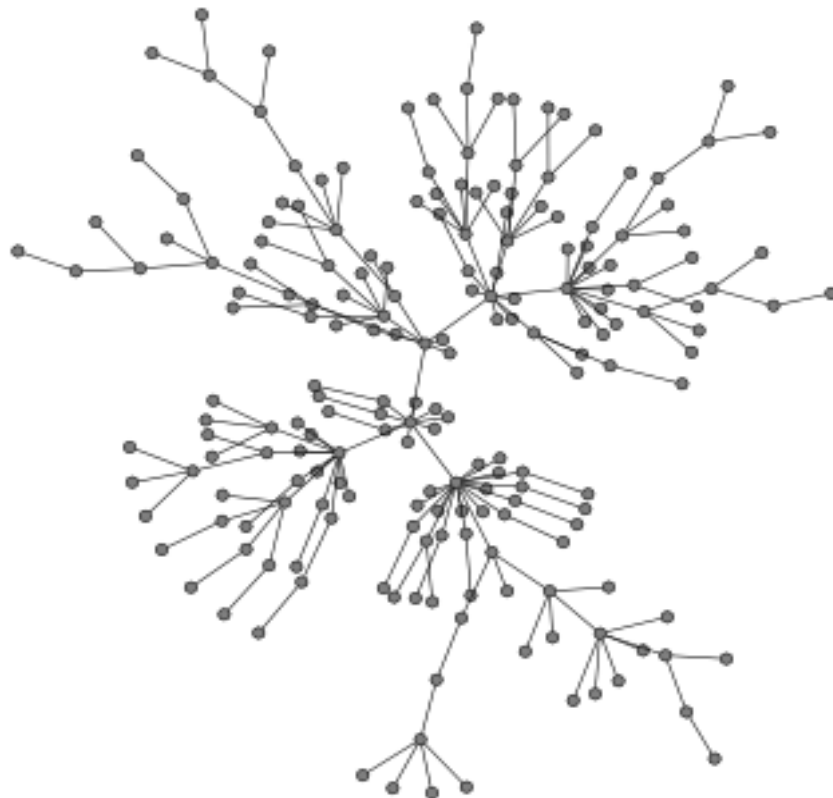


Рисунок 2.11 – Приклад безмасштабної мережі (дерева Барабаші-Альберт)

Близькою до моделі мережі ВА є модель Прайса, запропонована в 70-х роках XX століття для моделювання мережі цитування наукових робіт. Ця модель описує динаміку спрямованої (орієнтованої) мережі. В цьому випадку вузлу, що додається, відповідає m нових зв'язків (цитувань).

Імовірність посилення на i -тий вузол залежить від кількості наявних посилень на нього (k_i) і містить додатковий параметр $c > 0$, званий прихованою (латентною) привабливістю:

$$p_i = \frac{k_i + c}{\sum (k_i + c)}. \quad (2.29)$$

Введення прихованої привабливості обумовлено тим, що за її відсутності (тобто при використанні правила переважного приєднання в формі (2.20)) вийде, що ймовірність цитування статті пропорційна числу вже наявних посилень на неї, і тоді нова стаття (для якої $k = 0$) взагалі не буде ніколи цитуватися.

Оскільки кожен новий вузол додає в мережу m посилень, то знаменник виразу (2.29) дорівнює

$$\sum_{i=1}^n (k_i + c) = n(m + c). \quad (2.30)$$

Балансові рівняння щодо ступенів зв'язків вузлів p_k будуються аналогічно до (2.21)-(2.25) та матимуть вигляд

$$p_k = \frac{m}{m + c} ((k - 1 + c) p_{k-1} - (k + c) p_k), \quad k > m, \quad (2.31)$$

$$p_m = 1 - m p_m. \quad (2.32)$$

Рішення системи (2.31)-(2.32) має вигляд

$$p_k = \frac{B(k+c, \theta)}{B(c, \theta-1)}. \quad (2.33)$$

Вираз (2.33) також є окремим випадком закону розподілу Юла-Саймона $YS(\theta, k_{\min})$ при значеннях його параметрів

$$\theta = 2 + c/m, \quad k_{\min} = 0. \quad (2.34)$$

Ще однією моделлю безмасштабної мережі є модель Юла. Спочатку Юл розробив свою модель в 20-х роках XX століття для опису статистики біологічних таксонів квіткових рослин. Пізніше ця модель була успішно адаптована та узагальнена дослідниками для пояснення статечних законів у багатьох інших системах.

Відомо, що біологічний рід складається з видів. Нові види утворюються в результаті спеціалізації, тобто поділу одного виду на два. Юл припустив, що видоутворення відбувається з якоюсь стохастически постійною частотою, тобто кожен вид (в рамках одного роду, або отряду) має одні й ті ж шанси розділитися на два за якусь одиницю часу. Це означає, що ймовірність появи нового виду в роді, що складається з k видів, пропорційна k , тобто визначається правилом переважного приєднання (2.20).

Юл також постулював, що після в середньому m видоутворень в існуючих родах, утворюється новий рід, що складається з одного виду. Таким чином, на кожен новий рід виникає $m+1$ вид, і в середньому є $m+1$ видів в одному роді.

Відомо, що моделі Юла і Барабаші-Альберт збігаються при $m=1$ та породжують деревоподібну мережу (рис.2.11).

3 ОБЧИСЛЕННЯ ПОКАЗНИКА АСОРТАТИВНОСТІ В БЕЗМАСШТАБНИХ МЕРЕЖАХ

3.1 Вибір програмних засобів реалізації

У якості програмного засоба / середовища реалізації елементів системи моделювання еластичних мереж та оцінювання меж показника асортативності було обрано середовище MatLab [17], [18].

MatLab (скорочення від англ. «Matrix Laboratory – пакет прикладних програм для вирішення задач технічних обчислень, який має вбудовану мову програмування. MatLab використовують більше одного мільйона інженерних і наукових працівників, він працює на більшості сучасних операційних систем, включаючи Linux, Mac OS, Microsoft Windows.

MatLab як мова програмування був розроблений Клівом Моулером в кінці 1970-х років коли він був деканом факультету комп'ютерних наук в Університеті Нью-Мексико. Метою розробки було надати студентам факультету можливість використання програмних бібліотек Linpack і Eispack без необхідності вивчення Фортрана. Незабаром нова мова поширилася серед інших університетів і була з великим інтересом зустрінута вченими, які працюють в області прикладної математики. До сих пір в Інтернеті можна знайти версію 1982 року, написану на Фортрані, поширювану з відкритим вихідним кодом. Інженер Джон Літл познайомився з цією мовою під час візиту Кліва Моулера в Стенфордський університет в 1983 році. Зрозумівши, що нова мова має великий комерційний потенціал, він об'єднався з Клівом Моулером і Стівом Бангертом. Спільними зусиллями вони переписали MatLab на C і заснували в 1984 компанію The MathWorks для подальшого розвитку. Ці переписані на C бібліотеки довгий час були відомі під ім'ям JASCRAS. Спочатку MatLab призначався для проектування систем управління (основна спеціальність Джона Літла), але швидко завоював популярність у багатьох інших наукових і інженерних областях. Він також

широко використовувався і в освіті, зокрема, для викладання лінійної алгебри та чисельних методів.

Мова MatLab є високорівневою мовою програмування, яка інтерпретується. Середовище MatLab підтримує засновані на матрицях структури даних, широкий спектр функцій, інтегроване середовище розробки, об'єктно-орієнтовані можливості і інтерфейси до програм, написаних на інших мовах програмування.

Програми, написані на MatLab, бувають двох типів – функції і скрипти. Функції мають вхідні і вихідні аргументи, а також власний робочий простір для зберігання проміжних результатів обчислень і змінних. Скрипти ж використовують загальний робочий простір. Як скрипти, так і функції не компілюються в машинний код, а зберігаються у вигляді текстових файлів. Існує також можливість зберігати так звані pre-parsed програми – функції і скрипти, оброблені в вид, зручний для машинного виконання. У загальному випадку такі програми виконуються швидше звичайних, особливо якщо функція містить команди побудови графіків.

Основною особливістю мови MatLab є її широкі можливості щодо роботи з матрицями, які творці мови висловили в гаслі «думай векторно» (Think vectorized).

MatLab надає користувачеві велику кількість (кілька сотень) функцій для аналізу даних, які покривають майже всі області математики, зокрема:

- матриці і лінійна алгебра – алгебра матриць, лінійні рівняння, власні значення і вектора, факторизація матриць та інші;
- багаточлени і інтерполяція – корені многочленів, операції над багаточленами і їх диференціювання, інтерполяція і екстраполяція кривих і інші;
- математична статистика і аналіз даних – статистичні функції, статистична регресія, цифрова фільтрація, швидке перетворення Фур'є та інші;
- обробка даних – набір спеціальних функцій, включаючи побудову графіків, оптимізацію, пошук нулів, чисельне інтегрування та інші;

- диференціальні рівняння – рішення диференціальних і диференціально-алгебраїчних рівнянь, диференціальних рівнянь із запізненням, рівнянь з обмеженнями, рівнянь в часткових похідних і інші;
- розріджені матриці – спеціальний клас даних пакету MatLab, що використовується в спеціалізованих додатках;
- цілочисельна арифметика – виконання операцій цілочисельний арифметики в середовищі MatLab.

MatLab надає зручні засоби для розробки алгоритмів, включаючи високорівневі з використанням концепцій об'єктно-орієнтованого та функціонального програмування. У ньому є всі необхідні засоби інтегрованого середовища розробки, включаючи відлагоджувач та профайлер.

У складі пакету MatLab є велика кількість функцій для побудови графіків, в тому числі тривимірних, візуального аналізу даних і створення анімованих роликів. Вбудована середовище розробки дозволяє створювати графічні інтерфейси користувача з різними елементами управління, такими як кнопки, поля введення і іншими.

Програми MatLab, як консольні, так і з графічним інтерфейсом користувача, можуть бути зібрані за допомогою компоненти MatLab Compiler в незалежні від MatLab виконувані програми або динамічні бібліотеки, для запуску яких на інших комп'ютерах, проте, потрібна установка вільно розповсюджуваного середовища MatLab Compiler Runtime (MCR).

Пакет MatLab включає різні інтерфейси для отримання доступу до зовнішніх підпрограм, написаних на інших мовах програмування, даних. Підтримуються технології Component Object Model (COM) та Dynamic Data Exchange (DDE). Багато з цих можливостей відомі під назвою MATLAB API.

Пакет MatLab в Microsoft Windows надає доступ до програмній платформі .NET Framework. Є можливість завантажувати .NET збірки (Assemblies) і працювати з об'єктами .NET класів з середовища MatLab. У версії MatLab 7 підтримується .NET Framework версій 2.0, 3.0, 3.5 і 4.0.

Пакет MatLab містить функції, які дозволяють йому отримувати доступ

до інших додатків середовища Windows, так само як і цим програмам отримувати доступ до даних MatLab, за допомогою технології динамічного обміну даними (DDE).

В MatLab існує можливість викликати методи веб-сервісів. Спеціальна функція створює клас, ґрунтуючись на методах API веб-сервісу.

MatLab взаємодіє з клієнтом веб-сервісу за допомогою прийняття від нього посилань, їх обробки і посилань відповіді. Підтримуються наступні технології: Simple Object Access Protocol (SOAP) і Web Services Description Language (WSDL).

Пакет MatLab включає інтерфейс взаємодії з зовнішніми додатками, написаними на мовах С і Фортран. Здійснюється це взаємодія через MEX-файли. Існує можливість виклику підпрограм, написаних на С або Фортрані з MatLab, як ніби це вбудовані функції пакета. MEX-файли являють собою спільні бібліотеки, які можуть бути завантажені і виконані інтерпретатором, вбудованим в MatLab. MEX-процедури мають також можливість викликати вбудовані команди MatLab.

Інтерфейс середовища MatLab представлений на рисунку 3.1.

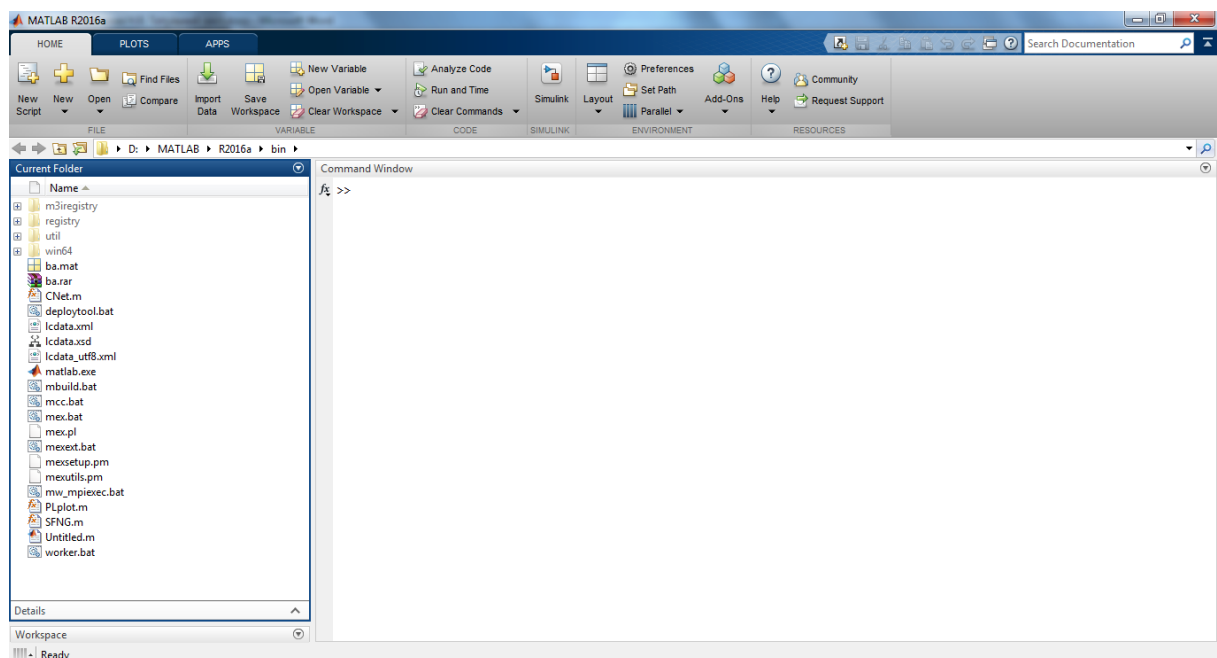


Рисунок 3.1 – Інтерфейс середовища MatLab

Інтерфейс MatLab дозволяє викликати функції, що знаходяться в звичайних DLL-бібліотеках, прямо з MatLab. Ці функції повинні мати С-інтерфейс.

Крім того, в MatLab є можливість отримати доступ до його вбудованих функцій через С-інтерфейс, що дозволяє використовувати функції пакета в зовнішніх додатках, написаних на С. Ця технологія в MatLab називається С Engine.

Для MatLab є можливість створювати спеціальні набори інструментів (toolbox), що розширюють його функціональність. Набори інструментів є колекції функцій, написаних на мові MatLab для вирішення певного класу задач. Компанія Mathworks поставляє набори інструментів, які використовуються в багатьох областях, включаючи наступні:

- цифрова обробка сигналів, зображень та даних: DSP Toolbox, Image Processing Toolbox, Wavelet Toolbox, Communication Toolbox, Filter Design Toolbox – набори функцій, що дозволяють вирішувати широкий спектр завдань обробки сигналів, зображень, проектування цифрових фільтрів і систем зв'язку;

- системи управління: Control Systems Toolbox, μ -Analysis and Synthesis Toolbox, Robust Control Toolbox, System Identification Toolbox, LMI Control Toolbox, Model Predictive Control Toolbox, Model-Based Calibration Toolbox – набори функцій, що полегшують аналіз і синтез динамічних систем, проектування, моделювання та ідентифікацію систем управління, включаючи сучасні алгоритми управління, такі як Робастне управління, H_{∞} -управління, μ -синтез та інші;

- фінансовий аналіз: GARCH Toolbox, Fixed-Income Toolbox, Financial Time Series Toolbox, Financial Derivatives Toolbox, Financial Toolbox, Datafeed Toolbox – набори функцій, що дозволяють швидко і ефективно збирати, обробляти і передавати різну фінансову інформацію;

- аналіз і синтез географічних карт, включаючи тривимірні: Mapping Toolbox;

– збір і аналіз експериментальних даних: Data Acquisition Toolbox, Image Acquisition Toolbox, Instrument Control Toolbox, Link for Code Composer Studio – набори функцій, що дозволяють зберігати і обробляти дані, отримані в ході експериментів, в тому числі в реальному часі. Підтримується широкий спектр наукового і інженерного вимірювального обладнання;

– візуалізація і уявлення даних: Virtual Reality Toolbox дозволяє створювати інтерактивні світи і візуалізувати наукову інформацію за допомогою технологій віртуальної реальності і мови VRML;

– засоби розробки: MatLab Builder for COM, MatLab Builder for Excel, MatLab Builder for NET, MatLab Compiler, Filter Design HDL Coder – набори функцій, що дозволяють створювати незалежні програми з середовища MatLab;

– взаємодія з зовнішніми програмними продуктами: MatLab Report Generator, Excel Link, Database Toolbox, MatLab Web Server, Link for ModelSim – набори функцій, що дозволяють зберігати дані в різних видів таким чином, щоб інші програми могли з ними працювати;

– бази даних: Database Toolbox - інструменти роботи з базами даних;

– наукові та математичні пакети: Bioinformatics Toolbox, Curve Fitting Toolbox, Fixed-Point Toolbox, Fuzzy Logic Toolbox, Genetic Algorithm and Direct Search Toolbox, OPC Toolbox, Optimization Toolbox, Partial Differential Equation Toolbox, Spline Toolbox, Statistic Toolbox, RF Toolbox – набори спеціалізованих математичних функцій, що дозволяють вирішувати широкий спектр наукових і інженерних задач, включаючи розробку генетичних алгоритмів, вирішення завдань в приватних похідних, цілочисельні проблеми, оптимізацію систем та інші;

– нейронні мережі: Neural Network Toolbox – інструменти для синтезу та аналізу нейронних мереж;

– нечітка логіка: Fuzzy Logic Toolbox – інструменти для побудови і аналізу нечітких множин;

– символічні обчислення: Symbolic Math Toolbox – інструменти для

символьних обчислень з можливістю взаємодії з символьним процесором програми Maple.

Крім перерахованих вище, існують тисячі інших наборів інструментів для MatLab, написаних іншими (сторонніми) розробниками.

І все-таки в першу чергу MatLab – це засіб математичного моделювання, який забезпечує проведення досліджень практично у всіх відомих областях науки і техніки. При цьому структура пакету дозволяє ефективно поєднувати обидва основні підходи до створення моделі: аналітичний і імітаційний.

Саме в сфері математичного моделювання MatLab дозволяє найбільш повно використовувати всі сучасні досягнення комп'ютерних технологій, в тому числі засоби візуалізації і аудіфікації (озвучування) даних, а також можливості обміну даними через Інтернет. Крім того, користувач має можливість створювати засобами MatLab власний графічний інтерфейс, який відповідає як його смакам, так і вимогам розв'язуваної задачі. Як впливає з назви пакета, він орієнтований в першу чергу на обробку масивів даних (матриць і векторів). Це дозволило його розробникам істотно підвищити ефективність процедур, що працюють із зазначеними типами даних, в порівнянні з мовами програмування «загального призначення».

З точки зору користувача, MatLab являє собою багатющу бібліотеку функцій (в MatLab 2015 їх більше 1000), єдина проблема при роботі з якою полягає в умінні швидко відшукати ті з них, які потрібні для вирішення даного завдання [12].

3.2 Генерація мереж Барабаші-Альберт

Модель БА являє собою випадковий динамічний граф, який вирощується з невеликого графа-затравки шляхом необмежено повторюваних кроків додавання до графа нової вершини з n ребрами. Таким чином, вузол тим швидше накопичує нові зв'язки, ніж їх у нього більше

На підставі проведеного аналізу був розроблений алгоритм генерації мережі Барабаш-Альберт, представлений на рисунку 3.2.

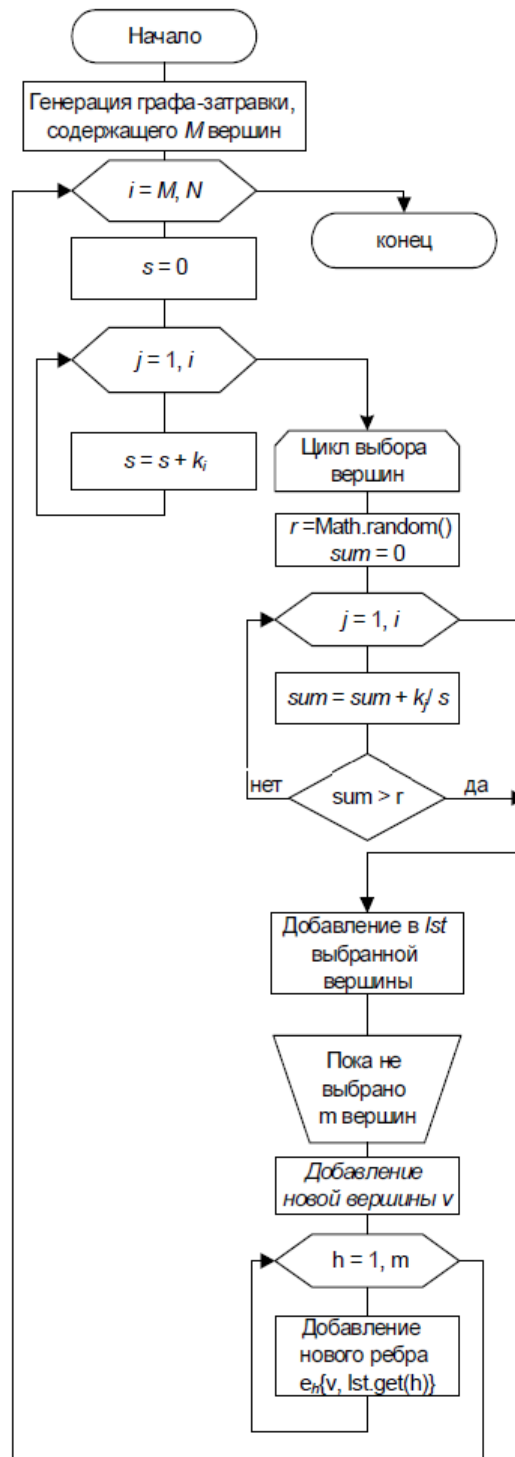


Рисунок 3.2 – Блок-схема алгоритма генерації моделі Барабаші-Альберт

Розроблена програма реалізує генерацію і візуалізацію моделі безмасштабної мережі Барабаш-Альберт. Вона складається з трьох функцій.

Функція SFNG використовується для імітації алгоритму БА і повертає безпосередньо БА-мережу заданого розміру. Цей код був розроблений таким чином, що можна створити мережу малого розміру, а потім використовувати цю мережу в якості графа-затравки для створення мережі більшого розміру, продовжуючи цей процес до тих пір, поки не буде досягнутий фактичний бажаний розмір мережі. Необхідно розуміти, що початковий граф-затравка не зобов'язаний мати властивостей безмасштабної мережі, в той час як з ростом розміру граф наближається до масштабно-інваріантного. Тому розумно не робити розмір початкового графа набагато більше, ніж кілька вузлів (найчастіше 3-10 взаємопов'язаних вузлів).

Функція CNet створює мережевий графік, використовуючи функцію gplot з круговими координатами. Він дозволяє отримати просту, але інтуїтивно зрозумілу візуалізацію мережі (рис. 3.3).

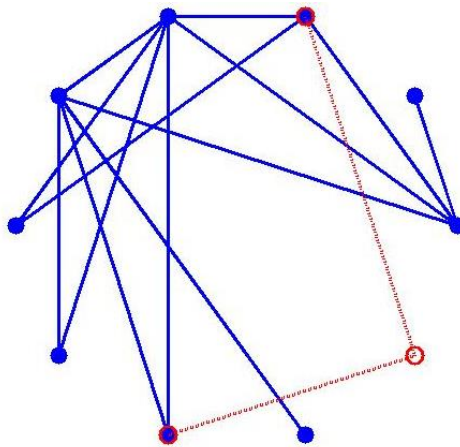


Рисунок 3.3 – Приклад згенерованої мережі

Функція PLplot приймає безмасштабну мережу в форматі матриці суміжності, візуалізує розподіл вузлів та ступенями та лінію, що апроксимує його до ступеневого розподілу (рис. 3.4). Функція повертає рівняння апроксимуючої лінії в змінній `cfit`.

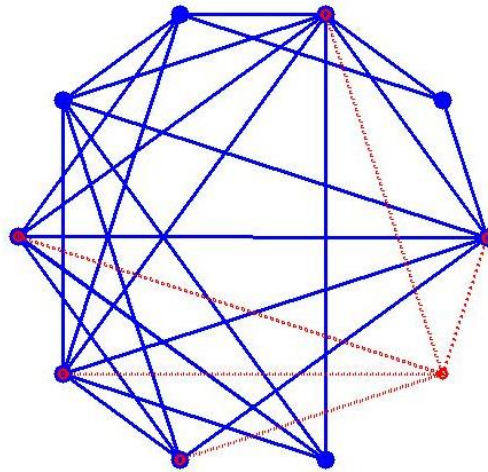


Рисунок 3.4 – Подальше генерування мережі

Під час завдання графа-затравки (seed) необхідно враховувати, що граф задається матрицею суміжності, яка повинна бути симетричною.

Приклад виклику описаних функцій (якому відповідають рис.3.3 та рис. 3.4) наведений на рисунку 3.5.

```
seed=[0 1 0 0 1;1 0 0 1 0;0 0 0 1 0;0 1 1 0 0;1 0 0 0 0]
Net = SFNG(50, 1, seed);
PL_Equation = PLplot(Net)
CNet(Net)
```

Рисунок 3.5 – Приклад виклику функцій

3.3 Розрахунок та модифікація асортативності мережі Барабаші-Альберт

Розрахунок коефіцієнта асортативності мережі Барабаші-Альберт проводився згідно з (2.6)-(2.7).

Модифікація асортативності згенерованої мережі проводиться шляхом перенаправлення зв'язків (рис.2.8). Зв'язки обираються випадково. Було передбачено три варіанти цього перенаправлення:

– нейтральне: процедура перенаправлення проводиться у циклі з заданою кількістю повторень;

– асортативне: процедура перенаправлення проводиться у циклі з заданою кількістю вдалих повторень. Перенаправлення вважається вдалим, якщо функція (2.19) позитивна;

– дизасортативне: процедура перенаправлення проводиться у циклі з заданою кількістю вдалих повторень. Перенаправлення вважається вдалим, якщо функція (2.19) негативна.

Задачі генерації мережі, розрахунку асортативності та перенаправлення зв'язків є взаємопов'язані, тому було розроблено єдиний додаток з графічним інтерфейсом (GUI MATLAB), в якому ці задачі реалізовані.

Скріншот вікна наведений на рисунку 3.6, а код програми- у додатку А.

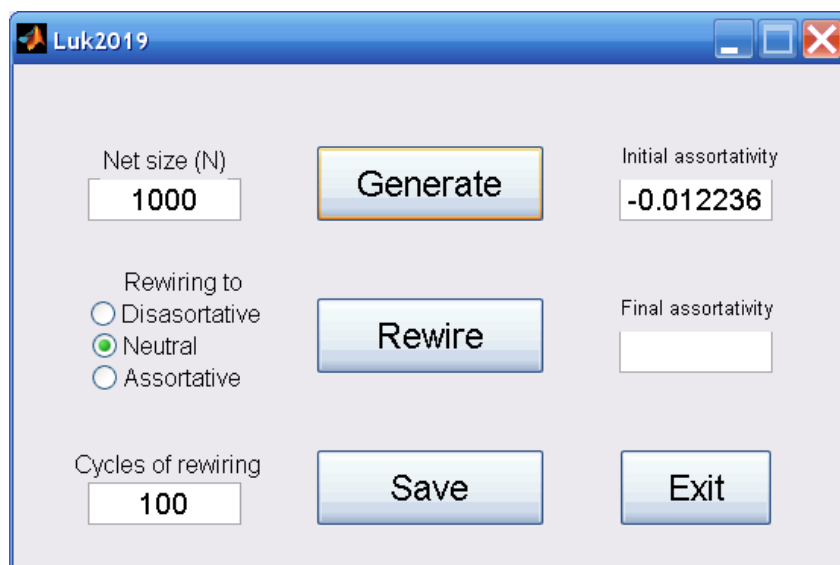


Рисунок 3.6 – Скріншот вікна додатку

ВИСНОВКИ

Проведено аналіз об'єкту досліджень – складних мереж. Проведена порівняльна характеристика існуючих мереж і їх характеристик. На основі існуючих досліджень була спроектована і змодельована імітаційна модель мережі Барабаші-Альберт.

Досліджено поняття асортативності, його числові характеристики, види, форми, способи їх обчислення, асортативність різних моделей мереж.

Розроблено алгоритм формування мережі Барабаші-Альберт та перенаправлення зв'язків задля досягнення заданої асортативності.

Здійснено програмну реалізацію генерації мережі Барабаші-Альберт, розрахунку асортативності та перенаправлення зв'язків у додатку з графічним інтерфейсом. Проведене пробне тестування підтвердило працездатність програми.

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ

1. Ландэ Д.В., Снарский А.А. Моделирование сложных сетей: учебное пособие. К. : НТУУ «КПИ», 2015. 212 с.
2. Dorogovtsev S.N., Mendes J.F.F. Evolution of Networks: From Biological Networks to the Internet and WWW. Oxford, USA : Oxford University Press, 2003. 280 p.
3. Newman M., Barabasi A.-L., Watts D.J. The Structure and Dynamics of Networks. Princeton, USA : Princeton university press, 2006. 624 p.
4. Евин И.А. Введение с теорию сложных сетей : учеб. пособие. М. : Университет, 2010. Том 2. С. 121-141.
5. Albert R., Barabasi A.- L. Statistical mechanics of complex networks // Rev. Mod. Phys. 2002. V. 74. P. 42-97.
6. Олемской А.И. Статистика сложных сетей (обзор). «Вісник СумДУ». 2006. №6 (90). С.21-47.
7. Структуры последовательной активности в нейронных сетях со случайными связями / Т.А. Леванова, М.А. Комаров, Е.Ю. Кадина, Г.В. Осипов. Н. : Университет, 2010. 131-139.
8. Безмасштабные сети URL : http://wiki.witology.com/index.php/Безмасштабные_сети (дата звернення 15.11.2019).
9. Масштабно-инвариантные сети. URL : http://www.cognitivist.ru/er/kernel/prologi_11_scale_free_network.xml (дата звернення 10.11.2019).
10. Модель Барабаши-Альберт. URL : http://https://ru.wikipedia.org/wiki/Модель_Барабаши-Альберт (дата звернення 12.11.2019).
11. Barabási A.L., Albert R. Emergence of scaling in random networks // Science. 1999. 286 (5439). P. 509–512.

12. Райгородский А. М. Модели случайных графов и их применения : учеб. пособие. Киев, 2009. 117 с.
13. Newman M. Networks. An Introduction. Oxford : Oxford University Press, 2010.
14. Newman M. Mixing patterns in networks // Phys. Rev. E. 2003. vol. 67.
15. Noldus R., Van Mieghem P. Assortativity in complex networks // J. Complex Networks. 2015. vol. 3. P.507-542.
16. Шергін В.Л., Удовенко С.Г., Загребельна М.Ф. Структура екстремально асортативних масштабно-інваріантних мереж // Фізико-технологічні проблеми передавання, оброблення та зберігання інформації в інфокомунікаційних системах. Чернівці, 2018.
17. Задорожный В.Н., Юдин Е.Б. Статистически однородные случайные графы: определение, генерация, применение. Омск : Омский научный вестник (настоящий выпуск), ОМГТУ. 2009.
18. Берновский М.М. Случайные графы, модели и генераторы безмасштабных графов : учеб. пособие. Москва, 2009.
19. MATLAB. URL : <https://ru.wikipedia.org/wiki/MATLAB> (дата звернення 25.11.2019).
20. Matlab. URL : <http://matlab.ru/products/matlab> (дата звернення 25.11.2019).