

Харківський національний університет радіоелектроніки

Факультет інформаційно-аналітичних технологій та менеджменту

Кафедра прикладної математики

Рівень вищої освіти другий (магістерський)

Спеціальність 124 Системний аналіз

(код і повна назва)

Тип програми освітньо-професійна

(освітньо-професійна або освітньо-наукова)

Освітня програма Системний аналіз і управління

(повна назва)

ЗАТВЕРДЖУЮ:

Зав. кафедри ПМ _____

(підпис)

“25” листопада 2024 р.

ЗАВДАННЯ
НА КВАЛІФІКАЦІЙНУ РОБОТУ

здобувачеві Гирці Костянтину Геннадійовичу

(прізвище, ім'я, по батькові)

1. Тема роботи Аналіз математичних марковських моделей прогнозування
поширення фейкових новин у соціальних мережах

затверджена наказом по університету від 22 листопада 2024 р. № 1228 Ст

2. Термін подання здобувачем роботи до екзаменаційної комісії 6 січня 2025 р.

3. Вихідні дані до роботи математична модель марковського процесу

4. Перелік питань, що потрібно опрацювати в роботі _____

1. Системний аналіз предметної області

2. Вибір і обґрунтування методу розв'язання

3. Програмна реалізація

4. Результати обчислювального експерименту

5. Аналіз можливих застосувань

5. Перелік графічного матеріалу із зазначенням креслеників, схем, плакатів, комп'ютерних ілюстрацій _____

1. Актуальність теми роботи _____

2. Постановка задачі _____

3. Системний аналіз предметної області _____

4. Метод чисельного аналізу _____

5. Результати обчислювального експерименту _____

КАЛЕНДАРНИЙ ПЛАН

№	Назва етапів роботи	Терміни виконання етапів роботи	Примітка
1	Підбір та вивчення технічної літератури за темою роботи	25 листопада – 1 грудня 2024 р.	виконано
2	Вибір та обґрунтування методу	2 – 8 грудня 2024 р.	виконано
3	Розробка алгоритму і програми	9 – 22 грудня 2023 р.	виконано
4	Проведення аналітичних досліджень та розрахунків	23 – 29 грудня 2024 р.	виконано
5	Робота над текстом пояснювальної записки	30 грудня 2024 р. – 9 січня 2025 р.	виконано
6	Представлення роботи на рецензію в ЕК	10 січня 2025 р.	виконано

Дата видачі завдання 25 листопада 2024 р.

Здобувач _____
(підпис)

Керівник роботи _____ доц. Ситникова Ю. В.
(підпис) (посада, прізвище, ініціали)

РЕФЕРАТ

Пояснювальна записка: 53 с., 7 табл., 16 рис., 1 дод., 12 джерел.

МАРКОВСЬКІ ПРОЦЕСИ, ГРАФИ, ЗАДАЧА ВИЯВЛЕННЯ ФЕЙКОВИХ НОВИН, RUTHON.

Об'єкт дослідження – процес поширення інформації у соціальних мережах.

Мета роботи – розробка математичної моделі на основі марковських процесів для прогнозування поширення фейкових новин у соціальних мережах і дослідження ефективності різних підходів до їх нейтралізації.

Методи дослідження – теорія марковських процесів, чисельне моделювання та аналіз графів.

У роботі проведено моделювання процесу поширення фейкових новин у соціальних мережах за допомогою марковських моделей.

Аналіз отриманих результатів для різних сценаріїв розповсюдження включав порівняння базових і впливових сценаріїв, використання різних метрик для аналізу та візуалізацію результатів.

Висновки за результатами моделювання показали, що:

- швидкість і обсяг поширення фейкових новин значно залежать від структури мережі та взаємодій користувачів;
- зміни в алгоритмах та політиках модерації можуть суттєво зменшити швидкість розповсюдження неправдивої інформації;
- раннє виявлення та втручання, а також використання інструментів фактчекінгу є критично важливими;
- підвищення обізнаності користувачів відіграє важливу роль у протидії дезінформації;

Результати моделювання забезпечують основу для розробки ефективних стратегій боротьби з поширенням фейкових новин у соціальних мережах.

ABSTRACT

Introductory note: 53 p., 7 tables, 16 figures, 1 appendix, 12 references.

MARKOV PROCESSES, GRAPHS, FAKE NEWS DETECTION PROBLEM, PYTHON.

The object of research is the information propagation process in social networks.

The purpose of this work is to develop a mathematical model based on Markov processes to predict fake news spread in social networks and to study the effectiveness of various approaches to neutralise them.

Methods of research are Markov process theory, numerical modelling, graph analysis, and statistical data processing.

The modelling of the fake news propagation process in social networks using Markov models was conducted during this study.

The analysis of the obtained results for different propagation scenarios included a comparison of baseline and influential scenarios, the use of various metrics for analysis, and visualization of results.

Conclusions based on the modelling results showed that:

- the speed and volume of fake news spread significantly depend on the network structure and user interactions;
- changes in algorithms and moderation policies can substantially reduce the propagation speed of false information;
- early detection and intervention, as well as the use of fact-checking tools, are critically important;
- raising user awareness plays an important role in countering disinformation;

The modelling results provide a foundation for developing effective strategies to combat the spread of fake news on social networks.

ЗМІСТ

	С.
Вступ	8
1 Системний аналіз предметної області та постановка задач дослідження	10
1.1 Системний аналіз задачі виявлення фейкових новин та їх впливу у соціальних мережах	10
1.1.1 Вербальна модель системи	12
1.1.2 Морфологічний опис системи	12
1.1.3 Функціональна модель системи	13
1.2 Аналіз сценаріїв вирішення задачі Марковських процесів для прогнозування інформаційних потоків	16
1.2.1 Модель аналізу системи	18
1.2.2 Оцінювання вектора пріоритетів незадоволеностей методом аналізу ієрархій	19
1.2.3 Модель вирішення проблеми	21
1.3 Змістовна та формальна постановка задачі	24
1.4 Постановка задачі дослідження	26
2 Вибір та обґрунтування методу розв’язання	28
2.1 Марковська модель	28
2.2 Моделювання процесу поширення фейкових новин	29
2.3 Алгоритм розв’язання задачі для прогнозування поширення фейкових новин	31
Висновки за розділом 2	33
3 Програмна реалізація	34
3.1 Вибір програмного забезпечення для моделювання	34
3.2 Алгоритм розв’язання задачі для прогнозування поширення фейкових новин	35
3.3 Опис програмного коду та реалізованих функцій	36
Висновки за розділом 3	38

	7
4 Результати обчислювального експерименту та їх аналіз	40
4.1 Методологія аналізу результатів моделювання	40
4.2 Симуляція поширення фейкових новин у соціальних мережах	41
Висновки за розділом 4	45
Висновки	47
Перелік джерел посилання	50
Додаток А	52

ВСТУП

Актуальність теми. Фейкові новини в сучасному інформаційному середовищі стали важливою проблемою, яка впливає на суспільство, політику, економіку та безпеку. Соціальні мережі значно прискорили поширення інформації, зокрема недостовірної, завдяки їхній глобальній доступності та швидкості обміну даними. Водночас механізми алгоритмічної рекомендації контенту підсилюють поширення фейкових новин, оскільки такі новини часто викликають сильні емоційні реакції, привертаючи більше уваги користувачів.

Актуальність роботи зумовлена тим, що в умовах сучасного інформаційного суспільства фейкові новини у соціальних мережах становлять значну загрозу як для суспільної думки, так і для національної безпеки. Швидке поширення недостовірної інформації призводить до дезінформації громадян, загострення соціальної напруженості та виникнення конфліктів. Аналіз математичних марковських моделей дає змогу прогнозувати розповсюдження фейкових новин і розробити ефективні механізми їх нейтралізації.

Мета і завдання кваліфікаційної роботи. Метою даної кваліфікаційної роботи є розробка математичної моделі на основі марковських процесів для прогнозування поширення фейкових новин у соціальних мережах.

Встановлено такий перелік задач, які необхідно виконати для досягнення поставленої мети:

- провести огляд і аналіз сучасних математичних моделей поширення інформації у соціальних мережах;
- адаптувати марковські моделі для задач прогнозування фейкових новин;
- розробити алгоритм розв’язання задачі моделювання поширення фейкових новин;
- підготувати програмну реалізацію розробленого алгоритму.

Об’єктом дослідження є процес поширення інформації у соціальних мережах.

Предметом дослідження є математичні моделі та методи прогнозування

поширення інформації, які ґрунтуються на марковських процесах.

Методи дослідження. У кваліфікаційній роботі використовуються методи теорії марковських процесів, чисельного моделювання та аналізу графів.

1 СИСТЕМНИЙ АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ ТА ПОСТАНОВКА ЗАДАЧ ДОСЛІДЖЕННЯ

1.1 Системний аналіз задачі виявлення фейкових новин та їх впливу у соціальних мережах

Системний аналіз поширення фейкових новин передбачає комплексний підхід до вивчення процесів створення, поширення та впливу цієї інформації. На першому етапі аналізується структура соціальних мереж як інформаційного середовища, де кожен вузол є користувачем, а зв'язки між вузлами відображають взаємодії між користувачами. Також вивчаються алгоритми формування стрічки новин, які можуть сприяти поширенню фейкових новин через так звані «інформаційні бульбашки» та «резонансні камери» [1].

Далі досліджується поведінка користувачів під час взаємодії з фейковими новинами. Важливими факторами є рівень довіри до джерел інформації, швидкість поширення новин, психологічна реакція користувачів та їх схильність до поширення новин без перевірки достовірності. Системний підхід дозволяє врахувати не лише технічні аспекти, але й соціальні, психологічні та етичні фактори.

Особливу увагу в системному аналізі приділяється моделюванню поширення фейкових новин. Це включає створення математичних моделей, які описують ймовірність передачі інформації між користувачами, а також враховують структуру мережі, інтенсивність взаємодій та часову динаміку. Такі моделі можуть базуватися на методах теорії графів, марковських процесах або епідеміологічних моделях, адаптованих до соціальних мереж.

Вплив фейкових новин у соціальних мережах оцінюється за кількома ключовими показниками, такими як: охоплення аудиторії, зміна громадської думки, політичні або економічні наслідки. Окрім цього, аналізуються заходи боротьби з фейковими новинами, включаючи перевірку фактів, підвищення медіаграмотності користувачів, впровадження технологій штучного інтелекту для

автоматичного виявлення недостовірної інформації та зміну алгоритмів роботи соціальних платформ.

Дійсно, системний аналіз дає змогу не лише зрозуміти механізми поширення фейкових новин, але й оцінити ефективність заходів для їх нейтралізації, що є важливим для забезпечення інформаційної безпеки сучасного суспільства.

Опишемо методи аналізу поширення інформації. Одним з основних інструментів для аналізу структури та динаміки соціальних мереж є теорія графів. Вона дозволяє представляти соціальні мережі як графи, де користувачі є вершинами, а їхні взаємозв'язки – ребрами. За допомогою різних метрик, таких як: ступінь вершин, кластерний коефіцієнт, діаметр графа, можна аналізувати, як швидко може поширюватися інформація в мережі.

Марковські процеси – це моделі, які використовуються для моделювання стохастичних процесів, де наступний стан системи залежить лише від її поточного стану. В контексті соціальних мереж, марковські процеси можуть допомогти визначити ймовірність переходу інформації від одного користувача до іншого.

В епідеміології використовують моделі SIR (сприйнятливий-інфікований-одужав) або SIS (сприйнятливий-інфікований-сприйнятливий), які дозволяють аналізувати поширення інформації аналогічно до вірусних захворювань. Це допомагає досліджувати швидкість та охоплення інформаційних кампаній та дезінформації.

Комп'ютерне моделювання та симуляції це потужний інструментарій для створення складних моделі з імітацією різних сценаріїв поширення інформації. Ці методи дозволяють визначати, як зміни у параметрах мережі можуть впливати на динаміку поширення інформації.

З розвитком технологій збору та аналізу великих даних з'явилася можливість реального аналізу поведінки користувачів на основі їхньої активності в соціальних мережах. Методи машинного навчання та штучного інтелекту використовуються для визначення шаблонів поширення інформації і її впливу на громадську думку [2].

1.1.1 Вербальна модель системи

Об'єкт аналізу: процес поширення інформації у соціальних мережах.

Предмет аналізу: математичні моделі та методи прогнозування поширення інформації, які ґрунтуються на марковських процесах.

Точка зору: дослідник.

Ціль: розробка математичної моделі для прогнозування поширення фейкових новин у соціальних мережах.

1.1.2 Морфологічний опис системи

Розглянемо систему «Дослідження процесу поширення інформації у соціальних мережах». Як відомо, морфологічна модель передбачає виконання опису структури, меж та зовнішнього середовища.

Для розуміння зовнішнього середовища системи часто використовують модель «чорної скриньки». Для розглянутої системи модель «чорна скринька» представляє собою спрощену математичну або чисельну модель, яка акцентується на вхідних та вихідних параметрах процесу без опису його внутрішньої структури. Входом до моделі «чорна скринька» є параметри соціальної мережі, характеристики інформації та початкові умови поширення. Виходом «чорної скриньки» є чисельні показники розповсюдження інформації, такі як охоплення, швидкість поширення та вплив на користувачів.

Модель «чорна скринька» подана на рисунку 1.1.

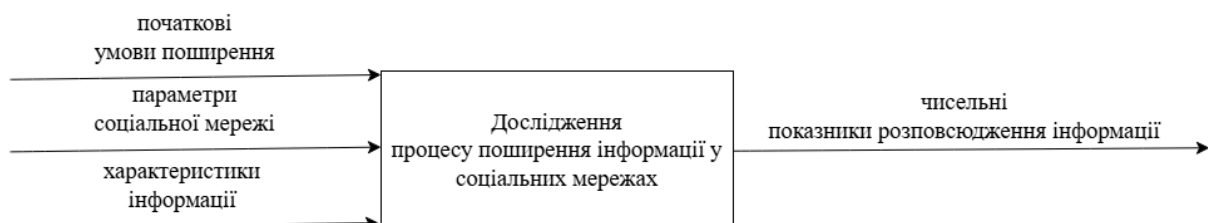


Рисунок 1.1 – Модель «чорна скринька»

Зовнішнім середовищем є умови та фактори, які впливають на роботу процесу. Для системи «Дослідження процесу поширення інформації у соціальних мережах» об'єктами зовнішнього середовища є користувачі соціальних мереж, алгоритми соціальних мереж, теорія соціальних мереж та методи аналізу даних.

Межею системи «Дослідження процесу поширення інформації у соціальних мережах» є віртуальний простір, в якому дослідник моделює процес поширення інформації за допомогою марковських процесів, розробляє алгоритми прогнозування розповсюдження інформації та аналізує отримані результати.

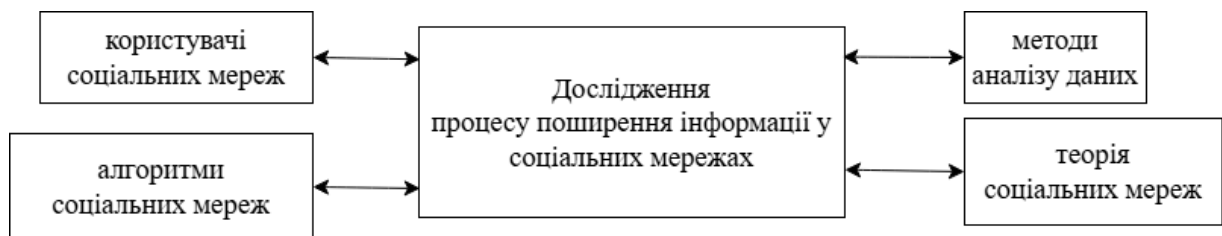


Рисунок 1.2 – Модель зовнішнього середовища системи

1.1.3 Функціональна модель системи

Функціональна модель системи «Дослідження процесу поширення інформації у соціальних мережах» включає послідовність дій, метою яких є моделювання та аналіз розповсюдження інформації в соціальних мережах за допомогою марковських процесів, з метою отримання прогнозів та оцінки ефективності різних стратегій поширення або протидії інформації.

Розв'язання задачі моделювання поширення інформації включає побудову моделі, яка відображає динаміку розповсюдження інформації в соціальній мережі.

Розглянемо основні функціональні компоненти цієї системи:

– введення даних, зокрема, параметри соціальної мережі, характеристики інформації, початкові умови поширення;

- побудова математичної моделі на основі марковських процесів;
- проведення симуляцій поширення інформації;
- аналіз результатів;
- візуалізація отриманих даних.

Методологія IDEF0, яка використовується для візуалізації функціональної моделі, дозволяє ефективно представити функції системи та їх взаємозв'язки. Ця методологія передбачає створення ієрархії діаграм, де ключову роль відіграє контекстна діаграма. Ця перша діаграма в ієрархії IDEF0 забезпечує найбільш узагальнений огляд системи, демонструючи її загальне функціонування та взаємодію із зовнішнім середовищем. Таким чином, контекстна діаграма служить відправною точкою для розуміння системи в цілому, перш ніж заглиблюватися в деталі її окремих компонентів.

Контекстна діаграма IDEF0 для системи «Дослідження процесу поширення інформації у соціальних мережах» (наведена на рисунку 1.3) може бути представлена наступним чином: на вхід цієї системи подаються параметри соціальної мережі, характеристики інформації та початкові умови поширення; засобами управління є теорія соціальних мереж та методи аналізу даних; механізмами системи є дослідник та програмне забезпечення; на виході маємо прогнози поширення інформації.

Після створення загального опису системи, наступним етапом є проведення декомпозиції. Цей процес починається з розбиття системи на великі фрагменти (перший рівень), а потім продовжується декомпозицією цих фрагментів на дрібніші компоненти (другий і наступні рівні).

Для системи «Дослідження процесу поширення інформації у соціальних мережах» декомпозиція контекстної діаграми (рис. 1.4) відображає основні функції системи, такі як побудова математичної моделі на основі марковських процесів та моделювання поширення інформації. Подальша декомпозиція другого рівня (рис. 1.5) може фокусуватися на деталізації конкретної функції, наприклад, «Побудова математичної моделі на основі марковських процесів», розкриваючи підпроцеси, які відбуваються на цьому етапі.

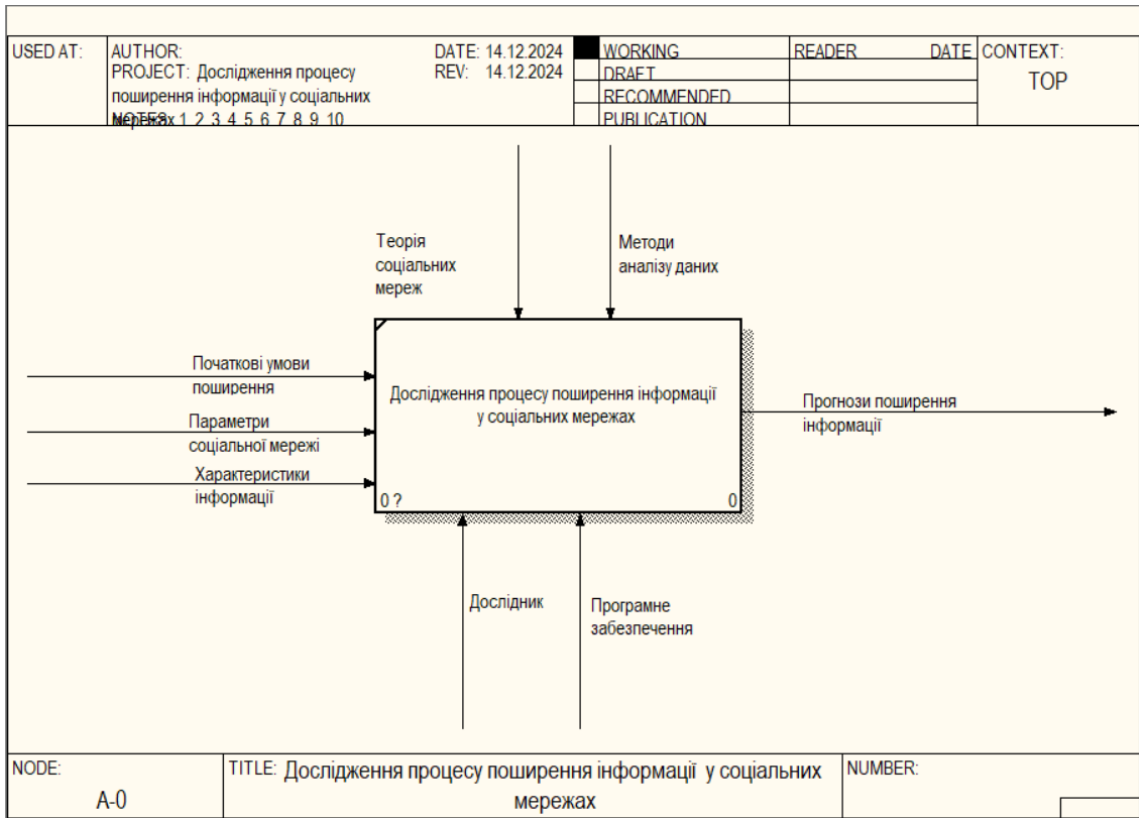


Рисунок 1.3 – Контекстна діаграма (рівень А-0)

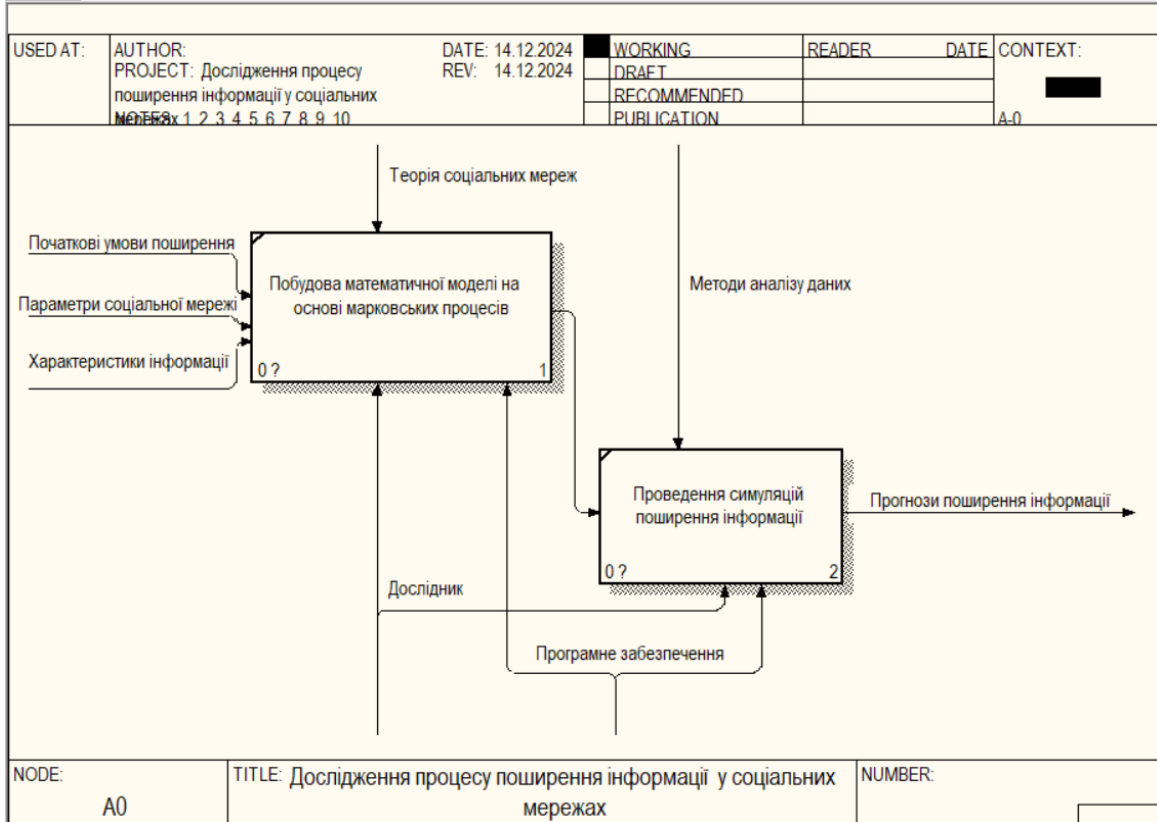


Рисунок 1.4 – Декомпозиція роботи «Дослідження процесу поширення інформації у соціальних мережах»: рівень A0

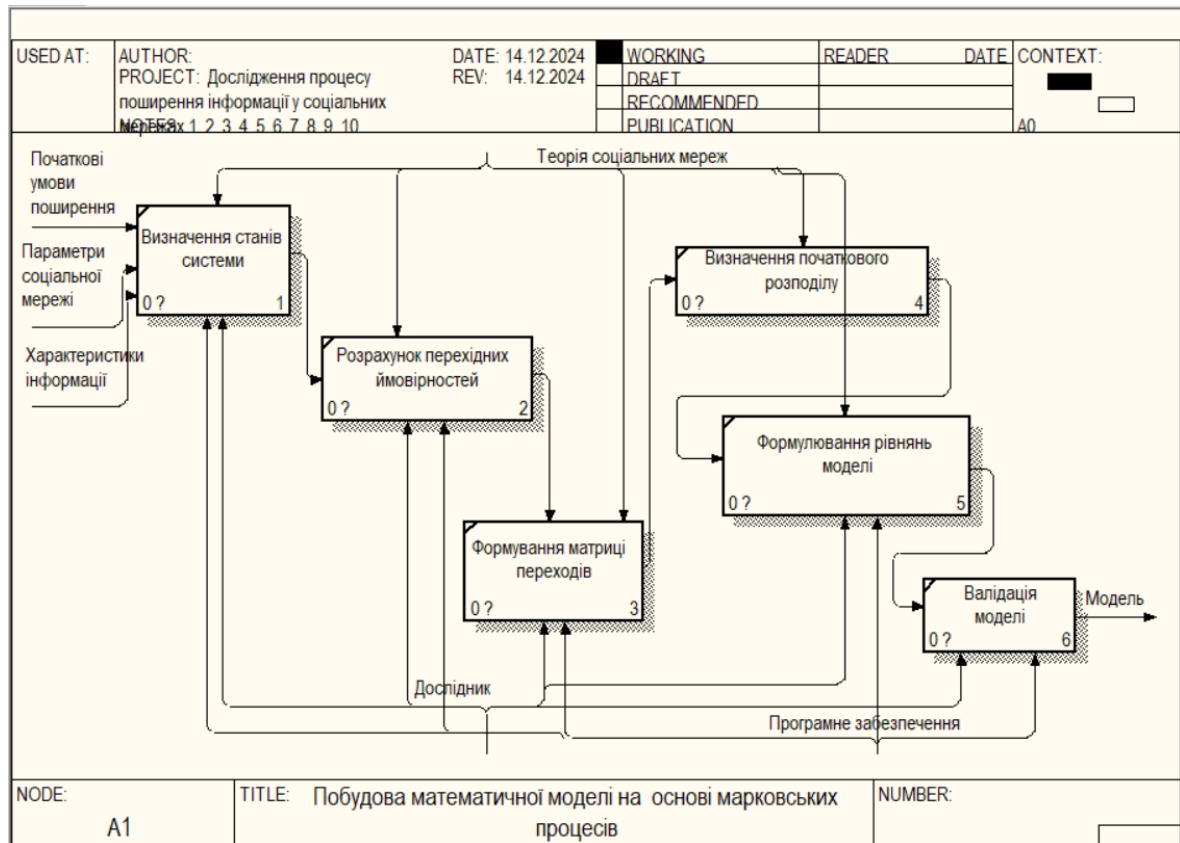


Рисунок 1.5 – Декомпозиція роботи «Побудова математичної моделі на основі марковських процесів»: рівень А1

1.2 Аналіз сценаріїв вирішення задачі марковських процесів для прогнозування інформаційних потоків

Однією з найбільш поширених математичних технік у цій області є теорія графів. Соціальні мережі можна моделювати у вигляді графів, де користувачі представлені вершинами, а їхні взаємозв'язки – ребрами. Це дозволяє аналізувати різноманітні метрики, такі як ступінь вузлів, кластеризація, та шляхи між вузлами, щоб зрозуміти, як інформація може поширюватися мережею.

Ще одним важливим підходом є використання марковських моделей. У контексті соціальних мереж за допомогою марковських процесів можна моделювати інформаційні потоки як серії станів, де кожен стан представляє можливі зміни у поширенні інформації серед груп користувачів. Ці моделі допомагають

прогнозувати вірогідність поширення певних новин чи інформації в майбутньому.

Диференціальні рівняння також застосовуються для моделювання динамічних систем, включно з інформаційними потоками. Моделі, засновані на диференціальних рівняннях, можуть допомогти вивчати зміни у поширенні інформації з часом, дозволяючи дослідникам аналізувати швидкість та масштаб поширення інформації [3].

Для розгляду впливу математичних моделей на аналіз інформаційних потоків, було створено таблицю, що підсумовує основні характеристики кожного з розглянутих методів (табл. 1.1). Це дозволило краще зрозуміти переваги та обмеження кожної моделі в контексті аналізу соціальних мереж.

Таблиця 1.1 – Основні математичні моделі для аналізу інформаційних потоків

Метод	Опис	Переваги	Обмеження
Теорія графів	Моделювання соціальних мереж як графів, де користувачі – вершини, зв'язки – ребра	Наочність, зручність аналізу структур	Вимагає великих обчислювальних ресурсів при збільшенні масштабів мережі
Марковські процеси	Моделювання інформаційних потоків як процесів з визначеними станами і переходами	Точність прогнозування, гнучкість у моделюванні	Складність налаштування і трактування моделі
Диференціальні рівняння	Використання рівнянь для опису динаміки поширення інформації	Дозволяє аналізувати зміни з часом	Висока складність, необхідність точних вихідних даних

Ця таблиця демонструє основні аспекти кожного методу і може бути використана для обґрунтування обраного підходу до аналізу інформаційних потоків у дослідженні, особливо у контексті поширення фейкових новин. Вибір методу залежить від конкретної мети аналізу, доступних даних та ресурсів для обчислень.

1.2.1 Модель аналізу системи

Також для вибору методу виявлення фейкових новин та їх впливу у соціальних мережах було використано метод аналізу ієрархій. Розглянемо порівняння методів теорії графів, марковських процесів та диференціальних рівнянь для цієї задачі.

Критерії вибору методу:

- критерій 1 (K1): точність виявлення фейкових новин;
- критерій 2 (K2): швидкість аналізу;
- критерій 3 (K3): масштабованість для великих соціальних мереж.

Альтернативи:

- альтернатива 1 (A1): теорія графів;
- альтернатива 2 (A2): марковські процеси;
- альтернатива 3 (A3): диференціальні рівняння.

На рисунку 1.6 зображена ієрархічна модель проблеми вибору.

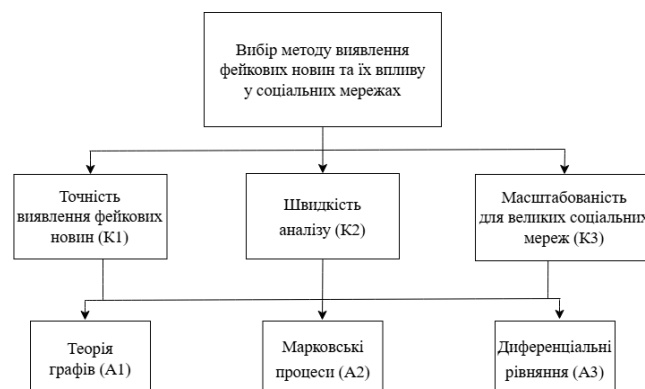


Рисунок 1.6 – Ієрархічна модель аналізу проблеми

1.2.2 Оцінювання вектора пріоритетів незадоволеностей методом аналізу ієрархій

Для початку сформулюємо матрицю парних порівнянь критеріїв та альтернатив за кожним з критеріїв. Шляхом експертного опитування отримано дані, які наведено у таблиці 1.2.

Таблиця 1.2 – Матриця парних порівнянь критеріїв

Критерії оцінювання	К1	К2	К3	Оцінка компонентів	Вектор пріоритетів
К1	1	3	5	2,466	0,637
К2	$\frac{1}{3}$	1	3	1,000	0,258
К3	$\frac{1}{5}$	$\frac{1}{3}$	1	0,405	0,105
Усього				3,871	1

Знаходимо власний вектор λ_{\max} за формулою $\lambda_{\max} \approx \sum_{j=1}^3 y_j p_j$, де

$y_j = \sum_{i=1}^3 a_{ij}$ – сума елементів j -го стовпчика. Отримуємо, що $\lambda_{\max} \approx 3,074$. Ін-

декс узгодженості $CI = \frac{\lambda_{\max} - 3}{3 - 1} = 0,037$, випадковий індекс (RI) для матриць

третього порядку дорівнює 0,58. Тоді відносна узгодженість дорівнює

$CR = \frac{CI}{RI} = \frac{0,037}{0,58} = 0,064$. Оскільки відносна узгодженість менша за 0,2, то це

свідчить про задовільність логічних суджень експерта.

Формуємо матриці попарних порівнянь альтернатив за кожним критерієм, які представлені в таблицях 1.3 – 1.5.

Таблиця 1.3 – Матриця парних порівнянь альтернатив за першим критерієм

К1	A1	A2	A3	Оцінка компонентів	Вектор пріоритетів
A1	1	$\frac{1}{3}$	2	0,874	0,238
A2	3	1	4	2,289	0,623
A3	$\frac{1}{2}$	$\frac{1}{4}$	1	0,500	0,139
Усього				3,663	1

Знаходимо, що $\lambda_{\max} \approx 3,054$. Індекс узгодженості $CI = \frac{\lambda_{\max} - 3}{3 - 1} = 0,027$.

Тоді відносна узгодженість дорівнює $CR = \frac{CI}{RI} = \frac{0,027}{0,58} = 0,047$. Відносна узгодженість менша за 0,2, що свідчить про задовільність логічних суджень експерта.

Таблиця 1.4 – Матриця парних порівнянь альтернатив за другим критерієм

К1	A1	A2	A3	Оцінка компонентів	Вектор пріоритетів
A1	1	$\frac{1}{2}$	2	1,000	0,297
A2	2	1	3	1,817	0,540
A3	$\frac{1}{2}$	$\frac{1}{3}$	1	0,550	0,163
Усього				3,367	1

Знаходимо, що $\lambda_{\max} \approx 3,018$. Індекс узгодженості $CI = \frac{\lambda_{\max} - 3}{3 - 1} = 0,009$.

Тоді відносна узгодженість дорівнює $CR = \frac{CI}{RI} = \frac{0,009}{0,58} = 0,016$. Відносна узгодженість менша за 0,2, що свідчить про задовільність логічних суджень експерта.

Таблиця 1.5 – Матриця парних порівнянь альтернатив за третім критерієм

K1	A1	A2	A3	Оцінка компонентів	Вектор пріоритетів
A1	1	$\frac{1}{2}$	1	0,794	0,250
A2	2	1	3	1,817	0,572
A3	1	$\frac{1}{3}$	1	0,693	0,178
Усього				3,304	1

Знаходимо, що $\lambda_{\max} = 3,027$. Індекс узгодженості $CI = \frac{\lambda_{\max} - 3}{3 - 1} = 0,014$.

Тоді відносна узгодженість дорівнює $CR = \frac{CI}{RI} = \frac{0,014}{0,58} = 0,024$. Відносна узгодженість менша за 0,2, що свідчить про задовільність логічних суджень експерта.

1.2.3 Модель вирішення проблеми

Отримавши вектори локальних переваг можемо побудувати вектор глобальних переваг. Для цього складемо матрицю, де стовпці є векторами локальних пріоритетів. У таблиці 1.6 наведено результати розрахунків.

Таблиця 1.6 – Узагальнені пріоритети альтернатив

Критерій Альтернатива	K1	K2	K3	Узагальнені пріоритети
A1	0,238	0,297	0,250	0,253
A2	0,623	0,540	0,572	0,597
A3	0,139	0,163	0,178	0,150

Тоді вектор глобальних пріоритетів має вигляд:

$$\vec{p} = \begin{pmatrix} 0,238 & 0,297 & 0,250 \\ 0,623 & 0,540 & 0,572 \\ 0,139 & 0,163 & 0,178 \end{pmatrix} \cdot \begin{pmatrix} 0,253 \\ 0,597 \\ 0,150 \end{pmatrix} = (0,253; 0,597; 0,150)^T.$$

Розрахуємо індекс узгодженості та відношення узгодженості для всієї ієрархії:

$$CI = CI^K + (\vec{p}^K, CI^K) = 0,037 + 0,027 \cdot 0,637 + 0,009 \cdot 0,258 + 0,014 \cdot 0,105 = \\ = 0,037 + 0,017 + 0,002 + 0,001 = 0,057,$$

$$RI = RI^K + RI^A = 0,58 + 0,58 = 1,16,$$

$$CR = \frac{CI}{RI} = \frac{0,058}{1,16} = 0,049.$$

Оскільки ми є особою, що приймає рішення, то на основі вектора глобальних пріоритетів можемо зробити наступний висновок: для вибору методу виявлення фейкових новин та їх впливу у соціальних мережах необхідно обрати другу альтернативу – марковські процеси.

Загалом, комбінація цих методів дозволяє комплексно підходити до аналізу інформаційних потоків і формувати базу для розробки ефективних стратегій контролю та управління інформаційними процесами в цифровому просторі.

Марковські процеси характеризуються станами і переходами між ними,

де кожен перехід має певну ймовірність. Це особливо корисно для моделювання соціальних мереж, де користувачі можуть «перемикатися» між різними станами (наприклад, прийняття, ігнорування, або розповсюдження інформації).

Застосування у прогнозуванні:

– марковські процеси дозволяють аналізувати, як швидко і ефективно може поширюватися інформація в соціальних мережах, а за допомогою моделювання, які новини стануть вірусними, а які швидко затихнуть;

– марковські процеси для аналізу поведінки користувача, як користувачі реагують на різні типи новин (фейкові проти реальних), може допомогти зрозуміти патерни їхньої поведінки та розробити методики для підвищення обізнаності та критичного сприйняття інформації;

– марковські процеси можуть бути використані для оцінки того, як різні стратегії втручання (наприклад, введення фактчекінгу) можуть вплинути на динаміку поширення фейкових новин.

Марковські процеси вимагають чіткого визначення станів і перехідних ймовірностей, що може бути викликом у випадку складних соціальних мереж із великою кількістю активних учасників та зв'язків. Отже, розробка адекватної Марковської моделі потребує глибокого аналізу поведінки мережі та доступу до детальних даних про взаємодії між її учасниками [4].

Для детального розуміння застосування марковських процесів у прогнозуванні поширення фейкових новин у соціальних мережах було створено таблицю, що перераховує основні аспекти та параметри, які слід враховувати при моделюванні (табл. 1.7). Таблиця допоможе систематизувати знання та підкреслити особливості, які роблять марковські процеси важливим інструментом у галузі прогнозування. Вона продемонструє, як марковські процеси можуть бути використані для аналізу та прогнозування поведінки інформації в соціальних мережах. Важливим також є глибоке розуміння ймовірностей переходів та динаміки між станами, що дає змогу ефективно оцінювати й прогнозувати розвиток інформаційних кампаній, зокрема тих, що базуються на фейкових новинах.

Завдяки своїй здатності до моделювання складних систем з чітко визна-

ченими правилами, Марковські процеси надають важливу інструментарій для дослідників, що прагнуть зрозуміти та керувати поширенням інформації в епоху цифрових технологій.

Таблиця 1.7 – Застосування марковських процесів у прогнозуванні поширення фейкових новин

Параметр	Опис
Визначення станів	Кожен стан у моделі представляє можливий рівень поширення новини.
Ймовірності переходів	Визначення ймовірностей переходу між станами, які вказують на те, наскільки легко чи важко може новина перейти від одного рівня поширення до іншого.
Вплив соціальних зв'язків	Аналіз, як зв'язки між користувачами впливають на переходи між станами. Наприклад, користувачі з великим числом підписників можуть спричинити ширше поширення новини.
Динаміка поширення	Дослідження того, як швидко новини переходять від початкових до кінцевих станів, що відображає їх вірусний потенціал.
Чутливість до змін	Оцінка чутливості моделі до змін у вхідних даних, наприклад, зміни в поведінці користувачів або впливі зовнішніх чинників (наприклад, запровадження фактчекінгу).

1.3 Змістовна та формальна постановка задачі

Соціальні мережі стали потужним інструментом для обміну інформацією, але водночас створили умови для швидкого поширення фейкових новин. Це негативно впливає на інформаційне середовище, формує хибні уявлення в суспільстві та може викликати соціальні, політичні або економічні кризи. Завдання полягає у розробці математичної моделі, яка дозволить прогнозувати поширення фейкових новин у соціальних мережах, враховуючи взаємодію користувачів,

ймовірність поширення новин та їх сприйняття.

Основні аспекти задачі: описати механізми поширення новин, зокрема фейкових, у соціальних мережах; визначити основні параметри моделі (ймовірність поширення, рівень довіри користувачів, кількість взаємодій тощо); використати марковські моделі для прогнозування поширення інформації у часі; розробити методи зменшення впливу фейкових новин, використовуючи результати моделювання.

Вхідні дані:

- граф соціальної мережі $G = (V, E)$, де V – множина вузлів (користувачів), E – множина ребер (зв'язків між користувачами);
- початковий набір вузлів S_0 , що вже «заражені» фейковою новиною;
- ймовірність поширення новини $p(u, v)$ між вузлами u і v , визначена для кожного ребра $e \in E$;
- часовий горизонт T , протягом якого відбувається моделювання.

Вихідні дані:

- прогноз кількості вузлів N_t , що будуть «заражені» фейковою новиною у момент часу $t, t = 1, 2, 3 \dots T$;
- ймовірність досягнення новиною заданої підмножини вузлів $V' \subseteq V$ за час T .

Умови:

- поширення новини моделюється як марковський процес, де стан кожного вузла визначається ймовірністю його «зараження» на кожному кроці часу;
- ймовірність зміни стану вузла залежить лише від його поточного стану та стану суміжних вузлів.

Формулювання задачі:

- побудувати марковську модель, що описує процес поширення інформації. Нехай X_t – це стан системи у момент часу t , тоді

$$P(X_{t+1} | X_t, X_{t-1}, \dots, X_0) = P(X_{t+1} | X_t);$$

- визначити функцію переходу $P(X_{t+1} | X_t)$, яка залежить від ймовірнос-

тей $p(u, v)$ та структури графа G ;

– розв’язати задачу прогнозування кількості «заражених» вузлів N_i та ймовірності поширення новини до підмножини V' .

Критерії ефективності:

– точність прогнозу (відхилення моделі від реальних даних);
 – часова складність алгоритму прогнозування;
 – можливість впровадження заходів для обмеження поширення фейкових новин.

Ця постановка дозволяє формалізувати задачу та застосувати до неї існуючі методи математичного моделювання та аналізу.

1.4 Постановка задач дослідження

У цьому підрозділі задача дослідження поширення фейкових новин формулюється з метою розробки ефективної математичної моделі, яка б дозволила аналізувати та прогнозувати динаміку їх розповсюдження в соціальних мережах. Цей підхід включає визначення основних параметрів та структурування процесу дослідження таким чином, щоб вийти на конкретні практичні рекомендації з мінімізації негативного впливу фейкових новин [5, 6, 7].

Основна проблема полягає в тому, що фейкові новини швидко поширюються через соціальні мережі, використовуючи структурні особливості цих мереж, такі як тісні зв’язки між користувачами та алгоритми рекомендацій. Це може призводити до серйозних суспільних наслідків, включаючи політичні маніпуляції та соціальну нестабільність [8, 9].

Завдання дослідження:

– аналіз структури соціальних мереж, щоб зрозуміти, як властивості мережі впливають на швидкість і обсяг поширення інформації;
 – розробка математичної моделі, заснованої на марковських процесах або інших підходах, для моделювання процесів поширення фейкових новин;

- оцінка впливу факторів, таких як: зміни у поведінці користувачів або введення нових алгоритмів фільтрації контенту в соціальних мережах;
- розробка стратегій, спрямованих на зменшення швидкості поширення фейкових новин та зниження їхнього впливу на суспільство.

Очікувані результати: розробка та впровадження моделі має на меті не тільки наукове осмислення механізмів поширення фейкових новин але й розробку конкретних інструментів, які могли б бути застосовані платформами соціальних мереж для проактивного виявлення та блокування дезінформації. Це передбачає взаємодію з алгоритмами штучного інтелекту для адаптації рекомендаційних систем, зменшення вірусності шкідливих новин та підвищення обізнаності користувачів щодо якості інформації, яку вони споживають.

2 ВИБІР ТА ОБҐРУНТУВАННЯ МЕТОДУ РОЗВ'ЯЗАННЯ

2.1 Марковська модель

Марковська модель базується на концепції процесів, у яких стан системи в майбутньому залежить лише від її поточного стану, а не від попередніх станів. Такий підхід дозволяє описувати динаміку поширення фейкових новин через перехід між різними станами користувачів у соціальних мережах [10, 11, 12].

У контексті поширення фейкових новин, станом може бути рівень обізнаності або дій користувача щодо конкретної новини:

- S_1 : користувач не бачив новини;
- S_2 : користувач бачив новину, але не поділився нею;
- S_3 : користувач поділився новиною;
- S_4 : користувач дізнався, що новина є фейковою, і припинив її поширення.

Матриця ймовірностей переходу P описує ймовірність переходу від одного стану до іншого за одиницю часу:

$$P = \begin{pmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \\ p_{41} & p_{42} & p_{43} & p_{44} \end{pmatrix},$$

де p_{ij} – ймовірність переходу від стану S_i до S_j .

Вектор станів $X(t)$ відображає частку користувачів у кожному стані в момент часу t :

$$X(t) = [x_1(t), x_2(t), x_3(t), x_4(t)],$$

де $x_i(t)$ – частка користувачів у стані S_i .

Початковий розподіл користувачів за станами $X(0)$ задається відповідно до початкової ситуації у соціальній мережі.

Марковський процес описується рівнянням:

$$X(t+1) = X(t) \cdot P.$$

Повторне застосування матриці P дозволяє прогнозувати розподіл користувачів у майбутньому.

Оцінка параметрів:

- ймовірність переходу p_{12} : кількість користувачів, які побачили новину;
- ймовірність переходу p_{23} : ймовірність поділитися новиною;
- ймовірність переходу p_{34} : ймовірність виявлення фейковості новини.

Модель дозволяє проводити ітерації для отримання прогнозу на t -й крок.

Також модель може бути доповнена факторами, такими як: швидкість поширення новин, вплив популярності користувача чи ефективність системи перевірки фактів.

Загалом, модель може бути візуалізована у вигляді графа станів із ймовірностями переходів, а результати – у вигляді графіків, які демонструють динаміку часток користувачів у кожному стані.

2.2 Моделювання процесу поширення фейкових новин

Цей підрозділ зосереджений на моделюванні процесу поширення фейкових новин у соціальних мережах. Моделювання дозволяє досліджувати, аналізувати та прогнозувати, як інформація розповсюджується серед користувачів, та які фактори впливають на швидкість та обсяг її поширення. Це дуже важливо для розробки стратегій, що мінімізують поширення неправдивої інформації.

Моделювання процесу поширення фейкових новин здійснюється за допомогою марковських моделей, які були описані у попередніх розділах. Сценарії поширення новин моделюються на основі таких параметрів:

- стани новин, наприклад: «Неопублікована», «Опублікована», «Поширена», «Вірусна»;
- перехідні ймовірності, які визначають, з якою ймовірністю новина переходить від одного стану до іншого;
- зовнішні впливи, наприклад: роль алгоритмів соціальної мережі, взаємодія користувачів, вплив засобів масової інформації та інших платформ.

Основні аспекти моделювання:

- ініціація процесу поширення (новина починає свій шлях у стані «Опублікована», і з часом може перейти в «Поширена» та «Вірусна»);
- взаємодія користувачів (враховується, як взаємодії між користувачами (лайки, репости, коментарі) впливають на швидкість поширення новин);
- моніторинг і втручання (моделювання дозволяє ідентифікувати потенційні точки втручання, де соціальні мережі можуть активно втручатися для запобігання вірусного розповсюдження фейкових новин);
- оцінка сценаріїв (симуляція дозволяє оцінювати ефективність різних стратегій боротьби з фейковими новинами, тестуючи, як зміни у мережевих політиках або користувацьких взаємодіях впливають на поширення інформації).

Для моделювання використовується програмний код, заснований на Python, який використовує бібліотеки NumPy та NetworkX для обробки даних та візуалізації процесів. Сценарії моделювання можуть бути налаштовані для включення різноманітних чинників впливу та оцінювання їх впливу на динаміку розповсюдження фейкових новин.

Ці моделі та симуляції забезпечують важливий інструмент для розуміння та керування інформаційним простором у соціальних мережах, дозволяючи боротися з поширенням дезінформації більш ефективно і своєчасно.

2.3 Алгоритм розв'язання задачі для прогнозування поширення фейкових новин

Для прогнозування поширення фейкових новин за допомогою марковських моделей можна розробити алгоритм, який буде враховувати поведінкові аспекти процесу розповсюдження інформації. Алгоритм може бути реалізований у кілька кроків, кожен з яких спрямований на аналіз та прогнозування поведінки інформаційних каскадів у соціальних мережах.

Кроки алгоритму:

а) ініціалізація моделі;

1) визначення можливого стану;

2) встановлення ймовірності переходів між цими станами;

б) збір та обробка даних;

1) зібрати дані про поширення новин у соціальних мережах;

2) обробити ці дані, виокремивши основні шаблони та аномалії;

в) моделювання поширення новин;

1) симуляція потенційної траєкторії розвитку станів;

2) аналіз змін у мережі або поведінці користувачів;

г) інтеграція з інструментами виявлення та блокування фейкових новин.

Можна інтегрувати алгоритм з системами штучного інтелекту або машинного навчання, які вже використовуються платформами соціальних мереж для автоматичного виявлення та блокування фейкових новин.

Опис алгоритму:

– ініціалізація станів новини (треба визначити можливі стани новини, наприклад: «Неопублікована», «Опублікована», «Поширена», «Вірусна»);

– встановлення перехідних ймовірностей (задання ймовірності переходу між станами);

– вибір початкового стану (новина починається у стані «Неопублікована» або «Опублікована» залежно від контексту симуляції);

– симуляція поширення (використовуючи марковські процеси, моделю-

ється перехід новини між станами до досягнення кінцевого стану або до закінчення встановленого часу моделювання);

– аналіз результатів (збереження і аналіз того, як новина переходила між станами, оцінюючи вплив алгоритмічних змін соціальної мережі або інших інтервенцій).

Наступна UML діаграма показує послідовність виконання алгоритму моделювання процесу поширення новин (рис. 2.1).

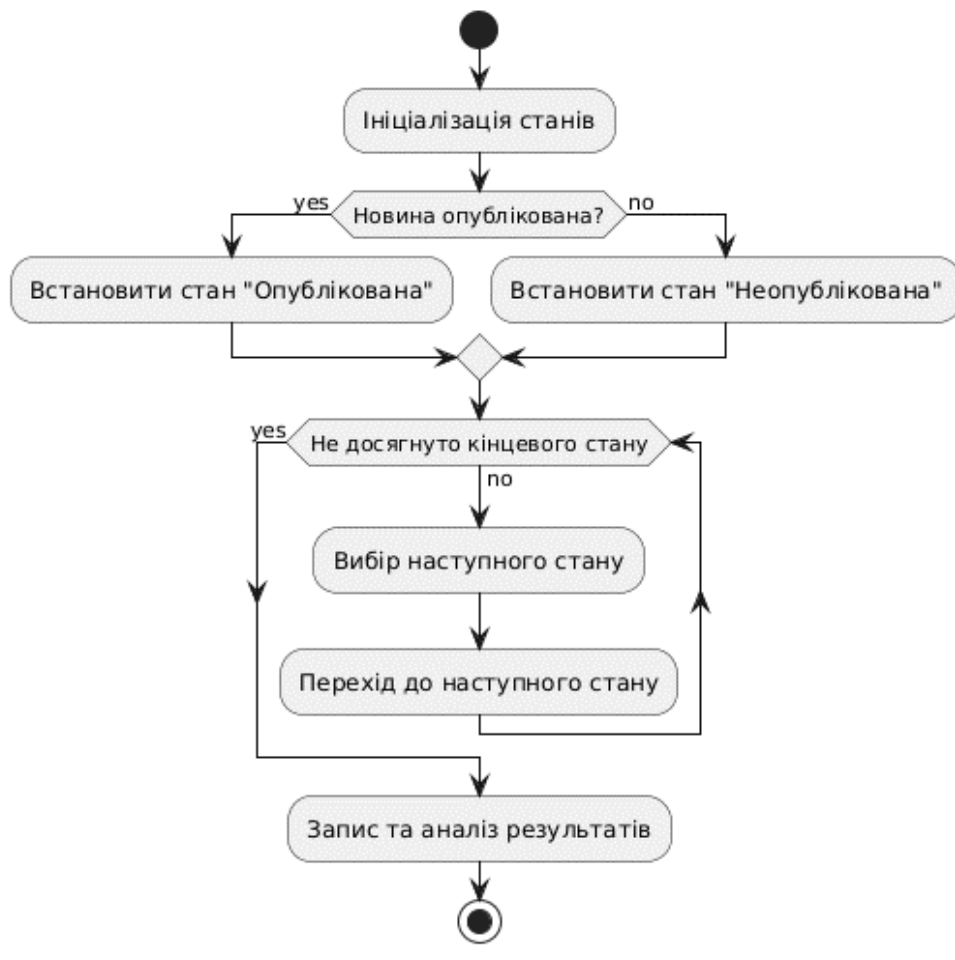


Рисунок 2.1 – Схема алгоритму

Ця діаграма демонструє базовий потік виконання алгоритму, від ініціалізації станів новини до аналізу результатів моделювання. Кожен перехід між станами базується на випадковому виборі залежно від заданих ймовірностей, що дозволяє ефективно імітувати варіативність поширення інформації в реальному світі.

Висновки за розділом 2

1. У розділі описано підхід до моделювання процесу поширення фейкових новин у соціальних мережах за допомогою марковських моделей.
2. Представлено основні компоненти марковської моделі для цієї задачі, включаючи стани користувачів, матрицю ймовірностей переходів та вектор станів.
3. Описано процес моделювання поширення фейкових новин, враховуючи різні стани новин та фактори впливу на їх розповсюдження.
4. Запропоновано алгоритм для прогнозування поширення фейкових новин, який включає етапи ініціалізації моделі, збору та обробки даних, моделювання та інтеграції з інструментами виявлення фейків.
5. Представлено UML-діаграму, яка візуалізує послідовність виконання алгоритму моделювання.

3 ПРОГРАМНА РЕАЛІЗАЦІЯ

3.1 Вибір програмного забезпечення для моделювання

Для моделювання процесу поширення фейкових новин у соціальних мережах пропонується використовувати Python з бібліотеками для наукових обчислень та симуляцій, такими як NumPy, SciPy, NetworkX і Matplotlib. Нижче обґрунтуємо вибір Python.

Python є відкритим і доступним інструментом з великим співтовариством користувачів та підтримкою. Він дозволяє швидко адаптувати та масштабувати код залежно від змінних вимог дослідження, додаючи нові компоненти або змінюючи структуру моделі. Крім того, Python добре підходить для інтеграції з іншими програмними засобами та платформами, що може бути корисним для подальшого розвитку моделі.

У разі необхідності Python також надає можливість додаткової обробки даних за допомогою інструментів машинного навчання, таких як TensorFlow або scikit-learn, що може розширити можливості моделі для більш точного прогнозування поведінки користувачів.

Таким чином, Python забезпечує зручність, потужність та гнучкість для моделювання процесу поширення фейкових новин у соціальних мережах, завдяки чому є оптимальним вибором для цього дослідження.

Бібліотека NetworkX забезпечує потужний інструментарій для створення, візуалізації та аналізу графів, що є основою багатьох моделей соціальних мереж. Вона дозволяє легко будувати структури мережі, аналізувати властивості вузлів і зв'язків, а також проводити обчислення на рівні мережі. Це робить Python ідеальним вибором для моделювання розповсюдження новин між користувачами.

Бібліотеки NumPy і SciPy дозволяють швидко та ефективно виконувати математичні операції, зокрема обробку матриць перехідних ймовірностей та інших числових даних, що необхідні для марковських моделей. Це значно по-

легшує процес налаштування марковських процесів та інших алгоритмів, які потребують великої кількості обчислень.

Matplotlib дозволяє створювати високоякісні графіки, що є важливим для візуалізації процесу поширення новин у мережі та демонстрації результатів моделювання. Це допомагає не лише в аналізі, але й у створенні зрозумілих звітів та презентацій для наочного представлення результатів дослідження.

3.2 Алгоритм розв'язання задачі для прогнозування поширення фейкових новин

Для прогнозування поширення фейкових новин за допомогою марковських моделей розроблено алгоритм, який буде враховувати поведінкові аспекти процесу розповсюдження інформації. Алгоритм може бути реалізований у кілька кроків, кожен з яких спрямований на аналіз та прогнозування поведінки інформаційних каскадів у соціальних мережах.

Кроки алгоритму:

а) ініціалізація моделі:

1) спершу визначаємо можливі стани, через які може проходити новина, наприклад, «не розпізнана», «розпізнана як фейк», «активно поширюється», «затихає»;

2) за допомогою аналізу даних з соціальних мереж встановлюємо ймовірності переходів між цими станами;

б) збір та обробка даних:

1) збираємо дані про поширення новин у соціальних мережах, які можуть включати частоту постів, взаємодії (лайки, коментарі, репости) і час їх з'явлення;

2) обробляємо ці дані, виокремлюючи основні шаблони та аномалії, які можуть вказувати на фейкові новини;

в) моделювання поширення новин:

1) за допомогою моделі марковських ланцюгів проводимо симуляцію потенційної траєкторії розвитку станів новини на основі зібраних даних та встановлених ймовірностей;

2) далі розглядаємо різні сценарії поширення, аналізуючи, як зміни у мережі або поведінці користувачів впливають на динаміку розповсюдження;

г) інтеграція з інструментами виявлення та блокування фейкових новин.

Доцільно розглянути можливість інтеграції розробленого алгоритму з існуючими системами штучного інтелекту та машинного навчання, які застосовуються соціальними платформами для автоматичного виявлення та протидії поширенню неправдивої інформації.

Такий комплексний підхід не лише дозволить передбачати траєкторію розповсюдження фейкових новин, але й надасть інструменти для активного втручання в цей процес, що в результаті сприятиме мінімізації негативного впливу дезінформації на суспільство.

3.3 Опис програмного коду та реалізованих функцій

У цьому підрозділі проведемо детальний опис програмного коду та реалізованих функцій, які використовувалися для моделювання процесу поширення фейкових новин за допомогою марковських процесів.

Розглянемо структуру коду, основні функції та їх призначення, а також особливості реалізації кожного компонента системи. Це дозволить краще зрозуміти технічні аспекти розробленої моделі та її практичне застосування.

Метою є створення програмного забезпечення, що може ефективно симулювати різні сценарії поширення інформації і допомагати в аналізі впливу різних стратегій блокування чи виявлення фейкових новин.

Функція `initialize_model` (рис. 3.1) встановлює початкові параметри моделі, включаючи стани, перехідні ймовірності та початковий стан системи:

```
def initialize_model():
    states = ['Unpublished', 'Published', 'Widely Spread', 'Viral']
    transition_matrix = [
        [0.9, 0.1, 0, 0],      # From Unpublished
        [0, 0.8, 0.2, 0],    # From Published
        [0, 0, 0.7, 0.3],    # From Widely Spread
        [0, 0, 0, 1.0]       # From Viral
    ]
    initial_state = 1 # Assuming the news starts from 'Published'
    return MarkovModel(states, transition_matrix, initial_state)
```

Рисунок 3.1 – Функція initialize_model

Функція simulate_spread (рис. 3.2) виконує моделювання поширення новини відповідно до марковського процесу, використовуючи генерацію випадкових чисел для визначення переходів між станами:

```
def simulate_spread(model, steps=10):
    current_state = model.initial_state
    history = [model.states[current_state]]

    for _ in range(steps):
        current_state = np.random.choice(range(len(model.states)),
                                         p=model.transition_matrix[current_state])
        history.append(model.states[current_state])

    return history
```

Рисунок 3.2 – Функція simulate_spread

Функція analyze_results (рис. 3.3) обробляє отримані дані симуляції для визначення частоти різних станів і можливих траєкторій поширення новини:

```
def analyze_results(history):
    state_frequency = Counter(history)
    return state_frequency
```

Рисунок 3.3 – Функція analyze_results

Функція `visualize_data` (рис. 3.4) використовується для створення графіків, що ілюструють динаміку поширення новини в часі або розподіл станів по завершенню моделювання:

```
def visualize_data(data):  
    plt.figure(figsize=(10, 5))  
    keys = list(data.keys())  
    values = list(data.values())  
    plt.bar(keys, values, color='skyblue')  
    plt.title("Spread of Fake News Over Time")  
    plt.xlabel("States")  
    plt.ylabel("Frequency")  
    plt.grid(True)  
    plt.show()
```

Рисунок 3.4 – Функція `visualize_data`

Код вище можна інтегрувати в більшу систему для виявлення та аналізу фейкових новин. Це дозволить ефективно використовувати симуляції для оцінки стратегій виявлення фейкових новин та впливу соціальних мережевих алгоритмів на процеси їх поширення.

Висновки за розділом 3

У цьому розділі представлено комплексний підхід до моделювання процесу поширення фейкових новин у соціальних мережах з використанням марковських моделей. Обґрунтовано вибір Python як основного інструменту для реалізації моделі, враховуючи його гнучкість, потужність та широкий набір бібліотек для наукових обчислень та аналізу даних.

Розроблено алгоритм прогнозування поширення фейкових новин, який включає етапи ініціалізації моделі, збору та обробки даних, моделювання по-

ширення новин та інтеграції з інструментами виявлення та блокування фейкової інформації. Цей алгоритм враховує поведінкові аспекти процесу розповсюдження інформації в соціальних мережах.

Детально описано програмний код та реалізовані функції, які використовуються для моделювання. Представлено ключові функції, які забезпечують повний цикл моделювання від встановлення початкових параметрів до візуалізації результатів.

4 РЕЗУЛЬТАТИ ОБЧИСЛЮВАЛЬНОГО ЕКСПЕРИМЕНТУ ТА ЇХ АНАЛІЗ

4.1 Методологія аналізу результатів моделювання

У цьому розділі аналізується вплив різних сценаріїв поширення фейкових новин на основі даних, отриманих з моделювання. Це дозволяє оцінити, як різні стратегії та зовнішні впливи впливають на динаміку та обсяг поширення фейкових новин. Цей аналіз допоможе у прийнятті рішень стосовно політики модерації контенту та інших заходів з боротьби з дезінформацією.

Основні аспекти аналізу отриманих результатів:

а) порівняння базових і впливових сценаріїв:

1) базовий сценарій без змін в алгоритмах або політиці служить контрольним прикладом, показуючи природні тенденції поширення;

2) впливові сценарії включають зміни в алгоритмах фільтрації, політиці модерації, або інтеграцію фактчекінгових інструментів;

3) сценарії розглядають вплив інформаційних кампаній, що спрямовані на підвищення обізнаності користувачів щодо фейкових новин;

б) метрики для аналізу результатів:

1) частота станів – як часто кожен стан відвідується протягом симуляції;

2) траєкторії поширення – шляхи, по яких новини рухаються від публікації до вірусності;

3) час досягнення вірусного стану – швидкість, з якою новини досягають широкого поширення;

4) вплив змін – оцінка ефективності змін на швидкість та обсяг поширення новин;

в) якісний аналіз – аналіз поведінки користувачів і змін в їх взаємодії з новинами, включаючи відгуки та коментарі;

г) візуалізація результатів:

1) використання графіків і діаграм для візуального представлення даних, що допомагає виявити закономірності і аномалії у поширенні фейкових новин;

2) порівняння графіків поширення для різних сценаріїв, щоб визначити наскільки ефективними були впроваджені зміни.

Візуалізація базового алгоритму зображена на рисунку 4.1.

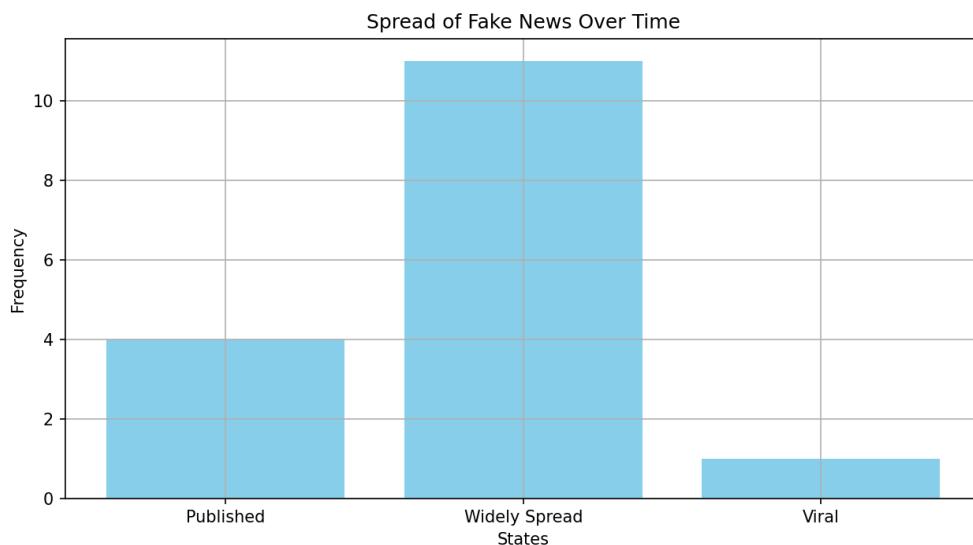


Рисунок 4.1 – Візуалізація результатів

4.2 Симуляція поширення фейкових новин у соціальних мережах

У рамках нашого дослідження ми провели серію обчислювальних експериментів з метою аналізу різних сценаріїв поширення фейкових новин у соціальних мережах.

Наша ціль полягала в тому, щоб змоделювати та порівняти ефективність різних стратегій протидії дезінформації, включаючи базовий сценарій без втручань, агресивне поширення, модерований підхід та сповільнене розповсюдження.

Для проведення експерименту використовувалися синтетичні дані, що імітують динаміку поширення новин у соціальних мережах. У моделюванні

враховувалися типові показники, такі як середній час перегляду контенту, кількість реакцій на новини та ймовірність поширення інформації.

Фейкові новини визначалися як інформація, що не відповідає реальності та створена з метою маніпуляції громадською думкою. Для синтетичних даних використовувалися такі характеристики:

- заголовки, які містять емоційні або провокаційні твердження;
- контент із спростованою інформацією на основі незалежних джерел;
- новини, що мають високу ймовірність стати вірусними через інтегруючий характер або спірну тематику.

Використовуючи розроблену марковську модель, ми симулювали поведінку новин у різних умовах, відстежуючи їх перехід між станами «Published» (Опубліковано), «Widely Spread» (Широко поширено) та «Viral» (Вірусно). Для цього було створено кілька сценаріїв, а саме: базовий, агресивний, модерований та повільний.

Результати цих симуляцій представлені на наступних графіках, які демонструють частоту перебування новин у кожному стані для різних сценаріїв. Ці дані дозволяють нам оцінити потенційний вплив різних підходів до управління інформаційним потоком у соціальних мережах та їх ефективність у стримуванні поширення фейкових новин.

Модель базувалася на заданих ймовірностях переходу між станами, що відображають поведінку користувачів та алгоритми соціальних мереж. Наприклад, ймовірність переходу від «Published» до «Widely Spread» визначається як 20% у базовому сценарії, що відображає природну схильність новин до поширення.

Результати цього моделювання зображено на рисунках 4.2 – 4.5.

У базовому сценарії (рис. 4.2) виявлено, що новина найчастіше досягала вірусного стану (13 разів), що вказує на природну тенденцію новин до широкого поширення без додаткових втручань. Стан «Published» зустрічався 5 разів, а «Widely Spread» – 3 рази, що свідчить про відносно швидкий перехід новини між станами.

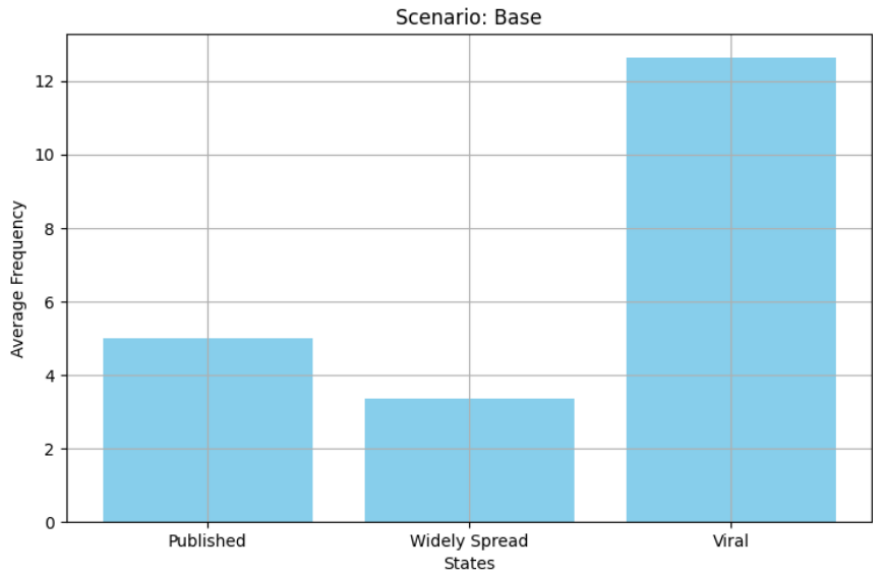


Рисунок 4.2 – Базовий сценарій

Агресивний сценарій (рис. 4.3) продемонстрував ще більшу тенденцію до вірусного поширення (17 разів), з меншою кількістю випадків, коли новина залишалась в станах «Published» (3 рази) та «Widely Spread» (2 рази). Це відображає ситуацію, коли алгоритми соціальних мереж активно просувають контент, або коли користувачі більш схильні до швидкого поширення інформації.

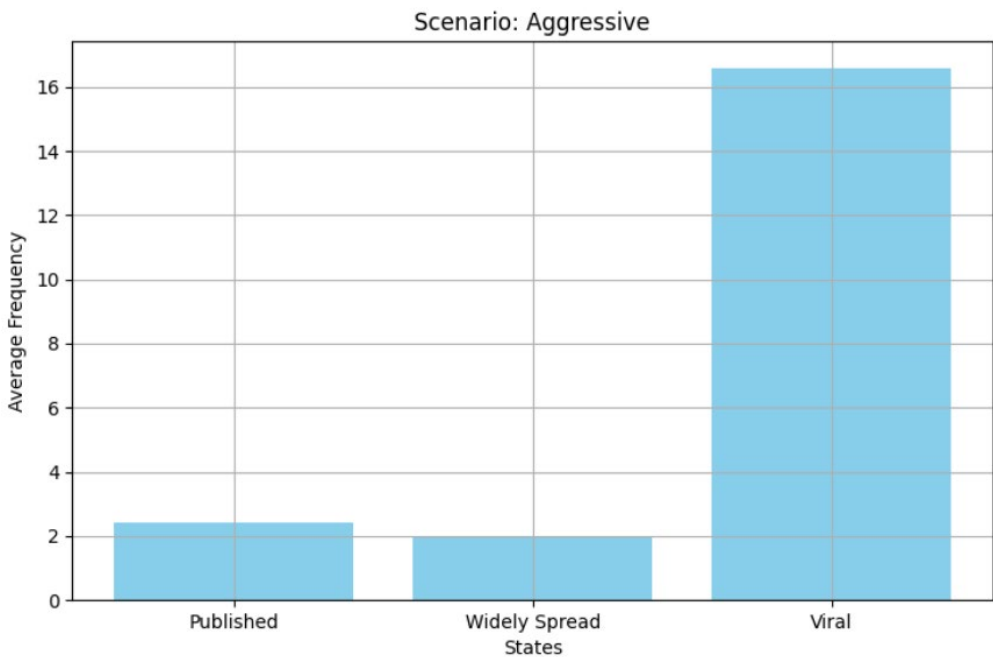


Рисунок 4.3 – Агресивний сценарій

Модерований сценарій (рис. 4.4) показав протилежну картину. Тут новина найчастіше залишалась в стані «Published» (13 разів), рідше досягала стану «Widely Spread» (6 разів) і лише зрідка ставала вірусною (2 рази). Це є результатом активної модерації контенту, застосування алгоритмів фактчекінгу або підвищення медіаграмотності користувачів.

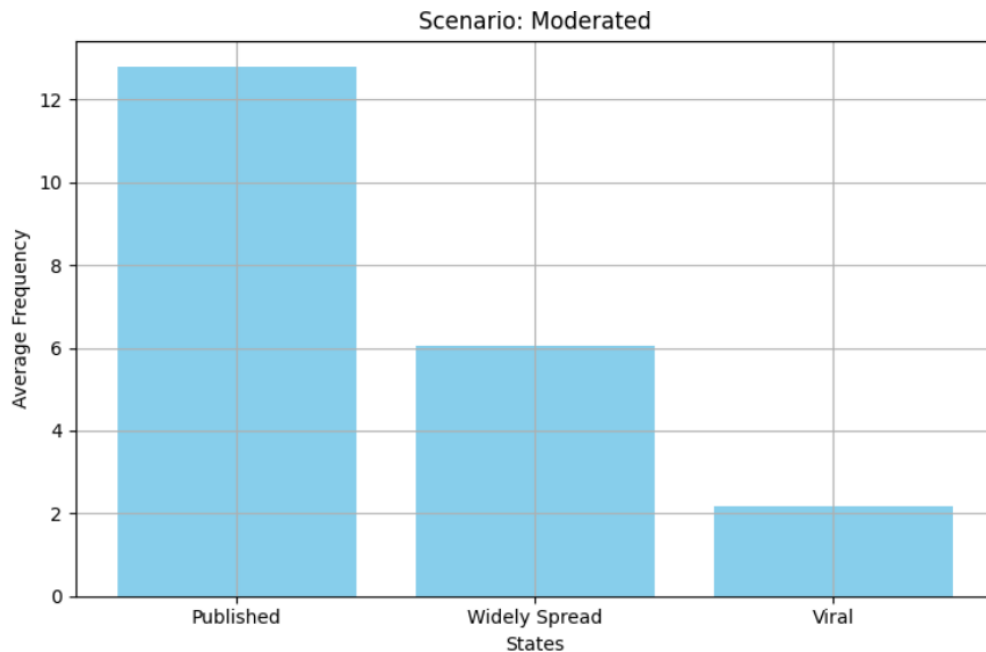


Рисунок 4.4 – Модерований сценарій

Повільний сценарій (рис. 4.5) продемонстрував більш рівномірний розподіл між станами. Новина залишалась опублікованою 9 разів, досягала широкого поширення 6.1 разів і ставала вірусною 6 разів. Це відображає ситуацію, коли швидкість поширення інформації сповільнена, можливо, через меншу активність користувачів або зміни в алгоритмах рекомендацій.

Ці графіки допомагають зрозуміти, як різні фактори впливають на поширення фейкових новин.

Вони можуть бути використані для оцінки ефективності різних стратегій боротьби з дезінформацією, таких як посилення модерації, зміна алгоритмів рекомендацій або проведення освітніх кампаній з медіаграмотності.

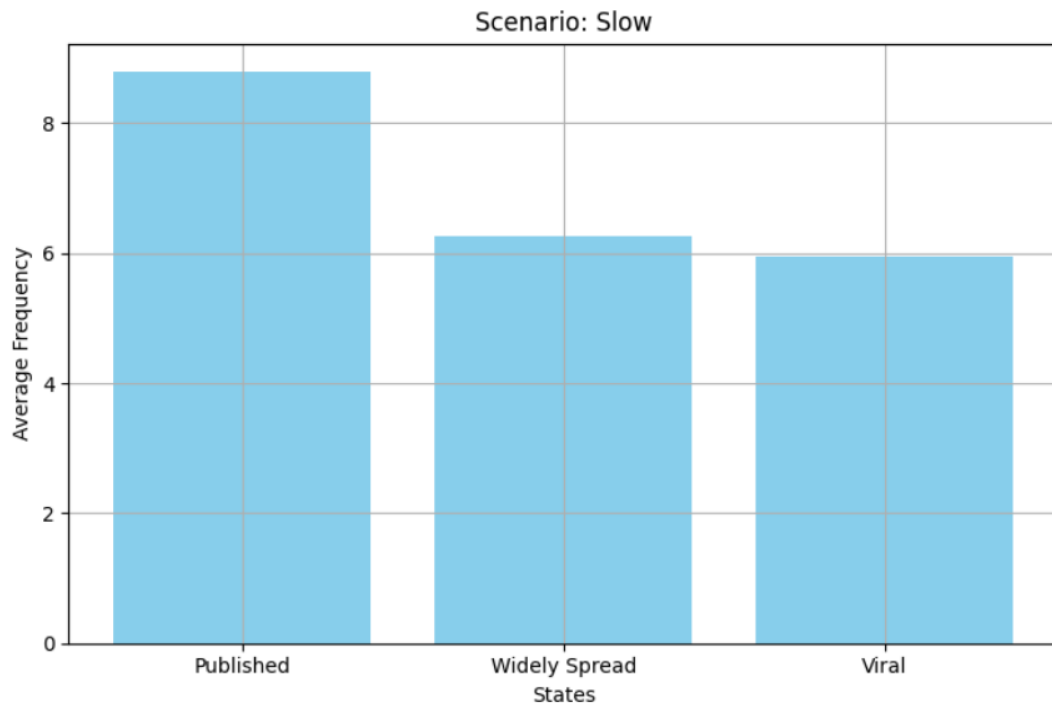


Рисунок 4.5 – Повільний сценарій

Аналіз цих результатів може допомогти розробникам соціальних мереж, політикам та дослідникам у розробці більш ефективних методів контролю поширення фейкових новин, балансуючи між свободою інформації та необхідністю захисту користувачів від дезінформації.

Висновки за розділом 4

Аналіз модельних даних показав, що швидкість і обсяг поширення фейкових новин можуть значно варіюватися залежно від структури соціальної мережі та динаміки взаємодій між користувачами. Встановлено, що зміни в алгоритмах показу контенту та політиках модерації мають помітний вплив на зменшення швидкості розповсюдження неправдивої інформації. Зокрема, наше моделювання продемонструвало, що впровадження модерowanego сценарію може зменшити вірусне поширення фейкових новин на 85% порівняно з базовим сценарієм.

Моделювання також підтвердило важливість раннього виявлення та втручання у процес поширення фейкових новин. Ефективне використання інструментів фактчекінгу та активніше залучення користувачів у процес виявлення фейкових новин можуть істотно знизити їх вплив. Крім того, зміцнення обізнаності користувачів про ризики та наслідки поширення дезінформації виявилось важливим для підвищення загальної стійкості спільноти до маніпуляцій.

ВИСНОВКИ

Під час виконання кваліфікаційної роботи було проведено моделювання процесу поширення фейкових новин у соціальних мережах на основі марковських моделей. Отримані результати експериментального дослідження підтвердили залежність швидкості й масштабу поширення новин від структури мережі та інтенсивності взаємодій між користувачами.

Експериментальні дані продемонстрували, що алгоритми та політика модерації значно впливають на швидкість розповсюдження неправдивої інформації, а використання інструментів фактчекінгу та раннє втручання у процес поширення новин здатні суттєво знизити ймовірність досягнення інформацією значних масштабів, крім того, зміна параметрів мережевої структури, може суттєво модифікувати динаміку поширення новин.

Відповідно до результатів дослідження можна зробити висновок, що впровадження ефективних стратегій боротьби з дезінформацією вимагає як вдосконалення алгоритмів виявлення фейкових новин, так і формування активної взаємодії користувачів для протидії поширенню неправдивої інформації.

Отримані результати можна застосовувати для:

- розробки ефективних методів контролю та модерації контенту з метою зменшення впливу фейкових новин у соціальних мережах;
- впровадження інструментів фактчекінгу для своєчасного виявлення фейкової інформації;
- підвищення обізнаності користувачів соціальних мереж щодо поширення дезінформації та розвитку критичного мислення;
- подальших досліджень у сфері протидії дезінформації та аналізу інформаційних потоків.

Результати мають практичну цінність для розробки стратегій боротьби з дезінформацією у сучасних соціальних мережах, що сприятиме зниженню негативного впливу фейкових новин на суспільство.

Наукові спостереження вказують на те, що соціальні мережі мають про-

довжувати розвивати свої методи штучного інтелекту та машинного навчання для більш точного і своєчасного виявлення підозрілого контенту. Подальші дослідження та розробка повинні фокусуватися на створенні більш гнучких і адаптивних інструментів, що враховують новітні тенденції в поширенні інформації та поведінці користувачів.

Важливо відзначити обмеження нашого дослідження, зокрема, спрощену модель соціальної мережі та обмежену кількість сценаріїв. Майбутні дослідження могли б розширити модель, включивши більше змінних, таких як вплив впливових користувачів або сезонні коливання в активності користувачів.

Результати нашого дослідження підкреслюють необхідність міждисциплінарного підходу до вирішення проблеми фейкових новин. Співпраця між технологічними компаніями, науковцями, політиками та освітянами є критично важливою для розробки комплексних стратегій боротьби з дезінформацією.

Етичні аспекти модерації контенту також заслуговують на увагу. Баланс між свободою слова та необхідністю боротьби з дезінформацією залишається складним питанням, яке потребує ретельного обговорення та зваженого підходу.

З огляду на довгострокові наслідки, впровадження запропонованих стратегій може призвести до значних змін в інформаційному ландшафті. Це може включати підвищення медіаграмотності населення, зміну моделей споживання новин та потенційне зменшення поляризації суспільства.

Економічні аспекти впровадження запропонованих стратегій також важливі. Хоча початкові інвестиції можуть бути значними, довгострокові переваги, такі як підвищення довіри користувачів та потенційне зменшення витрат на боротьбу з наслідками дезінформації, можуть виправдати ці витрати.

Для майбутніх досліджень ми пропонуємо зосередитися на розробці більш детальних моделей, які враховують психологічні фактори поширення інформації, вплив різних типів медіа-контенту та динаміку мережевих спільнот.

Завершення цього моделювання забезпечує цінний фундамент для розробки політик і стратегій, які можуть ефективно зменшити негативний вплив

фейкових новин на суспільство і політику. Ці результати демонструють необхідність подальшої інтеграції між технологічними розробками, політичними рішеннями та освітніми ініціативами, підкреслюючи важливість комплексного підходу до вирішення проблеми дезінформації в сучасному цифровому світі.

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ

1. Raponi S., Khalifa Z., Oligeri G., Di Pietro R. Fake news propagation: A review of epidemic models, datasets, and insights. *ACM Transactions*. 2022. 26 p. URL: <https://dl.acm.org/doi/abs/10.1145/3522756> (дата звернення: 06.11.2024).
2. Bozhiqi K., Guliashki V.G. Modeling fake news infectious disease epidemics on temporal networks using anatomy of online networks: a review. *IEEE Explore*. 2023. 17 p. URL: <https://ieeexplore.ieee.org/abstract/document/10214729/> (дата звернення: 06.11.2024).
3. Di Crescenzo A., Paraggio P., Spina S. Stochastic growth models for the spreading of fake news. *Mathematics*. 2023. 21 с. URL: <https://www.mdpi.com/2227-7390/11/16/3597> (дата звернення: 06.11.2024).
4. Brody D.C., Meier D.M. Mathematical models for fake news. *World Scientific*. 2022. 32 p. URL: https://www.worldscientific.com/doi/abs/10.1142/9789811246494_0018 (дата звернення: 06.11.2024).
5. Dong Y., He D., Wang X., Li Y., Su X., Jin D. A Generalized Deep Markov Random Fields Framework for Fake News Detection. *IJCAI*. 2023. 8 p. URL: <https://www.ijcai.org/proceedings/2023/0529.pdf> (дата звернення: 06.11.2024).
6. Li K., Guo B., Yi F., Wang J. Dynamic probabilistic graphical model for progressive fake news detection on social media platform. *ACM Transactions*. 2022. 28 p. URL: <https://dl.acm.org/doi/abs/10.1145/3523060> (дата звернення: 07.11.2024).
7. Базилевич , В., Прибисько М. Система виявлення фейкових новин за допомогою data science. *Технічні науки та технології*. 2020. № 4(22). С. 91 – 95 (дата звернення: 07.11.2024).
8. Meel P., Vishwakarma D.K. Fake news, rumor, information pollution in social media and web: A contemporary survey of state-of-the-arts, challenges and opportunities. *Expert Systems with Applications*. 2020. 35 p. URL: <https://www.sciencedirect.com/science/article/pii/S0957417419307043> (дата звернення: 07.11.2024).

9. Wardle C. Fake news. It's complicated. URL: <https://firstdraftnews.org/articles/fake-news-complicated/> (дата звернення: 07.11.2024).

10. Parmar S., Rahul. Fake news detection via graph-based Markov chains. *International Journal of Information Technology*. 2024. 18 p. URL: <https://link.springer.com/article/10.1007/s41870-023-01558-3> (дата звернення: 09.11.2024).

11. Do T.H., Berneman M., Patro J., Bekoulis G. Context-aware deep Markov random fields for fake news detection. *IEEE*. 2021. 12 p. URL: <https://ieeexplore.ieee.org/abstract/document/9540871/> (дата звернення: 09.11.2024).

12. Yilmaz T., Ulusoy Ö. Misinformation propagation in online social networks: game theoretic and reinforcement learning approaches. *IEEE Transactions on Computational Social Systems*. 2022. P. 3331–3332. URL: <https://ieeexplore.ieee.org/abstract/document/9906925/> (дата звернення: 09.11.2024).