ДИПЛОМНА РОБОТА
БАКАЛАВРА

ДРБ. ІП - 11.00.00.000 ПЗ

Група ІП-18-1

Яковишин Богдан

2022

Міністерство освіти і науки України
Івано-Франківський національний технічний університет нафти і газу

Інститут інформаційних технологій

Кафедра інженерії програмного забезпечення

УДК 004.4

ДИПЛОМНА РОБОТА БАКАЛАВРА

Тема: " Адаптивні алгоритмічні стратегії інтелектуалізації Web "

(назва згідно з наказом ректора)

Спеціальність – 121 "Інженерія програмного забезпечення"

ПОЯСНЮВАЛЬНА ЗАПИСКА

ДРБ.ІП - 11.00.00.000 ПЗ

(позначення)

Студент

ІП-18-1 /Б.Т.Яковишин/
(шифр групи) (підпис) (дата) (розшифровка підпису)

Керівник роботи

доц., к.ф.-м.н. /Р.І.Храбатин/
(посада) (підпис) (дата) (розшифровка підпису)

Консультанти:

з питань нормоконтролю

доц., к.т.н /В.Я. Піх/
(посада) (підпис) (дата) (розшифровка підпису)

Допускається до захисту
Завідувач кафедри

проф., д.т.н. /В.І.Шекета/
(посада) (підпис) (дата) (розшифровка підпису)

2022
(рік)

Івано-Франківський національний технічний університет нафти і газу

Інститут, факультет інформаційних технологій

Кафедра інженерії програмного забезпечення

Ступінь вищої освіти бакалавр

Спеціальність 121 – Інженерія програмного забезпечення

ЗАТВЕРДЖУЮ:
Зав. кафедрою **ІПЗ**
проф. **В.І. Шекета**
“ ____ ” _____2022 р.

ЗАВДАННЯ

НА ДИПЛОМНУ РОБОТУ БАКАЛАВРА СТУДЕНТОВІ

Яковишину Богдану Тарасовичу

(прізвище, ім’я, по-батькові)

1. Тема проекту (роботи) "Адаптивні алгоритмічні стратегії інтелектуалізації Web"

керівник проекту (роботи) Храбатин Роман Ігорович, доцент

затверджені наказом закладу вищої освіти від **“** _ ” _____ 2022 р. № _____

2. Строк подання студентом проекту (роботи) 23 черня 2022 р.

3. Вихідні дані до проекту (роботи) Результати і матеріали отримані під час проходження переддипломної практики

4. Зміст розрахунково - пояснювальної записки (перелік питань, які потрібно розробити)

1 Огляд існуючих програмних рішень та сервісів в предметній області

2 Алгоритмічні та математичні рішення для задач даного класу

3 Власне оригінальне алгоритмічне та програмне рішення задачі

4 Тестування програмного продукту

5. Перелік графічного матеріалу (з точним зазначенням обов’язкових креслень)

1. Приклад соціальної мережі

2. Огляд підходів до прогнозування конфіденційності

3. Огляд процедури пошуку даних

4. Розподіл ступенів між усіма користувачами мережі

6. Консультанти розділів проекту (роботи)

Розділ	Консультант	Підпис, дата	
		Завдання видав	Завдання прийняв

7. Дата видачі завдання **14 травня 2022р.**

Керівник _____ (підпис)

Завдання прийняв до виконання _____ (підпис)

КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назва етапів дипломного проекту (роботи)	Строк виконання етапів проекту	Примітка
1	Визначення та обгрунтування теми роботи	15.05.2022	виконано
2	Огляд існуючих концепцій, рішень та сервісів в даній області	20.05.2022	виконано
3	Побудова моделі або алгоритму власного рішення	25.05.2022	виконано
4	Документування реалізації власного оригінального рішення вибраними засобами	08.06.2021	виконано
5	Оформлення пояснювальної записки кваліфікаційної роботи	13.06.2021	виконано

Студент _____
(підпис)

Керівник роботи _____
(підпис)

АНОТАЦІЯ

Дипломна робота пов’язана з інтелектом, який може з’явитися з соціальних веб-платформ. Зокрема, явища соціальної конфіденційності та онлайн-переконання визначаються як два стовпи соціального веб-інтелекту, вивчення якого є важливим для просування та розвитку як колективного, так і ділового інтелекту.Враховуючи постійно зростаючу інформацію, яку генерують через різноманітні соціальні мережі, аналітичні дослідження соціальних мереж за останні кілька років активізувалися з усіх верств суспільства. Зокрема, роботи щодо соціального веб-інтелекту сприяють і отримують користь від інтелекту натовпу та намагаються отримати корисну інформацію з таких даних. У формі колективного інтелекту натовпи збираються разом і сприяють вирішенню проблем, які можуть бути складними або неможливими для вирішення окремих людей та окремих комп’ютерів. Крім того, розуміння споживачів, отримане з соціальних слідів, можна використати для створення потужних інструментів бізнес-аналітики, що забезпечують ефективні процеси прийняття рішень.

КЛЮЧОВІ СЛОВА: СОЦІАЛЬНА МЕРЕЖА, ІНТЕЛЕКТ, СОЦІАЛЬНА КОНФІДЕНЦІЙНІСТЬ, ПЕРЕКОНЛИВИЙ ДИСКУРС, ШТУЧНИЙ ІНТЕЛЕКТ

ANNOTATION

The thesis deals with the intelligence that can emerge from social web platforms. In particular, the phenomena of social privacy and online trust are identified as two pillars of social web intelligence, the study of which is important for the promotion and development of both collective and business intelligence. Given the ever-increasing information generated through various social networks, social network analytics research in the last few years have become more active from all walks of life. In particular, works on social web intelligence contribute to and benefit from crowd intelligence and attempt to extract useful information from such data. In a form of collective intelligence, crowds come together and contribute to solving problems that may be difficult or impossible for individual humans and individual computers to solve. In addition, consumer insights gained from social footprints can be used to create powerful business intelligence tools that enable effective decision-making processes.

KEY WORDS: SOCIAL NETWORK, INTELLIGENCE, SOCIAL PRIVACY, PERSUASIVE DISCOURSE, ARTIFICIAL INTELLIGENCE

ЗМІСТ

ПЕРЕЛІК УМОВНИХ СКОРОЧЕНЬ……...7

ВСТУП……8

1 АНАЛІЗ ПРОБЛЕМИ СОЦІАЛЬНОЇ КОНФІДЕНЦІЙНОСТІ

1.1 Конфіденційність у соціальних мережах……..10

1.2 Огляд літератури……13

1.3 Аналіз основних ідей ……19

1.5 Атрибути конфіденційності та профілю в Twitter……...23

1.6 Супутні проблеми що потребують рішення…….……....25

1.7 Джерело даних. Вибір користувача……...27

1.8 Аналіз атрибутів профілю……..31

1.9 Класифікація користувачів…….....35

2 ВИРІШЕННЯ ПРОБЛЕМИ КОНФІДЕНЦІЙНОСТІ НА ПРИКЛАДІ TWITTER

2.1. Основні поняття персоналізації та конфіденційності даних…….37

2.2 Структура наборів даних……45

2.3 Аналіз поведінки користувача……...58

2.4 Визначення параметрів конфіденційності за допомогою спільної фільтрації ……..76

2.5 Модель латентного фактора на основі сусідства ……79

2.6 Атрибути профілю для передбачення конфіденційності ……80

2.7 Графік конфіденційності та приховані атрибути ……... 82

3 РІШЕННЯ ЗАДАЧІ ОБЧИСЛЮВАЛЬНОГО АНАЛІЗУ КОЛЕКТИВНОЇ ІНТЕЛЕКТУАЛІЗАЦІЇ WEB

3.1. Концепція інтелектуалізації ……..86

3.2. Розробка методології ……89

3.3. Представлення результатів тестування …….103

ВИСНОВКИ ……...105

ПЕРЕЛІК ВИКОРИСТАНИХ ДЖЕРЕЛ…...……106

БІБЛІОГРАФІЧНА ДОВІДКА

ПЕРЕЛІК УМОВНИХ СКОРОЧЕНЬ

ШІ – штучний інтелект

БД – База даних

СУБД - Система управління базами даних

ПЗ - Програмне забезпечення

ВСТУП

Поява і постійні досягнення комп'ютерно-опосередкованого спілкування зробили революцію в нашому соціальному житті, вплинувши на те, як ми сприймаємо час, відстань та глобальні кордони. Поява технологій Web 2.0, зокрема, дозволила все більшій кількості людей брати участь у соціальній діяльності в Інтернеті, що призвело до безпрецедентної кількості щоденного створення вмісту та спілкування в Інтернеті. Вікіпедія, форуми питань і відповідей, соціальні мережі та веб-сайти для обміну мультимедійним вмістом – все це приклади таких соціальних платформ. Ця робота в основному присвячена появі інтелекту з таких великих соціальних слідів і вивчає два соціальні явища, які можуть глибоко вплинути на інтелектуальну силу соціальної мережі.

Незважаючи на довгу історію досліджень у різних дисциплінах, стандартного та усталеного визначення інтелекту не існує . У вузькому погляді існує майже стільки ж визначень інтелекту, скільки експертів попросили дати його визначення. Для цього дослідження ми адаптували відоме визначення психолога Говарда Гарднера. Він визначає інтелект як «здатність вирішувати проблеми або створювати продукти, які цінуються в межах однієї чи кількох культурних умов». Соціальна мережа може використовувати два типи інтелекту, які відрізняються насамперед з точки зору культурних умов, для яких вони мають цінність, а саме колективний та бізнес-аналіз.

У своїй традиційній формі колективний інтелект відноситься до інтелекту, який виникає внаслідок локальних взаємодій між окремими людьми. За останні кілька десятиліть соціальні веб-технології уможливили нові форми колективного інтелекту, які дозволяють величезній кількості слабо організованих індивідів взаємодіяти один з одним, вирішувати проблеми та створювати високоякісні інтелектуальні артефакти. Дозволяючи співпрацю та обговорення у масових масштабах, такі соціальні платформи виходять за межі можливостей вирішення проблем невеликої групи 1 влади та може ефективно

використовувати колективну владу великої групи людей з різним походженням та знаннями. Зараз багато соціальних платформ сприяють виникненню та розвитку колективного інтелекту. У Вікіпедії, наприклад, колективні зусилля користувачів зі створення та редагування вмісту привели до створення такої масивної енциклопедії. Внески користувачів базуються один на одному для створення високоякісних проектів з відкритим кодом, таких як Linux. Ринки прогнозів, які дозволяють людям робити ставки на результати таких подій, як президентські вибори, часто призводять до дивно точних результатів.

На додаток до явища колективного інтелекту, підприємства використовують соціальні медіа як важливий компонент своєї бізнес-аналітики наступного покоління, яка може допомогти їм у розробці продуктів, управлінні відносинами зі споживачами, а також цільовому маркетингу та рекламі. Дані соціальних мереж можуть бути дуже цінними для бізнесу в усій галузі, починаючи від власників харчових вантажівок, які оновлюють місцезнаходження своїх вантажівок, до некомерційних організацій, які діляться історіями, які можуть торкнутися сердець мільйонів. Крім того, мільярдні бренди активно інвестують в аналітичні платформи соціальних мереж, щоб постійно спілкуватися з наявними лояльними клієнтами, а також рекомендувати та рекламувати відповідні продукти чи послуги потенційним клієнтам. Наприклад, Apple Inc. придбала аналітичну платформу Twitter Topsy Labs Inc. за понад 200 мільйонів доларів у 2012 році. Гігантська аналітична платформа соціальних медіа під назвою Networked Insight допомагає Samsung, Revlon і Disney приймати ефективні маркетингові рішення. Ford Motor Company використовує соціальні мережі, щоб створити очікування та емоційну прихильність до своїх автомобілів навіть за місяці до їх випуску.

1 АНАЛІЗ ПРОБЛЕМИ СОЦІАЛЬНОЇ КОНФІДЕНЦІЙНОСТІ

1.1 Конфіденційність у соціальних мережах

Соціальні мережі зазвичай представлені у вигляді графіка G = hV, Ei, де кожен користувач відповідає вузлу i ∈ V. Ребро (i, j) ∈ E на графіку вказує на певний соціальний зв'язок між двома користувачами i і j . Функцію маркування F можна визначити як F : V → R, де V — набір зареєстрованих користувачів, а R = {r1,r2,...,rn} — скінченний набір можливих зв'язків, що з'єднують користувачів. Відносини rk можуть бути як двосторонніми (наприклад, стосунки з друзями у Facebook), так і однонаправленими (наприклад, стосунки підписників у Twitter). Крім того, кожен користувач може мати набір властивостей та елементів профілю P = {p1, p2,..., pn}, які вказують, хто є користувачем у соціальній мережі, наприклад, його особистість та особисту інформацію. Користувачі також можуть бути пов'язані з набором вмісту C = {c1,c2,...,cn}, який описує те, що користувач відкриває в соціальній мережі, наприклад, завантажений текст, зображення, відео та інші елементи даних, створені за допомогою різноманітні заходи в мережі.

У більшості сучасних соціальних мереж налаштування конфіденційності описуються з точки зору контролю доступу до спільного профілю та елементів вмісту. Більшість популярних сайтів соціальних мереж, таких як Facebook і Google+, дозволяють користувачам вказувати свої налаштування конфіденційності, контролюючи «хто що бачить» їхніх даних. Різні ступені деталізації специфікації конфіденційності зазвичай підтримуються як для змінних «хто», так і «що». Наприклад, користувачі можуть встановити видимість інформації як публічну, так і приватну; або вони можуть призначати різні специфікації для різних груп своїх соціальних контактів, або навіть різні соціальні контакти окремо. Аналогічно, користувачам може бути дозволено вказувати атрибути конфіденційності для

всіх своїх опублікованих елементів одночасно, різних категорій елементів або кожного елемента спільного доступу окремо. Крім того, контроль доступу до інформації може бути визначений у двійковому значенні (наприклад, дозволити і заборонити) або в номінальному масштабі (наприклад, перегляд, коментар і повторний доступ).

Наприклад, розглянемо часткову мережу користувача, представлену на рисунку 1.1, де фокусний користувач U пов’язаний з трьома соціальними контактами u1,u2,u3 через подібний тип відносин (тобто дружба). Припустимо, що цей користувач має один опублікований елемент профілю p = {p1} і два спільних елементи вмісту c = {c1,c2}. Таблиця 1 представляє можливу специфікацію конфіденційності для користувача U у соціальній мережі, де контроль доступу вказується на двійковому рівні для кожного із соціальних контактів і кожного з опублікованих елементів окремо. Як обговорювалося раніше, управління такими налаштування конфіденційності можуть бути громіздким і виснажливим завданням для сотень соціальних контактів і спільних елементів, тому користувачі часто ігнорують їх.

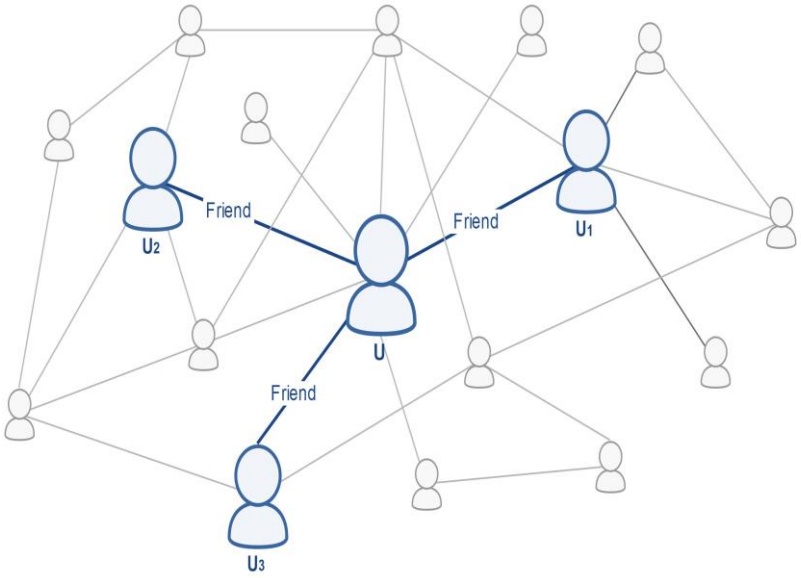


Рисунок 1.1 - Приклад соціальної мережі з основним користувачем U

Таблиця 1.1 - Приклад специфікації конфіденційності в соціальній мережі

	u_1	u_2	u_3
p_1	allow	deny	deny
c_1	allow	allow	allow
c_2	deny	allow	deny

Таким чином, деякі дослідники використовували соціальні сліди в Інтернеті та запропонували напівавтоматичні або автоматичні методи для розробки політики конфіденційності, подібної до того, що надають сучасні соціальні мережі як налаштування конфіденційності. Замість того, щоб зосередитися на конфігураціях конфіденційності, характерних для певного сайту соціальної мережі, низка досліджень намагалася використати такі відбитки для характеристики загальних уподобань користувачів щодо конфіденційності. У цих роботах переваги конфіденційності користувачів можна визначити шляхом відображення користувачів у двійковій, числовій або порядковій шкалі бажаної конфіденційності. Тут ми намагаємося надати вичерпний огляд досліджень, які використовують будь-який із підходів для визначення атрибутів конфіденційності.

1.2 Огляд літератури

Через зростання інтересу до регулярного спілкування та обміну інформацією через онлайн-соціальні мережі, дослідження на цих платформах привернули велику увагу в останні кілька десятиліть. Багата інформація, доступна в соціальних мережах, може принести велику користь окремим особам, громадам і суспільствам, підприємствам, політикам і урядам, а також науковцям. Попередні роботи з автоматичного визначення параметрів

конфіденційності класифікуються на основі типу даних, які використовуються для прогнозування.

Існує велика кількість досліджень, які пов’язують демографічну інформацію, а також особисті риси з уподобаннями щодо конфіденційності. Наприклад, різні опитування встановили позитивне співвідношення віку, освіти та доходу, пов’язане з проблемами приватності. У контексті соціальних медіа були проведені різноманітні дослідження для виявлення можливих зв’язків між статтю та віком та поведінкою щодо приватності. Що стосується статі, результати є непереконливими, оскільки деякі дослідження не виявили гендерних відмінностей у налаштуваннях конфіденційності, тоді як деякі виявили, що користувачі-жінки більш приватні. Подібність, хоча Dey et al. показали, що дорослі, як правило, більш приватні в соціальних мережах, дослідження, проведене Christofides et al. показали, що дорослі та підлітки демонструють подібну поведінку щодо конфіденційності. Подібні спроби також були зроблені для вивчення можливих зв’язків між місцем розташування, а також етнічністю з атрибутами конфіденційності.

Крім того, було показано, що особистісні властивості людей пов’язані з їхнім ставленням до конфіденційності та поведінкою. Наприклад, у контексті послуг на основі розташування, Junglas et al. дослідили можливі зв’язки так званих рис особистості «Великої п’ятірки» (тобто, приємність, екстраверсія, емоційна стабільність, відкритість досвіду та сумлінність) та проблем конфіденційності. Вони виявили, що люди, які набрали високі бали за приємність, сумлінність та відкритість, висловлювали менший рівень занепокоєння щодо конфіденційності.

Незважаючи на велику кількість робіт, які пов’язують демографічні та особистісні характеристики з проблемами конфіденційності, наскільки нам відомо, існує лише кілька робіт, які використовують цей тип інформації для рекомендації функцій конфіденційності. Використовуючи онлайн-опитування, Мінкус і Мемон дослідили налаштування конфіденційності

користувачів Facebook і пов’язали їхній вибір з демографічними та особистісними особливостями. Пізніше результати опитування використовуються для створення та розгортання онлайн-додатка під назвою MyPrivacy, який автоматично рекомендує налаштування конфіденційності. Результати їхнього опитування надають докази для попередніх досліджень, які вказують на те, що риси особистості та демографічні характеристики пов’язані з поведінкою щодо конфіденційності. Зокрема, вони виявили, що нейротизм, вік, етнічна приналежність та занепокоєння щодо конфіденційності пов’язані з налаштованими налаштуваннями конфіденційності користувачів у Facebook. Тому MyPrivacy спочатку ставить кілька запитань від користувачів, щоб визначити ці атрибути, а потім використовує контрольований алгоритм навчання до рекомендованих налаштувань конфіденційності. Оцінка MyPrivacy показала позитивні суб’єктивні думки реальних користувачів Facebook щодо інструменту. Щоб рекомендувати налаштування конфіденційності для певного спільного елемента, також запропонував контрольований метод. Їхній алгоритм побудований на наборі демографічних ознак, включаючи вік, стать та місце розташування; разом із набором метаданих, пов’язаних із спільним елементом. Відсутність підходів, зосереджених на особистісних атрибутах, може бути пов’язана з суперечливими та непереконливими результатами досліджень, які аналізують зв’язок між такими атрибутами та перевагами конфіденційності. Ці суперечливі емпіричні відмінності можуть випливати з відмінностей у тому, що вони вимірюють як переваги конфіденційності. Хоча деякі дослідники можуть вимірювати поведінку конфіденційності, щоб вказати переваги конфіденційності, деякі можуть бути зосереджені на ставленні до конфіденційності. Крім того, у цих роботах процес збору даних часто обмежується конкретними і часто досить невеликими пулами учасників, такими як люди, які живуть у Нью-Йорку або студенти коледжу. Подальші дослідження з великими та різноманітними наборами учасників можуть

привести до послідовних результатів, які можна надійно використовувати в інструментах автоматичного прогнозування. Крім того, демографічні й особистісні атрибути можуть бути недоступні безпосередньо через профілі соціальних мереж, що призводить до наявності невеликого набору атрибутів. Незважаючи на те, що були зроблені успішні спроби витягти цю інформацію з діяльності користувачів у їхній соціальній мережі, ці підходи часто є складними та вимагають додаткових обчислювальних ресурсів.

Порівняно з використанням особистих атрибутів, велика кількість досліджень була зосереджена на соціальному контексті основного користувача, щоб проаналізувати та передбачити особливості, пов'язані з конфіденційністю. Ці дослідження можна розділити на дві основні групи. Перший набір робіт переважно зосереджується на конфіденційності з точки зору видимості інформації для різних груп соціальних контактів, які часто називають соціальними колами. Таким чином, вони пропонують підходи, щоб допомогти користувачам створити та підтримувати такі соціальні кола та відповідну політику конфіденційності.

Хоча вищезгадана група робіт зосереджена на розподілі та кластеризації соціальних контактів користувачів, вони не використовують цінну інформацію, приховану в їхньому соціальному контексті. Таким чином, інша група дослідників адаптувала методи з області спільної фільтрації, щоб призначити політику конфіденційності для користувача на основі переваг інших користувачів. Один із підходів до визначення цієї групи користувачів полягає в тому, щоб вибрати їх із соціальних контактів основного користувача (наприклад, друзів у Facebook або підписників у Twitter). Цей метод дотримується принципу , який відноситься до схильності людей асоціюватися з подібними особами і спостерігається в контексті онлайн-соціальних мереж. Як альтернативу використанню соціальних контактів, ряд дослідників розробив і використав набір заходів подібності, щоб вибрати користувачів зі схожим походженням і характеристиками з основним користувачем.

Враховуючи, що користувачі мають у середньому сотні друзів , визначення політики, яка керує доступом до різних елементів інформації, є складним і виснажливим завданням навіть для користувачів, які стежать за конфіденційністю. В результаті, з метою полегшення процесу управління політикою конфіденційності, були спроби автоматично класифікувати соціальні контакти користувачів у значущі соціальні кола. Деякі дослідження вийшли за межі кластеризації та запропонували методи для визначення бажаних налаштувань конфіденційності користувача для створених кіл контактів.

Аду-Оппонг та ін. запропонував, що алгоритм кластеризації, представлений у [3], може бути використаний для ефективного створення соціальних кіл щільно і тісно пов'язаних контактів в односпрямованих мережах. Дотримуючись цього підходу, (α,β)-кластери будуть сформовані так, що будь-який вузол в кластері сусідить щонайменше з β-часткою кластера, а будь-який вузол за межами кластера сусідить щонайменше з α-часткою кластера. кластер. У дещо схожому підході Денезіс запропонував алгоритм групування своїх соціальних контактів у кола, які тісно пов’язані один з одним і мають багато зв’язків всередині себе, але мають менше зв’язків з тими, хто не входить до кола. У роботі для кластеризації соціальних контактів враховується велика кількість унікальних характеристик, таких як освіта, захоплення та вік. Модифікована версія апріорних алгоритмів використовується для динамічного вибору ознак кластеризації на основі атрибутів соціальних контактів основного користувача.

Джонс і О’Ніл провели дослідження та інтерв’ю користувачів, щоб зрозуміти обґрунтування користувачів під час групування їхніх соціальних контактів з метою управління конфіденційністю. У результаті цього експерименту було визначено набір із шести критеріїв, які зазвичай розглядаються користувачами. Оскільки ці критерії пов’язані зі стосунками

між користувачами, алгоритм кластеризації мережі, який називається SCAN, використовується для групування соціальної мережі в різні кола.

Деякі дослідники запропонували додаткові методи кластеризації, щоб рекомендувати налаштування конфіденційності для створених кластерів. У наприклад, після формування кластерів контактів користувачеві пропонується позначити кількість випадково вибраних контактів з кожного кластера. У процесі маркування користувач вказує на своє бажання поділитися з ним певним предметом. Класифікатор навчається на атрибутах профілю, а також на мережевих атрибутах позначених контактів, щоб передбачити переваги конфіденційності користувача для немаркованих контактів щодо певного об'єкта інформації. Вони досягли точності 83% при 20% навчання.

Fang і LeFevre створили майстер конфіденційності, який ітераційно просить користувача позначати ретельно відібрані інформаційні контакти. У цих питаннях користувач вказує на готовність поділитися певною інформацією профілю з соціальним контактом. Щоб автоматично позначати інші контакти, ця позначена інформація використовується в класифікаторі, де контакти представлені у вигляді векторів функцій, які охоплюють структуру спільноти користувачів і особливості профілю, такі як вік, стать та освіта.

У [2] розроблено класифікатор, щоб вирішити, чи повинен елемент даних бути видимим для контакту користувача. На основі припущення, що мітки конфіденційності, які були явно призначені друзям, правильні, поточні налаштування конфіденційності користувачів використовуються як позначені дані. Подібно до [3], друзі представлені у вигляді вектора ознак, який включає атрибути їхньої спільноти та їхні особисті риси. Потім на основі цього набору функцій створюється класифікатор для призначення міток конфіденційності контактам без міток.

Замість категоризації та маркування соціального контексту деякі дослідники запропонували методи визначення уподобань конфіденційності на основі характеристик конфіденційності соціального контексту. Виявилося, що

поведінка користувачів щодо обміну інформацією знаходиться під великим впливом найближчого кола близьких друзів [9]. Наприклад, кількість приватної інформації, якою поділилися користувачі, як виявилося, корелює з кількістю приватної інформації, якою поділилися друзі. Аналогічно, люди з подібним досвідом, як правило, мають подібні занепокоєння щодо конфіденційності [4]. Ці висновки спонукали дослідників адаптувати спільні методи фільтрації та визначати власні переваги конфіденційності за атрибутами його/її мережі. Спільна фільтрація використовує відомі уподобання групи користувачів для створення рекомендацій щодо невідомих уподобань інших користувачів і в основному використовується в контексті рекомендаційних систем [5].

Squicciarini та ін. надають алгоритм формування соціальних кіл на основі характеристик користувачів, таких як стать, хобі та рід занять. Ці кола також використовуються, щоб рекомендувати політику конфіденційності для нещодавно доданих об’єктів (тобто доданих контактів або завантажених елементів даних). Коли завантажується новий об’єкт, система спочатку шукає соціальні кола, які, швидше за все, мають справу з об’єктом так само, як користувач. Тоді політика конфіденційності, яку використовує вибране коло, є основою для прогнозування політики конфіденційності для щойно доданого об’єкта. Подібна ідея застосовується до зображень, завантажених користувачами, в яких алгоритм прогнозування політики призначає політику щойно завантаженому зображенню на основі інформації, отриманої від соціальних кіл.

У [4] активне навчання та властивості соціального графа вперше використовуються для виявлення набору найбільш інформативних контактів, які мають бути позначені як навчальні зразки. У процесі маркування користувач вказує, чи бажає він/вона поділитися певним елементом даних із вибраним контактом. Потім застосовується ітераційний напівконтрольований підхід для позначення інших контактів користувача, де мітки поширюються

від позначених екземплярів до немічених екземплярів у соціальному графі. Це розповсюдження керується показником подібності користувача, який представлений через ваги краю. Обчислення подібності базується на інформації про профіль контактів, метриках їх мереж, а також на структурі спільноти. Результати оцінювання цього підходу забезпечили більш високу точність і точність у порівнянні з навчанням під наглядом і підходом на основі випадкового блукання.

У контексті соціальної мережі, заснованої на місцезнаходження, Toch et al. [5] надають користувачам рекомендовану політику конфіденційності, яку раніше обрали подібні користувачі. Великий набір політик конфіденційності спочатку групується на основі їх розташування, часу та властивостей соціальної групи. Політики в кожному кластері потім ранжуються відповідно до кількості політик, на які вони подібні, і ступеня їх схожості. Щоб рекомендувати політику конфіденційності, кластери, які стосуються поточного користувача, вибираються на основі політики, обраної подібними користувачами (наприклад, користувачами, які перебувають у тій самій мережі Facebook). Нарешті, як рекомендації представлені політики з найвищим рейтингом вибраних кластерів.

Спільна фільтрація також дотримується в [6], де автори використовують переваги набору функцій профілю, інтересів користувачів, які мають на увазі в їхніх соціальних мережах, та їх конфігурації конфіденційності, щоб знайти набір користувачів, подібних до користувача, який у фокусі уваги. У їхньому підході користувачі спочатку характеризуються відповідно до їхніх уподобань щодо конфіденційності як фундаменталістів конфіденційності, прагматиків конфіденційності або байдужих до конфіденційності. Рішення щодо конфіденційності та налаштування користувачів щодо їхніх фотоальбомів розглядаються як ознака їхньої конфіденційності. Зокрема, користувачів розподіляють до цих трьох категорій на основі кількості їхніх загальнодоступних, налаштованих та

приватних фотоальбомів. Потім використовується алгоритм K-найближчого сусіда, щоб визначити, до якої категоризації конфіденційності належить фокусний користувач. На основі функцій призначеної категорії система рекомендує налаштування конфіденційності.

Частою діяльністю користувачів у соціальних мережах є публікація та обмін вмістом, таким як повідомлення про статус, коментарі, зображення та відео. Усі екземпляри спільних типів даних можна використовувати, щоб зробити висновки про особистість та уподобання користувачів. Зокрема, показано, що природна мова є відображенням і посередником внутрішніх станів [8]. Наші слова можуть розкривати особистість, емоційні стани та почуття, моделі уваги, думки та соціальні ситуації. Тому для вимірювання таких психометричних показників природною мовою було розроблено різноманітні методи автоматизованого аналізу контенту. Ці методи варіюються від використання попередньо визначених словників і таксономій, таких як мовний запит і підрахунок слів (LIWC), до більш складних обчислювальних алгоритмів, які використовують складні методи аналізу даних і машинного навчання.

У контексті конфіденційності Gill et al. надають набір категорій, пов'язаних з конфіденційністю, кожна з яких пов'язана з низкою слів, які є релевантними для семантичного аналізу домену конфіденційності. Словник складається з 388 слів, пов’язаних із конфіденційністю, які згруповані у вісім високорівневих теоретично обґрунтованих категорій на основі їх семантичної схожості. LIWC містить велику кількість семантичних категорій, які можуть мати відношення до функцій конфіденційності. Тому LIWC використовується як базова лінія для оцінки словника конфіденційності. Результати оцінки показують, що словник конфіденційності здатний охопити унікальні лінгвістичні особливості мови конфіденційності та є більш надійним у виявленні вмісту, орієнтованого на конфіденційність.

Каліскан-Іслам та ін. [9] використовував словник конфіденційності разом із різноманітними методами та інструментами, включаючи моделювання тем, розпізнавання іменованих об’єктів та аналіз настроїв, щоб автоматично визначити, чи містить твіт особисту інформацію. Анотовані дані були зібрані з Amazon Mechanical Turk (AMT), де працівників AMT попросили позначити зібрані твіти відповідно до категорій конфіденційності. Потім користувачі отримують оцінки конфіденційності на основі кількості приватної інформації, яку вони опублікували у своїй хронології Twitter. Часові шкали цих позначених користувачів потім використовуються в техніці машинного навчання під наглядом, щоб призначити оцінки конфіденційності користувачам без міток на основі їх спільного текстового вмісту.

Модель прогнозування, запропонована в [6], дотримується підходу машинного навчання з наглядом, щоб рекомендувати налаштування конфіденційності для певної публікації у Facebook. Окрім демографічних характеристик, вони використовували набір функцій на основі вмісту, пов’язаних із публікацією. Включається оцінка настроїв до публікації, а також деякі актуальні атрибути. Крім того, береться до уваги весь пакет слів, де розглядається лише набір слів з високим балом tf-idf. Використовуються також деякі контекстні елементи метаданих, наприклад час доби, до якого поширюється публікація.

Враховуючи неструктурований лінгвістичний контент, опублікований користувачем, [8] спочатку виявляє конфіденційну інформацію, таку як номер телефону, адреса та місцезнаходження, у тексті. Тоді для кількісної оцінки ризику конфіденційності користувача приймається модель, запропонована в [9], де ідентифіковані чутливі частини обробляються так само, як інформаційні елементи в [9]. У [2] розроблено математично обґрунтовану модель з урахуванням чутливості та видимості спільних елементів. Запропонована модель надає користувачам оцінку конфіденційності, яка кількісно визначає потенційний ризик конфіденційності користувача. Передумовою їхньої моделі

є те, що чим більш конфіденційну інформацію розкриває користувач, тим вище ризик його конфіденційності. Крім того, чим більш помітною стає спільна інформація в мережі, тим вище ризик конфіденційності.

У контексті обміну зображеннями [4] використовує попередні зображення користувачів і відповідні їм політики конфіденційності, щоб призначити політику конфіденційності новому зображенню. Для створення конфіденційності використовуються кластеризація зображень і аналіз асоціацій політик. Однак, якщо користувач новий або відбулися значні зміни в тенденціях щодо конфіденційності користувача, для прогнозування політики використовується соціальний контекст користувачів, як пояснюється в Розділі Щоб виявити зображення з приватним вмістом, Zerr et al. [5] використовували різноманітні візуальні особливості, такі як наявність облич, в алгоритмі навчання під керівництвом. Вони також використали текстові метадані, пов’язані із зображеннями, і виявили кореляції між темами та вмістом приватних зображень. Наприклад, було показано, що теми, які використовуються для опису особистих концепцій, емоцій та настроїв, а також людського тіла, здебільшого використовуються для особистих зображень. З іншого боку, теми, пов’язані з природою, архітектурою та неживими об’єктами, здебільшого зустрічаються в неприватних зображеннях.

1.3 Обговорення результатів

Автоматичному висновку про конфіденційність приділено відносно мало уваги. Крім того, наш аналіз відповідної літератури показує, що зусилля дослідження в основному зосереджені на передбаченні та рекомендації налаштувань конфіденційності, які є специфічними для базової соціальної мережі. З іншого боку, були зроблені спроби кількісно оцінити ризики конфіденційності користувачів [2, 6] та поточний рівень конфіденційності користувачів [9]. Газінур та ін. [6] характеризує людей, класифікуючи їх за

різними рівнями проблем конфіденційності на основі їхніх налаштувань конфіденційності в їхніх фотоальбомах. Ця відсутність роботи може бути пов'язана з тим, що конфіденційність залежить від контексту [7], що ускладнює розробку загальних методів. Однак нещодавнє дослідження в контексті мобільних додатків показало, що, незважаючи на різноманітність уподобань конфіденційності, користувачів можна об'єднати в набір значущих профілів конфіденційності, які ефективно фіксують їхню бажану конфіденційність [7]. Ці дослідження передбачають можливість характеристики користувачів відповідно до їхніх уподобань щодо конфіденційності, незалежних від платформи.

Ще один недолік у дослідженні, який ми виявили, — це обмежений набір типів даних, на яких були зосереджені попередні дослідження. Такими типами даних часто є функції профілю користувачів, атрибути соціальних і мережевих мереж, а також вміст комунікацій. Багато інших типів даних залишаються невивченими в літературі [4]. Наприклад, рейтинги/інтереси користувачів можуть бути корисними для розуміння власних уподобань конфіденційності та прихованих атрибутів, і часто легко доступні в профілях користувачів. Однак, наскільки нам відомо, лише [6] використав його для висновку про конфіденційність. Інший приклад – контекстні дані. Цей тип даних посилається на властивість елемента, яка стає явною та надається семантикою, як-от теги, надані в зображеннях Facebook і повідомленнях про статус або згадки в твітах. Хоча налаштування конфіденційності користувачів можуть бути виявлені з цих контекстних даних, дослідники не використовували їх для виявлення функцій конфіденційності.

Аналізуючи використані підходи, можна побачити, що в існуючій літературі бракує вивчення гібридних методів та змішаних типів даних. Однак

у подібних областях, таких як рекомендаційні системи, гібридні підходи були продемонстровані як дуже ефективні та можуть компенсувати обмеження будь-якого підходу та покращити ефективність прогнозування [5].

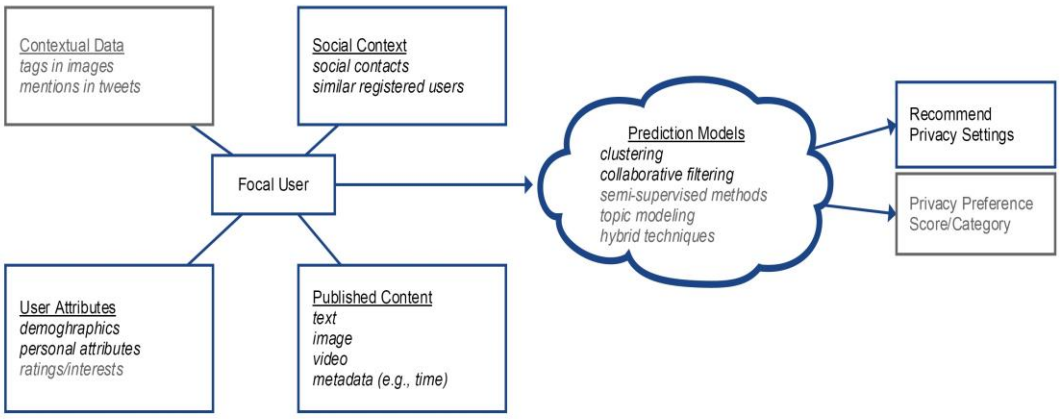


Рисунок 1.2 - Огляд підходів до прогнозування конфіденційності з точки зору їх входу, методів та цілей.

Багато з розглянутих досліджень використовували контрольовані методи для класифікації та прогнозування атрибутів конфіденційності. Однак вони вимагають маркованого введення, і можуть здатися нездійсненними в контексті соціальних мереж, де позначена інформація зазвичай становить дуже малу частину доступних даних. У такому контексті неконтрольовані та напівконтрольовані методи можуть представляти великий інтерес. Зокрема, напівконтрольовані методи, які мають перевагу використання меншої кількості мічених даних для досягнення кращих прогнозів, можуть бути потенційним напрямком дослідження для подальшого вивчення. Напівконтрольований метод на основі графіків був запропонований для ефективного визначення переваг конфіденційності [4], а інші методи, такі як очікування та максимізація (EM), тематичне моделювання та спільне навчання, потребують подальшого дослідження. Наприклад, спільне навчання успішно використовується для виявлення прихованих особистісних атрибутів користувачів у соціальних мережах [5].

На рисунку 1.2 наведено огляд наших розглянутих досліджень з точки зору їх вкладу, запропонованих методів, а також їх цілей і цілей. Також наведено приклади кожного з цих елементів. На рисунку також показано

прогалини в дослідженні, які ми обговорювали вище, сірими рамками. Обговорені обмеження попередніх досліджень вимагають подальших спроб глибоко проаналізувати, як різні аспекти великого соціального впливу в Інтернеті можуть бути використані для ефективної характеристики уподобань користувачів щодо конфіденційності.

Видобуток скарбів, які існують у соціальних мережах, має величезний потенціал для компаній, щоб покращити досвід клієнтів за допомогою персоналізації та цільового маркетингу. Однак клієнти можуть не захотіти бути профільованими в Інтернеті через свої проблеми з конфіденційністю.

З іншого боку, проблеми конфіденційності користувачів часто погано втілюються в конфігурацію конфіденційності їх соціальних мереж, що призводить до створення даних, які є доступними для громадськості. Для компаній виникає дилема: використовувати свої загальнодоступні дані для надання цінного персоналізованого вмісту чи не використовувати такі дані, щоб уникнути нехтування перевагами конфіденційності. Одним із потенційних напрямків вирішення цієї дилеми є розробка нових алгоритмів і методів, які використовують переваги соціального впливу користувачів і характеризують їх поведінку та ставлення до конфіденційності. У цьому документі ми розглянули існуючу літературу щодо автоматичного визначення параметрів конфіденційності в контексті соціальних мереж.

Ми класифікували та переглянули існуючі дослідження щодо визначення переваг конфіденційності відповідно до типу даних, а саме: демографічні характеристики та особливості профілю, соціальний контекст та особливості мережі, а також спільний вміст. Потенціал та обмеження підходів далі обговорюються, де в літературі виявлено ряд прогалин. Виходячи з нашого вивчення літератури, ми закликаємо до додаткових досліджень загального моделювання та характеристики користувачів відповідно до їхніх уподобань щодо конфіденційності. Крім того, дослідникам, які вивчають виявлення конфіденційності, рекомендується використовувати більш

широкий спектр типів даних, доступних у соціальних мережах, а також гібридні методи для прогнозування.

1.5 Атрибути конфіденційності та профілю в Twitter

Зростання рівня участі в онлайнових соціальних мережах призвело до накопичення великих соціальних слідів, які залишають мільйони користувачів щодня. Це величезне джерело інформації може надати релевантну інформацію в правильному контексті, відкриваючи величезні можливості як для бізнесу, так і для окремих осіб. Однак, щоб ефективно використовувати цю скарбницю даних, необхідно вирішити можливі ускладнення з конфіденційністю. Такі проблеми конфіденційності викликають особливе занепокоєння через невідповідність між поведінкою користувачів щодо конфіденційності та їхнім ставленням у соціальних мережах [6], що робить їхні поточні налаштування конфіденційності ненадійними. Таким чином, потрібні методи, які можуть визначити параметри конфіденційності користувачів, щоб можна було відкинути дані, пов’язані з користувачами, які піклуються про конфіденційність.

Ця дихотомія конфіденційності може бути пов’язана з помилковими уявленнями користувачів щодо видимості їхніх даних [2], складними інтерфейсами специфікації конфіденційності [4], а також їх хибним, але поширеним сприйняттям налаштувань за замовчуванням як рекомендованої політики конфіденційності [2]. Крім того, навіть користувачі, які знають конфіденційність, можуть вирішити вибрати загальнодоступні налаштування для очікуваної соціальної вигоди, в той час як вони можуть не захотіти бути профільованими в Інтернеті бізнесом і компаніями.

Ми стверджуємо, що соціальні сліди користувачів [7] у середовищі соціальних медіа можуть характеризувати їх уподобання щодо конфіденційності, пропонуючи альтернативне та надійне джерело для

виявлення уподобань конфіденційності. Соціальні сліди доступні в трьох типах даних соціальних мереж: атрибути профілю користувачів, їхній соціальний контекст та зв'язки, а також опублікований вміст. У цьому дослідженні ми зосередимося на аналізі атрибутів профілю, щоб вивчити їх потенційні зв'язки.

Версія цієї глави була опублікована в супровідному томі конференції World Wide Web (WWW MSM'6).29 налаштування конфіденційності користувачів. Зокрема, ми проаналізували атрибути профілю облікових записів Twitter, щоб перевірити, чи люди з різними рівнями конфіденційності налаштовують ці атрибути по-різному.

Налаштування конфіденційності в Twitter є відносно простим і відповідає бінарній специфікації. Користувачі Twitter можуть дотримуватися загальнодоступних налаштувань за замовчуванням, які вказують на те, що їхні твіти та списки підписників/друзів доступні для всіх. Крім того, вони можуть змінити налаштування на захищені, що робить їхні твіти та списки підписників/друзів доступними лише для їхніх затверджених підписників. Примітно, що атрибути профілю користувачів видимі та доступні для громадськості та API Twitter незалежно від їх налаштувань конфіденційності.

Ми проаналізували набір функцій та описів профілів користувачів, які легко доступні в їхніх акаунтах у Twitter. Ми також розробили та проаналізували три додаткові функції на основі наявних атрибутів профілю. На основі результатів аналізу розроблено набір функцій і використовується в кількох класифікаторах для автоматичного виявлення користувачів із захищеними налаштуваннями конфіденційності. У порівнянні зі структурою соціальних мереж користувачів та їхніми функціями, пов'язаними з вмістом, їхні профілі Twitter містять дуже обмежену інформацію. Незважаючи на це обмеження, наш класифікатор отримав F-бал 0,71, що покращує випадкову та наївну базову лінію більш ніж на 20%. Цей висновок може мати значення для розробки інструментів персоналізації, що зберігають конфіденційність, і

вказує на значення атрибутів профілю при виявленні користувачів, які піклуються про конфіденційність.

Виявлення атрибутів конфіденційності в соціальних мережах. Для вирішення проблем конфіденційності користувачів у літературі повідомлялося про три основних підходи: використання принципів підвищення конфіденційності для розробки систем персоналізації, проектування інтерфейсів і візуалізацій, які дозволяють користувачам визначити свою політику конфіденційності, а також обчислювальні методи, які автоматично прогнозують переваги конфіденційності користувачів . Незалежно від його потенційної цінності, підходу прогнозування приділено набагато менше уваги порівняно з двома іншими [8]. Крім того, більшість моделей прогнозування структуровано в соціальних мережах користувачів та їх створеному контенті, тоді як лише
деякі з них використовують профіль

Мінкус і Мемон [3] провели онлайн-дослідження анкети, щоб вивчити налаштування конфіденційності користувачів Facebook і пов’язали налаштування з їхніми демографічними та особистісними особливостями. Пізніше їх результати опитування були використані для створення та розгортання онлайн-додатка під назвою MyPrivacy, який автоматично рекомендує налаштування конфіденційності. Зокрема, MyPrivacy спочатку ставить кілька запитань від користувачів, щоб визначити демографічні та особистісні атрибути, а потім використовує контрольований алгоритм машинного навчання, щоб давати рекомендації на основі налаштувань конфіденційності користувачів. Оцінка MyPrivacy показала, що реальні користувачі Facebook мали позитивні суб’єктивні думки щодо інструменту. Однак цей інструмент є напівавтоматизованим і вимагає безпосереднього введення від користувачів.

Аналогічно, в [4] запропоновано алгоритм навчання з наглядом, побудований на великому наборі функцій, щоб рекомендувати налаштування

конфіденційності. Ці функції включають елементи метаданих щодо спільного елемента, а також демографічні характеристики та характеристики профілю користувачів, наприклад кількість публікацій користувачів у Facebook та їхніх друзів. Варто зазначити, що алгоритм розроблено для того, щоб рекомендувати налаштування конфіденційності для певного спільного елемента, наприклад окремих дописів у Facebook, на відміну від прогнозування налаштувань конфіденційності користувачів загалом.

В [3] запропонував модель прогнозування конфіденційності, яка враховує поведінкові аналоги соціальних медіа до психологічних змінних, які, як відомо, впливають на поведінку користувачів у розкритті інформації. Деякі з виявлених аналогів засновані на особливостях профілю користувачів. Наприклад, надійність користувача розраховується на основі співвідношення його підписників до загальної кількості їхніх соціальних контактів.

Виявлення атрибутів користувача Twitter. Для виявлення прихованих атрибутів користувачів Twitter використовувалися різноманітні методи та типи соціальних даних. Наприклад, використовує поля профілю, поведінку твітів, вміст твітів, а також структуру мережі, щоб зрозуміти політичну приналежність, етнічну приналежність, а також спорідненість користувачів із певним бізнесом.

Іншим прикладом є робота Nguyen et al. [5], в якій у контексті Twitter досліджується зв’язок між вживанням мови та віком. Автори створили класифікатор на основі уніграм твітів і класифікували користувачів за віковою категорією, життєвим етапом та віком. У своєму дослідженні Rao et al. [7] виявили відмінні варіації у мовленні користувачів Twitter різної статі, віку, регіонального походження та політичної орієнтації.

Ці дослідження щодо автоматичного визначення налаштувань конфіденційності та атрибутів користувачів Twitter дають цінне уявлення про можливості та обмеження доступних функцій і методів у виявленні прихованих атрибутів користувачів. Наша мета використання атрибутів

профілю для визначення загальних уподобань користувачів щодо конфіденційності приділяється дуже мало уваги в літературі [3]. Хоча стверджується, що поля профілю користувачів можуть не включати достатньо якісної інформації для цілей класифікації користувачів [6], наші результати свідчать про те, що вони можуть бути перспективними для виявлення соціальної конфіденційності.

1.7 Джерело даних. Вибір користувача

Ми створили каталог користувачів Twitter, скануючи кілька відомих акаунтів Twitter і збираючи їхніх підписників. У таблиці представлено ці облікові записи Twitter, кількість зібраних підписників кожного облікового запису, а також відсоток кількості захищених підписників до кількості зібраних підписників облікового запису. Зверніть увагу, що в таблиці числа округлені (наприклад, кількість підписників CNN становить 12 246 514, а ми округлили її до 12,2 млн), а відсотки розраховуються на основі точних чисел, а потім округляються.

Відомо, що велика кількість облікових записів Twitter неактивні, тому, швидше за все, не підтримують жодного облікового запису [1]. Крім того, такі облікові записи з більшою ймовірністю відповідатимуть загальнодоступним налаштуванням конфіденційності за замовчуванням порівняно з активними обліковими записами. Наш базовий набір облікових записів Twitter слідує принаймні одному обліковому запису. Тому очікується, що відсоток захищених облікових записів буде вищим, ніж облікових записів Twitter загалом. У середньому 4,8% користувачів Twitter мають захищені акаунти [1]. Як видно з таблиці 1.2, відсоток захищених облікових записів подібний або вище середнього для всіх наших груп підписників, що підтверджує наші очікування.

Як показано в таблиці, відсоток захищених облікових записів для «Пошукових новин CNN» становить 11%, що значно вище, ніж у інших груп підписників і середній відсоток у Twitter. Одна з можливостей полягає в тому, що твіти CNN можуть охоплювати більше тем, які приваблюють приватних користувачів у Twitter, наприклад, новини, пов'язані з конфіденційністю, порівняно з іншими обліковими записами в таблиці. Це також може бути справедливим і для інших акаунтів новин Twitter, оскільки ми очікуємо, що новинні акаунти охоплюватимуть більш широкий спектр тем, які пов'язані з та мають потенційний вплив на життя та суспільства користувачів. Очікується, що подальший аналіз відсотка захищених облікових записів Twitter, які слідують за іншими акаунтами новин у Twitter, та аналіз новинного вмісту дадуть більше розуміння цього питання.

Як зазначалося у вступному розділі, налаштування конфіденційності користувачів можуть не відображати їхні фактичні налаштування конфіденційності [6]. Іншими словами, користувачі Twitter, які мають загальнодоступні налаштування конфіденційності за замовчуванням, можуть бути приватними. Таким чином, атрибути профілю цих користувачів відображають конфігурації приватних користувачів. Щоб вивчити можливі відмінності в налаштуваннях захищених і загальнодоступних профілів, нам потрібен набір облікових записів, у яких така ситуація мінімізована.

Таблиця 1.2 - Набір популярних облікових записів у Twitter і статистика їх зібраних наборів підписників.

Account	#Followers	#Protected	%Protected
Facebook	4.8M	261K	5%
CNN Breaking News	12.2M	1.5M	11%
Youtube	14.8M	788K	5%
Bill Gates	5M	374K	7%
Obama	23.8M	52M	7%

Щоб гарантувати, що неактивні користувачі не враховуються, ми відфільтрували список підписників CNN, щоб включити лише тих, хто опублікував принаймні десять твітів. У Twitter облікові записи, які належать ключовим особам і брендам, позначаються як підтверджені. Щоб зосередитися на широкому загалі, ми видалили ці підтверджені облікові записи з набору. Нарешті, ми відфільтрували набір, щоб включити лише ті облікові записи, мова яких вибрано англійською. Застосовуючи ці три критерії, ми змогли вибрати приблизно 850 тисяч захищених облікових записів із початкових 1,5 мільйонів облікових записів. Наші публічні облікові записи також скоротилися з 12,2 млн до майже 10 млн. Щоб мати відносно збалансований набір облікових записів, ми випадковим чином витягли 1 млн публічних облікових записів із цього набору.

Кожен обліковий запис Twitter пов’язаний з набором атрибутів профілю. Набір функцій профілю налаштовується власником облікового запису, часто під час створення облікового запису, і призначений для представлення того, хто є користувачем у мережі. Прикладами таких атрибутів профілю є ім’я користувача, зображення профілю та інформація про місцезнаходження. Інший набір атрибутів, який також вказує користувач, пов'язаний з налаштуваннями його облікового запису. Наприклад, вони можуть вказати, чи хочуть вони, щоб їхні твіти позначалися географічними тегами, встановивши значення атрибута geo-enabled. Інші приклади таких атрибутів включають їхню бажану мову інтерфейсу та те, чи слід закрити їхній обліковий запис у певних країнах. Нарешті, Twitter визначає набір контекстних атрибутів. Наприклад, час створення облікового запису, кількість твітів, опублікованих кожним користувачем, і кількість підписників/друзів.

Підмножина доступних атрибутів профілю вважається релевантною для нашої мети та обраною в нашому аналізі. Цей список наведено нижче разом із коротким описом атрибутів. Зверніть увагу, що описи адаптовані зі специфікації API Twitter.

• Ім'я: ім'я користувача.

• Ім'я користувача: псевдонім, за яким ідентифікують себе користувачі.

• Опис: фрагмент тексту, який користувачі надають для опису свого облікового запису.

• URL: URL-адреса, надана користувачем у зв'язку з його профілем.

• Місцезнаходження: визначене користувачем місцезнаходження профілю цього облікового запису.

• Geo-enabled: якщо значення true, це означає, що користувач увімкнув можливість позначення геотегів своїх твітів.

• Зображення за замовчуванням: якщо значення true, це означає, що користувач не завантажив своє зображення, а замість нього використовується аватар за замовчуванням.

• Профіль за замовчуванням: якщо значення true, це означає, що користувач не змінював тему або фон свого профілю користувача.

• Кількість обраних: кількість твітів, доданих цим користувачем у вибране за весь час існування облікового запису

• Кількість твітів: кількість твітів, виданих користувачем.

• Кількість підписників: кількість підписників на даний обліковий запис.

• Кількість друзів: кількість користувачів, за якими слідкує цей обліковий запис.

• Кількість списків: кількість загальнодоступних списків, членом яких є цей користувач.

1.8 Аналіз атрибутів профілю

У нашому аналізі атрибут geo-enabled, профіль за замовчуванням і зображення за замовчуванням є бінарними атрибутами і вивчаються як

двійкові змінні. Аналогічно, числові атрибути кількості улюблених, твітів, підписників/друзів, а також кількість списків аналізуються як є.

На основі оголошеного імені в обліковому записі Twitter ми створили двійковий і числовий атрибут: ми зіставили назву облікового запису з каталогом англійських імен, щоб перевірити, чи якась частина їх оголошеного імені справді є ім'ям людини в словнику. Ми також підрахували кількість частин у назві облікового запису, які доступні в словнику. Наприклад, ім'я облікового запису, у якому збігається лише ім'я, має значення 1, тоді як ім'я облікового запису, яке містить у словнику і ім'я, і прізвище, має значення 2. Для імені користувача облікового запису Twitter ми перевірили щоб перевірити, чи містить оголошене ім'я користувача. Для атрибутів опису, URL-адреси та розташування ми просто перевірили, чи надано користувачем відповідну інформацію.

Нарешті, ми використали інструмент лінгвістичного аналізу для вивчення описів профілю облікового запису, щоб зрозуміти, як користувачі різних налаштувань конфіденційності описують себе в Twitter.

У таблиці 1.3 представлено вибрані двійкові функції, а також відсоток захищених і загальнодоступних облікових записів, для яких ці бінарні атрибути є. Хоча результати тесту хі-квадрат свідчать про статистичну значущість для всіх функцій, значення розміру ефекту свідчать про те, що лише три функції мають практично різні значення в загальнодоступних і захищених облікових записах: має місце розташування, увімкнено гео та є профілем за замовчуванням. Ми розрахували величину ефекту за допомогою V Крамера і дотримувались умов для інтерпретації значення [2]. Щоб показати практично значущий ефект у реальності, V Cramer має бути не менше .1. Як показано в таблиці, більший відсоток захищених облікових записів увімкнув функцію геотегування та надав інформацію для атрибута розташування. Крім того, більш захищені облікові записи змінили налаштування профілю за замовчуванням порівняно з загальнодоступними обліковими записами.

У таблиці 1.3 наведено середнє значення наших числових характеристик у двох типах рахунків. Ми розрахували величину ефекту, використовуючи d Коена, і дотримувались конвенції для інтерпретації значення [2]. Зокрема, в контексті нашого дослідження значення d функції Коена має бути щонайменше 0,2, щоб розглядатися як практично корисна функція, яка відрізняє захищені облікові записи від загальнодоступних. Хоча результати t-тесту показують статистичну значущість для всіх функцій, значення розміру ефекту свідчать про те, що лише функція підрахунку твітів має практично інше значення в загальнодоступних і захищених облікових записах. Результати показують, що в середньому захищені облікові записи твітують частіше, а ефект цієї функції близький до середнього (d = 0,29) (див. таблицю 1.3). Здається, що захищений обліковий запис має більшу кількість улюблених твітів, хоча ефект все ще досить незначний (d = 0,09).

Загалом результати цікаві і суперечать тому, що ми очікували до аналізу. Наприклад, ми очікували, що, оскільки захищені облікові записи представляють собою більш приватну групу або групу, яка більше обізнана з конфіденційністю, вони з меншою ймовірністю включатимуть функцію відстеження місцезнаходження або змінюють тему профілю за замовчуванням чи навіть часто твітитимуть. Однак ці висновки свідчать про інше.

Таблиця 1.3 - Аналіз атрибутів бінарного профілю захищених і публічних облікових записів.

Binary Attributes	%Protected	%Public	Effect Size
Has Name	71.17	68.86	0.02
Username Has Name	3.01	3.45	0.01
Has Description	56.79	51.69	0.05
Has URL	15.01	16.78	0.02
Has Location	64.70	49.05	0.15
Is Geo-enabled	39.78	25.59	0.15
Is Default Profile	33.26	71.17	0.38
Is Default Image	6.43	8.80	0.04

Таблиця 1.4 - Аналіз числових атрибутів профілю захищених та публічних облікових записів.

Numeric Attributes	Protected	Public	Effect Size
Favirote Count	189.32	115.43	0.09
Tweet Count	1389.16	384.55	0.29
Follower Count	80.71	166.78	0.03
Friend Count	255.78	242.76	0.03
List Count	1.01	0.93	0.0006
Name Count	1.10	1.07	0.04

Для аналізу мовних категорій у цих описах ми використали мовний запит і кількість слів. Програма LIWC обробляє кожен текстовий файл слово за словом і порівнює їх із попередньо створеним словником, щоб визначити категорію LIWC, до якої належить це слово. Після обробки всіх слів у тексті LIWC обчислює та виводить відсоток кожної категорії LIWC. Перед проведенням лінгвістичного аналізу LIWC ми застосували наступні етапи попередньої обробки описів:

• видалені символи HTML

• замінено вилучення апострофа (наприклад, I’m -¿ I am).

• замінено URL-адреси словом “url”

• замінили смайли на відповідні значення (наприклад, :) -¿ усмішка)

• вилучено розділові знаки

• замінено обробники користувачів словом «згадати»

Словник LIWC структурований в ієрархічному форматі, де кожна категорія може охоплювати кілька підкатегорій. Детальну інформацію про ці категорії можна знайти в LIWC.

Оскільки описи профілів користувачів зазвичай дуже короткі (зазвичай від 8 до 10 слів), відсотки, надані LIWC, часто дуже малі для більшості категорій. Тому ми зосередилися лише на категоріях вищого рівня, які знаходяться у верхній частині ієрархії LIWC.

Таблиця 1.5 - Категорії LIWC та їх відповідний відсоток для охоронюваних та загальнодоступних описів

LIWC Category	Protected	Public	Effect Size
Function Words	37.51	33.68	0.15
Affect	8.68	7.50	0.10
Social Processes	11.08	10.75	0.02
Cognitive Processes	7.86	6.71	0.09
Drives and Needs	11.01	11.38	0.02
Relativity	10.03	10.23	0.0006

У таблиці 1.5 наведено ці категорії, а також відповідні відсотки для захищених і публічних рахунків. Тут ми виключили ті категорії LIWC, які мали менше 5% відповідних слів у всьому корпусі описів. Крім того, LIWC виводить набір підсумкових розмірів разом із відсотком відповідних слів. У таблиці 3.5 наведено підсумкові змінні, які вважаються релевантними, та їх відповідні відсотки для двох наборів рахунків. Для цих категорій виконується t-тест разом із величиною ефекту, виміряною d Коена. Усі категорії мають статистично значущі різні значення між захищеними та публічними рахунками, але ці відмінності невеликі на основі d Коена. Все ще цікаво відзначити, що захищений обліковий запис містить більший відсоток функціональних слів і слів, що впливають, що, оскільки подібне до наших висновків щодо поверхневих атрибутів, суперечить нашим попереднім очікуванням.

На додаток до основних категорій LIWC, аналіз підсумкових параметрів показує, що захищені облікові записи містять меншу кількість довгих слів (тобто слів із шістьма або більше літерами). Вони використовують менше слів, що представляють аналітичне мислення та вплив. Однак у них вищий відсоток слів, які мають емоційний тон та автентичність. Відмінності є статистично значущими на основі результатів t-критерію, але практично не значущими від значення d Коена.

Таблиця 1.6 - Змінні LIWC та відповідні їм значення для захищених та відкритих описів.

Summary Dimension	Protected	Public	Effect Size
Six Letter Words	22.73	26.69	0.14
Analytical Thinking	75.64	84.04	0.15
Clout	66.83	72.99	0.10
Emotional Tone	98.58	96.55	0.05
Authentic	28.89	21.24	0.13

1.9 Класифікація користувачів

Ми використали двійкові та числові функції, а також функції LIWC, які у кількох класифікаторах для ідентифікації захищених облікових записів. Класифікації проводяться за допомогою інструментарію машинного навчання Weka, а результати оцінюються за допомогою стратифікованої 10-кратної перехресної перевірки.

Для класифікації ми змінили деякі функції LIWC. Ми змінили категорії LIWC, які відповідали менш ніж 25% всього корпусу, на бінарні атрибути (див. таблицю 1.7 для відсотка категорій), такі як вплив, соціальні процеси, а також спонукання та потреби. Наприклад, якщо опис містить будь-яке слово, яке віднесено до категорії афектного слова в LIWC, відповідна функція встановлюється на 1; в іншому випадку він встановлюється на 0. Решта атрибутів LIWC зберігаються як є.

Ми також додали наявність чотирьох ключових слів як додаткових функцій. Протягом нашого аналізу ключових слів описів було виявлено, що два ключові слова follow та business зазвичай присутні в загальнодоступних описах порівняно із захищеними. Крім того, в публічних описах, схоже, згадуються інші облікові записи частіше, ніж їхні захищені аналоги. З іншого боку, слово посмішка часто використовується в захищених описах, які можуть бути або використанням слова безпосередньо, або смайликом, заміненим словом посмішка на етапі очищення.

Оскільки наша класифікація проводиться для користувачів, які мають опис, функція, яка перевіряє наявність опису, тут не має значення, тому її видаляють. Крім того, користувачі Twitter із захищеними обліковими записами повинні схвалити своїх підписників, а будь-який користувач може миттєво стежити за загальнодоступними обліковими записами. Таким чином, відмінності в кількості послідовників можуть не обов'язково випливати з відмінностей у ставленні до конфіденційності; отже, кількість підписників також видаляється з нашого набору функцій. Підсумовуючи, наші класифікатори побудовані на основі 26 об'єктів профілю, витягнутих з кожного облікового запису, які складаються з 11 поверхневих атрибутів, 11 обговорень.

Таблиця 1.7 - Оцінка результатів класифікації

Algorithm	Precision	Recall	F-score
Naive Bayes	0.66	0.67	0.66
Regression	0.71	0.70	0.71
Logistic	0.69	0.70	0.69
J48	0.68	0.66	0.67
KNN	0.67	0.59	0.63

У таблиці 1.7 наведено наші результати оцінки кількох класифікаторів. Наші найкращі результати класифікації отримані за допомогою ClassificationViaRegression з продуктивністю 0,71. Оскільки в літературі немає порівнянного дослідження, ми використовували випадковий класифікатор як базовий рівень. Виходячи з точної кількості захищених і загальнодоступних облікових записів у нашому базовому наборі, 48% набору складається із захищених облікових записів, тоді як 52% — це користувачі з загальнодоступними обліковими записами. Таким чином, випадковий класифікатор позначатиме захищені екземпляри F-показом 0,48. Наш набір функцій перевершує цей базовий рівень для всіх алгоритмів і покращує результати на 23% у кращому випадку.

Крім того, ми використали наївну базову лінію для порівняння результатів. Ця базова лінія вирішує позначати користувачів на основі їхньої функції з підтримкою географічного розташування. Це правило встановлюється на основі того, як влаштований інтерфейс сторінки конфігурації налаштувань Twitter. На цій сторінці розділ, призначений для зміни атрибута geo-enabled, розміщений у верхній частині розділу, призначеного для зміни налаштувань конфіденційності. Таким чином, ця базова лінія наївно припускає, що користувачі, які змінили своє поле за замовчуванням із геоданими з false на true, знають про налаштування конфіденційності за замовчуванням і, таким чином, змінили свої налаштування конфіденційності на захищений режим. Наївний базовий рівень досягає F-балла 50,77, що приблизно на 21% гірше, ніж наш найкращий алгоритм. Ці результати свідчать про те, що навіть покладаючись лише на атрибути профілю, можна автоматично визначити налаштування конфіденційності користувачів у соціальних мережах. Цей висновок є обнадійливим. Ми закликаємо до аналізу інших атрибутів користувачів Twitter та їх потенційного зв'язку з їхньою поведінкою щодо конфіденційності.

Характеризувати переваги конфіденційності користувачів у соціальних мережах є складним і непростим завданням, яке вимагає ретельного вивчення різних аспектів соціального впливу користувачів. Один клас таких слідів можна знайти в тому, як користувачі формують і створюють профілі своїх акаунтів. У цьому дослідженні ми виявили можливі зв'язки між атрибутами профілю користувачів у Twitter та їх налаштуваннями конфіденційності.

Зокрема, ми виявили, що захищені облікові записи частіше включають функцію гео-відстеження, ніж загальнодоступні. Крім того, вони, як правило, надають інформацію про своє місцезнаходження та змінюють тему профілю за замовчуванням. Крім того, вони твітують набагато частіше. Ці відмінності, поряд із загальною наявністю емоцій та афектних слів в описах захищених профілів, можуть бути пов'язані з екстравертною особистістю. Виходячи з цієї

інтерпретації, наші висновки суперечать попереднім дослідженням конфіденційності в традиційних умовах, де стверджується, що інтроверти, як правило, більше піклуються про конфіденційність і, швидше за все, відчувають вторгнення, коли їх просять розкрити особисту інформацію [9].

Можлива припущення: оскільки користувачі із захищеними обліковими записами знають, що їхні твіти та облікові записи є конфіденційними, вони почуваються в безпеці в цьому середовищі та готові добровільно розкрити більше інформації про себе та активніше брати участь у мережі. З іншого боку, користувачі, які свідомо слідкують за загальнодоступними налаштуваннями, використовують іншу стратегію для захисту своєї конфіденційності, яка не включає інформацію про своє місцезнаходження, використання теми за замовчуванням, менше твітів або використання меншої кількості функцій і слів у своїх описах. Якщо це так, то користувачі, які активніше беруть участь у соціальних мережах (наприклад, змінюють атрибути свого профілю, діляться особистою інформацією та розкривають емоційні стани), як правило, відчувають себе в безпеці в оточенні. Тому вони, швидше за все, відчують вторгнення цільової реклами та маркетингових повідомлень.

Як обговорювалося раніше, поведінка користувачів щодо конфіденційності може не обов’язково збігатися з їхнім ставленням до конфіденційності. Незважаючи на наші зусилля вибрати набір облікових записів з мінімальною кількістю таких помилкових спрацьовувань, наш базовий набір все ще може включати загальнодоступні облікові записи, які мали бути захищені. Незважаючи на цю проблему, ми отримали F-показник 0,71, що вказує на значення та важливість атрибутів профілю у виявленні поведінки конфіденційності. На основі цього висновку можна розробити методики без нагляду або напівнагляду, щоб ефективно ідентифікувати публічні облікові записи, які належать людям, які піклуються про конфіденційність, з урахуванням атрибутів профілю користувача.

РОЗДІЛ 2 ВИРІШЕННЯ ПРОБЛЕМИ КОНФІДЕНЦІЙНОСТІ В TWITTER

2.1. Основні поняття персоналізації та конфіденційності даних

Поняття персоналізації та конфіденційності даних завжди переплітаються, оскільки ефективна персоналізація значною мірою залежить від збору та обробки особистих і потенційно конфіденційних даних користувачів. Напруження між персоналізацією та конфіденційністю часто називають парадоксом персоналізації та конфіденційності [5]. Якщо питання конфіденційності будуть належним чином вирішені, персоналізація може зблизити клієнтів і бренди і значно збільшити прибутки фірми. Однак персоналізація за рахунок порушення конфіденційності може легко відштовхнути клієнтів і навіть спонукати їх виступати проти фірми. Тому пошук солодкого місця між «індивідуалізованими» та «інвазивними» комунікаціями є центральним для будь-яких успішних персоналізованих і індивідуальних маркетингових спроб [7].

Узгодження парадоксу персоналізації та конфіденційності може бути особливо складним у контексті веб-сайтів соціальних мереж. На таких платформах користувачі можуть захистити свої дані за допомогою засобів контролю конфіденційності; однак дослідження поведінки в Інтернеті показують, що існує невідповідність між ставленням користувачів соціальних мереж до конфіденційності та їх фактичною поведінкою при визначенні політики конфіденційності. Ці рішення щодо конфіденційності часто є складними і вимагають ретельного аналізу компромісів між потенційною соціальною вигодою та можливими ризиками для конфіденційності. Таким чином, багато користувачів уникають проблем із налаштуванням конфіденційності та дотримуються налаштувань за замовчуванням, які часто неправильно сприймаються як рекомендований курс Версія цієї глави була подана для публікації в ACM Transactions on Social Computing 45 дії [4]. Тим

не менш, користувачі зазвичай не знають, що налаштування конфіденційності за замовчуванням зазвичай відкриті та дозволені (наприклад, Twitter та Instagram). Навіть серед тих, хто докладає зусиль, щоб керувати своєю конфіденційністю, багато хто досі не усвідомлюють наслідків своїх рішень [2].

Тим часом люди використовують соціальні платформи для спілкування з друзями та родиною, для обміну особистими даними, переконаннями, думками, поточною діяльністю і навіть даними про своє місцезнаходження. Такі масивні джерела даних, що накопичуються щодня, можуть дати уявлення про різні аспекти поведінки клієнтів і надати безпрецедентні маркетингові можливості для бізнесу. Наприклад, компанії можуть використовувати релевантну інформацію про кожного клієнта та надавати цінні пропозиції щодо наступної найкращої пропозиції, маркетингові повідомлення, пристосовані до кожної окремої людини, та правильні відповіді на служби підтримки клієнтів. Однак через проблему дихотомії конфіденційності в соціальних мережах дані користувачів можуть бути ненавмисно видимими для громадськості, і тому їх не слід використовувати для цілей профілювання користувачів. Таким чином, ефективне використання соціальних даних вимагає заходів із захисту конфіденційності, які виходять за межі поточних налаштувань конфіденційності користувачів.

Рішення, запропоновані для вирішення парадоксу персоналізації та конфіденційності в соціальних медіа, включають дослідження, які представляють набір принципів підвищення конфіденційності для проектування систем персоналізації, роботи, які пропонують використані інтерфейси та візуалізації для специфікації конфіденційності, а також моделі для прогнозування та визначення переваг конфіденційності на основі діяльності користувачів. Останній напрям мав скромний успіх у прогнозуванні налаштувань конфіденційності користувачів, однак його дослідження поки що було дуже обмеженим. Тому саме цей підхід є в центрі уваги цього дослідження. Зокрема, ми прагнемо створити набір функцій для

характеристики онлайн-поведінки загальнодоступних облікових записів користувачів, які мають бути приватними. Такий набір функцій можна використовувати в моделі прогнозування, щоб підняти «червоний прапор», попереджаючи підприємства про те, що користувач, який у фокусі уваги, ймовірно, не бажає бути профільованим в Інтернеті.

Наше дослідження проведено в Twitter, де контроль конфіденційності відповідає простій бінарній специфікації. У Twitter користувачі можуть дотримуватися загальнодоступних налаштувань за замовчуванням, які вказують, що їхні твіти та списки контактів є загальнодоступними та доступними через API Twitter. Крім того, вони можуть змінити налаштування конфіденційності на захищені, завдяки чому їхні твіти та контакти будуть доступними лише для їх затверджених підписників. За винятком атрибутів профілю (наприклад, імені та імені користувача), дані, пов'язані із захищеними обліковими записами, також недоступні через Twitter API, що унеможливлює пряме порівняння та аналіз захищених та загальнодоступних профілів. Натомість ми в першу чергу зосередилися на параметрах конфіденційності як локалізованому атрибуті, щоб отримати уявлення про атрибути користувачів із різними параметрами конфіденційності. Наш аналіз, експерименти та моделі засновані на соціальній мережі 23 тисяч користувачів Twitter, зібраних із випадкових користувачів — це соціальне явище, яке ґрунтується на тому, що «подібність породжує зв'язок» [3]. Попередні дослідження в соціальних мережах свідчать про те, що контекст сусідства містить значну інформацію про користувачів і може бути використаний для визначення уподобань користувачів [3]. Крім того, з часом на переваги соціальних мереж можуть впливати їхні соціальні контакти в соціальній мережі. Наш первинний аналіз сусідніх районів показує, що налаштування конфіденційності дійсно локалізовані в Twitter [4]. Крім того, ми опублікували твіти з підмножини зібраних облікових записів на краудсорсинговій платформі та попросили людей-анотаторів оцінити переваги конфіденційності

авторів твітів. Ми обрали Amazon Mechanical Turk (AMT) як платформу фокусу. Результати експерименту AMT також показують, що параметри конфіденційності локалізовані в Twitter.

Тому ми використовуємо той факт, що параметри конфіденційності локалізовані в мережі, і характеризуємо їхні переваги конфіденційності за ступенем занепокоєння конфіденційністю, яке висловлюють його/її соціальні контакти. У цій роботі поточне налаштування конфіденційності соціальних контактів (тобто публічних і захищених) розглядається як ознака їхньої переваги щодо конфіденційності. Зокрема, ми використали показник конфіденційності, який кількісно визначає відсоток захищених соціальних контактів до всіх контактів користувача. Таким чином, користувачі з вищим показником конфіденційності – це ті, хто перебуває в більш приватних районах порівняно з користувачами з нижчим показником конфіденційності. Виходячи з нашого попереднього аналізу, ми припускаємо, що загальнодоступні облікові записи, розташовані в приватних районах, швидше за все, належать користувачам, які піклуються про конфіденційність, і страждають від проблеми дихотомії конфіденційності. Далі ми використовуємо це відкриття, щоб охарактеризувати поведінку таких користувачів.

Щоб знайти набір атрибутів, пов’язаних із налаштуваннями конфіденційності, ми досліджуємо, чи відрізняється поведінка публічних користувачів у Twitter, якщо вони розташовані в різних районах (з різним ступенем занепокоєння конфіденційністю). Щоб дослідити ці відмінності, ми досліджуємо різні показники поведінки, включаючи атрибути профілю, які легко доступні з облікових записів Twitter (наприклад, кількість твітів і описи профілів), лінгвістичні аспекти твітів (наприклад, використання функціональних слів, ступінь автентичності та настрої пов’язані характеристики), а також атрибути, пов’язані з їх поведінкою в спілкуванні (наприклад, кількість відповідей і ретвітів). Згідно з результатами, публічні

акаунти, розташовані в приватних кварталах, мабуть, поділяють контент, який є більш приватним відповідно до суспільного консенсусу [4]. Цей висновок вказує на те, що такі користувачі можуть бути стурбовані конфіденційністю, хоча у них може складатися неправильне враження, що їхні дані захищені в мережі.

Щоб виявити приховані функції, які представляють інтерес для користувачів, які піклуються про конфіденційність, ми перетворили кожного користувача в мережі Twitter на набір атрибутів, що представляють користувача (наприклад, хештеги, уніграми та теми, які використовуються в їхніх твітах). Для кожного вузла атрибутів ми обчислюємо показник конфіденційності на основі кількості захищених і загальнодоступних облікових записів, що оточують цю функцію. Подібно до нашого попереднього аналізу твітів, атрибути, які є більш приватними, часто зустрічаються в приватних районах. Наприклад, хештеги #Annoyed і #SoTired частіше використовуються публічними акаунтами з великим відсотком захищених контактів. Оскільки наша робота заснована на локальності переваг і використовує відомі переваги користувачів для прогнозування невідомих переваг інших користувачів, її можна вважати методом спільної фільтрації [40].

Попередні дослідження щодо автоматичного визначення переваг соціальної конфіденційності використовували різноманітні типи даних на веб-сайтах соціальних мереж. Ряд досліджень спирався на потенційні зв’язки між атрибутами профілю користувачів та їхніми вподобаннями щодо конфіденційності. Деякі дослідники запропонували алгоритми, які в першу чергу орієнтовані на соціальний контекст та зв’язки користувачів. Деякі дослідження використовували вміст (наприклад, текст і зображення), опублікований користувачами, щоб отримати бажані функції конфіденційності.

Атрибути профілю. У більшості веб-сайтів соціальних мереж кожен обліковий запис користувача пов’язаний з набором атрибутів профілю. Підмножина таких функцій профілю налаштовується власником облікового запису, часто під час створення облікового запису, і призначена для представлення того, хто є користувачем у мережі. Приклади таких атрибутів профілю включають ім’я користувача, зображення профілю та інформацію про місцезнаходження користувача. Інша підмножина атрибутів, які також вказуються користувачем, пов'язана з налаштуваннями його облікового запису. Наприклад, вони можуть вказати бажану мову інтерфейсу або хочуть, щоб їхня спільна інформація була позначена географічними тегами. Нарешті, набір контекстних атрибутів розраховується самим веб-сайтом соціальної мережі на основі діяльності користувачів. Наприклад, час створення облікового запису, кількість елементів, опублікованих кожним користувачем, і кількість соціальних контактів. Обмежена кількість досліджень було зосереджено на потенційних відносинах атрибутів профілю та поведінки та ставлення до конфіденційності.

У контексті Facebook алгоритм контрольованого навчання побудований на великому наборі функцій, щоб рекомендувати налаштування конфіденційності [5]. Ці функції включають елементи метаданих щодо спільного елемента, а також демографічні характеристики та особливості профілю користувачів. Варто відзначити що алгоритм розроблено для того, щоб рекомендувати налаштування конфіденційності для певного спільного елемента, наприклад окремих дописів у Facebook, на відміну від прогнозування загальних налаштувань конфіденційності. Зосередившись на Twitter, Донг та ін. [6] запропонував модель прогнозування конфіденційності, яка враховує поведінкові аналоги соціальних медіа до психологічних змінних, які, як відомо, впливають на поведінку користувачів у розкритті інформації. Деякі з виявлених аналогів засновані на особливостях профілю користувачів.

Наприклад, надійність користувачів розраховується на основі співвідношення їхніх підписників до загальної кількості їхніх соціальних контактів.

Соціальний контекст. Щодо кількості досліджень, зосереджених на атрибутах профілю, великий набір підходів зосереджено на соціальному контексті основного користувача, щоб передбачити особливості конфіденційності. Ці дослідження можна розділити на дві основні групи. Перший набір робіт переважно зосереджується на конфіденційності з точки зору видимості інформації для різних груп соціальних контактів, які часто називають соціальними колами. Таким чином, вони пропонують підходи, щоб допомогти користувачам створити та підтримувати такі соціальні кола та визначити бажані налаштування конфіденційності для створених кіл контактів. Друга група робіт зосереджена насамперед на та запропонованих методах спільної фільтрації для визначення бажаних функцій конфіденційності.

Запропоновано різні алгоритми створення кіл соціальних контактів, які тісно пов’язані між собою і мають багато зв’язків всередині себе, при цьому мають менше зв’язків з тими, хто не входить до кола [5, 2]. У [8] після формування кластерів контактів користувачеві пропонується позначити кількість випадково вибраних контактів з кожного кластера з точки зору його/її бажання поділитися з ними певним елементом. Класифікатор навчається на атрибутах профілю, а також на мережевих атрибутах позначених контактів, щоб передбачити налаштування конфіденційності користувача для немаркованих контактів. У доволі актуальній роботі Фанг і ЛеФевр [7] створили майстер конфіденційності, який ітераційно просить користувача позначати ретельно відібрані інформаційні контакти з точки зору його/її бажання поділитися з ними певною інформацією профілю. Щоб автоматично позначати інші контакти, ця позначена інформація використовується в класифікаторі, у якому контакти представлені з точки зору їхньої структури спільноти та атрибутів профілю.

Напрям досліджень, тісно пов’язаний з цим дослідженням, адаптував методи з області спільної фільтрації для прогнозування власних уподобань щодо конфіденційності. Ці дослідження мотивовані гомофілією, яка відноситься до схильності людей спілкуватися з подібними людьми [3]. Крім того, виявилося, що на поведінку користувачів, що обмінюються інформацією, значно впливає внутрішнє коло близьких друзів [4]. Тому один набір підходів зосереджений на прогнозуванні переваг конфіденційності на основі уподобань соціальних контактів основного користувача (наприклад, друзів у Facebook або підписників у Twitter). Подібність: люди зі схожим походженням, як правило, мають подібні проблеми щодо конфіденційності [3]. В якості альтернативи використанню соціальних контактів ряд дослідників розробив методи визначення переваг конфіденційності на основі відомих уподобань користувачів зі схожим походженням і характеристиками з основним користувачем.

Наприклад, Squicciarini et al. надають алгоритм формування соціальних кіл на основі характеристик користувачів, таких як стать, хобі та рід занять. Коли новий об’єкт (тобто контакти або елементи даних) завантажується, система спочатку шукає соціальні кола, які, швидше за все, мають справу з об’єктом так само, як користувач. Тоді політика конфіденційності, яку використовує вибране коло, є основою для прогнозування політики конфіденційності для щойно доданого об’єкта. У [9] фокусний користувач спочатку вказує, чи бажає він/вона поділитися певним елементом даних із вибраним контактом. Потім застосовується ітераційний напівконтрольований підхід для позначення інших контактів користувача, де мітки поширюються від позначених екземплярів до немічених екземплярів у соціальному графі. Це поширення керується показником подібності користувачів, який базується на інформації про профіль контактів, показниках їхніх мереж, а також на структурі спільноти.

У [1] користувачі спочатку характеризуються відповідно до їхніх уподобань щодо конфіденційності як фундаменталістів конфіденційності, прагматиків конфіденційності або байдужих до конфіденційності. Рішення щодо конфіденційності та налаштування користувачів щодо їхніх фотоальбомів вважаються ознакою їхньої переваги щодо конфіденційності. Зокрема, користувачів розподіляють до цих трьох категорій на основі кількості їхніх загальнодоступних, налаштованих та приватних фотоальбомів. Потім використовується алгоритм K-найближчого сусіда, щоб визначити, до якої категоризації конфіденційності належить основний користувач. На основі функцій призначеної категорії система рекомендує налаштування конфіденційності. У контексті соціальної мережі, заснованої на місцезнаходження, Toch et al. [6] надають користувачам рекомендовану політику конфіденційності, яку раніше обрали подібні користувачі.

Опублікований вміст. Частою діяльністю користувачів у соціальних мережах є публікація та обмін вмістом, таким як текстові повідомлення, зображення та відео. Ці спільні типи даних можна використовувати, щоб зробити висновки про особистість та уподобання користувачів. У контексті текстового контексту та конфіденційності Gill et al. [1] надають набір категорій слів, пов'язаних з конфіденційністю, які є релевантними для семантичного аналізу домену конфіденційності. LIWC також містить велику кількість семантичних категорій слів, які можуть мати відношення до функцій конфіденційності. Каліскан-Іслам та ін. [4] використовував словник конфіденційності разом із різноманітними методами та інструментами, включаючи моделювання теми, розпізнавання іменованих об’єктів та аналіз настроїв, щоб автоматично визначити, чи містить твіт особисту інформацію. Потім користувачам надаються оцінки конфіденційності на основі кількості приватної інформації, яку вони опублікували у своїх графіках у Twitter. Часові шкали позначених користувачів потім використовуються в контрольованій техніці для призначення оцінок конфіденційності користувачам без міток.

Модель прогнозування, запропонована в [5], також дотримується контрольованого підходу до рекомен- змінити налаштування конфіденційності для певної публікації у Facebook. Оцінка настроїв публікації включена до їхнього набору функцій разом із деякими актуальними функціями, які вказують, чи пов'язана публікація з набором попередньо визначених тем (наприклад, сім'я, робота, подорожі тощо) чи ні. Крім того, враховується повне представлення вмісту. Враховуючи неструктурований лінгвістичний контент, опублікований користувачем, [3] спочатку виявляє конфіденційну інформацію, таку як номер телефону, адреса та місцезнаходження, у тексті. Тоді модель, запропонована в [9], приймається для кількісної оцінки потенційного ризику конфіденційності користувача в мережі.

У контексті обміну зображеннями [3] використовує попередні зображення користувачів і відповідні їм політики конфіденційності, щоб призначити політику конфіденційності новому зображенню. Для створення конфіденційності використовуються кластеризація зображень і аналіз асоціацій політик. Однак, якщо користувач новий або відбулися значні зміни в тенденціях щодо конфіденційності користувача, для прогнозування політики використовується соціальний контекст користувачів. Щоб виявити зображення з приватним вмістом, Zerr et al. [4] використовували різноманітні візуальні особливості, такі як наявність облич, під наглядом. Вони також використали текстові метадані зображень і виявили кореляції між темою та вмістом приватного зображення. Наприклад, було показано, що тема, яка використовується для опису особистих концепцій, емоцій та настроїв, а також людське тіло в основному використовується для особистих зображень. З іншого боку, теми, пов'язані з природою, архітектурою та неживими об'єктами, здебільшого зустрічаються в неприватних зображеннях.

Згідно з нашим попереднім аналізом літератури [3], більшість досліджень зосереджено на прогнозуванні налаштувань конфіденційності,

характерних для конкретного веб-сайту соціальної мережі, тоді як загальним моделюванням та характеристикам користувачів приділено менше уваги. Незважаючи на те, що нашою основною платформою є Twitter, ми прагнемо охарактеризувати загальні переваги конфіденційності загальнодоступних профілів користувачів. Крім того, на відміну від багатьох попередніх спроб, які зосереджені на одному або кількох типах даних, ми включаємо різноманітні типи даних (наприклад, особливості профілю, соціальний контекст та опублікований вміст). Нарешті, багато з розглянутих досліджень використовували контрольовані методи для класифікації та прогнозування атрибутів конфіденційності. Однак такі методи вимагають маркованого введення і можуть здатися неможливими в контексті соціальних мереж, де позначена інформація зазвичай становить дуже малу частину доступних даних. Замість цього ми пропонуємо неконтрольований підхід до спільної фільтрації, який мотивується місцевістю уподобань у соціальних мережах.

2.2 Структура наборів даних

Щоб зібрати та створити соціальну мережу з Twitter, ми спочатку вибрали випадкового користувача, згенерувавши випадковий ідентифікатор Twitter. Ми забезпечили, щоб цей початковий користувач був загальнодоступним, оскільки соціальні контакти захищених облікових записів недоступні через API Twitter, що унеможливлює розширення мережі із захищеного вузла користувача. Після вибору цього користувача ми ітераційно створювали мережу користувачів методом пошуку в ширину (BFS). З огляду на те, що наш підхід використовує місцевість переваг, ми зосередилися лише на взаємних стосунках замість асиметричних стосунків із наступними або дружніми стосунками. Очікується, що взаємні відносини вказують на міцніші стосунки між двома користувачами, і вони відрізняють розділ соціальної мережі Twitter-сфери від її інформаційної мережі [4]. Оскільки ми зосереджені

лише на цих взаємних контактах, відтепер, коли ми використовуємо слово контакт, ми маємо на увазі соціальні контакти фокусного користувача із взаємними стосунками.

Перш ніж додавати кожного загальнодоступного користувача до мережі, ми отримуємо та обчислюємо набір метаданих про користувача. Спочатку підраховуємо відсоток захищених контактів до всіх контактів основного користувача. Наприклад, якщо користувач має 100 соціальних контактів, з яких 20 захистили свої акаунти, користувачу буде присвоєно значення 20%. Цей відсоток, який називається коефіцієнтом конфіденційності, є основним показником для нашого подальшого аналізу. Крім того, ми збираємо атрибути профілю Twitter (наприклад, місцезнаходження та кількість твітів) і останні 500 твітів, опублікованих користувачем як метадані вузла. Після того, як розширений користувальницький вузол додається до мережі, ми перевіряємо, чи має новий вузол взаємні відносини з будь-яким із існуючих вузлів, і додаємо відповідні ребра. Цей процес повторюється з новим відкритим користувачем, вилученим із черги BFS. На рисунку 2.1 показано огляд нашого процесу збору даних.

Слід зазначити, що користувачі, які мають менше 10 твітів або менше 30 підписників/друзів, вважаються неактивними і, таким чином, не включаються в процес збору даних.

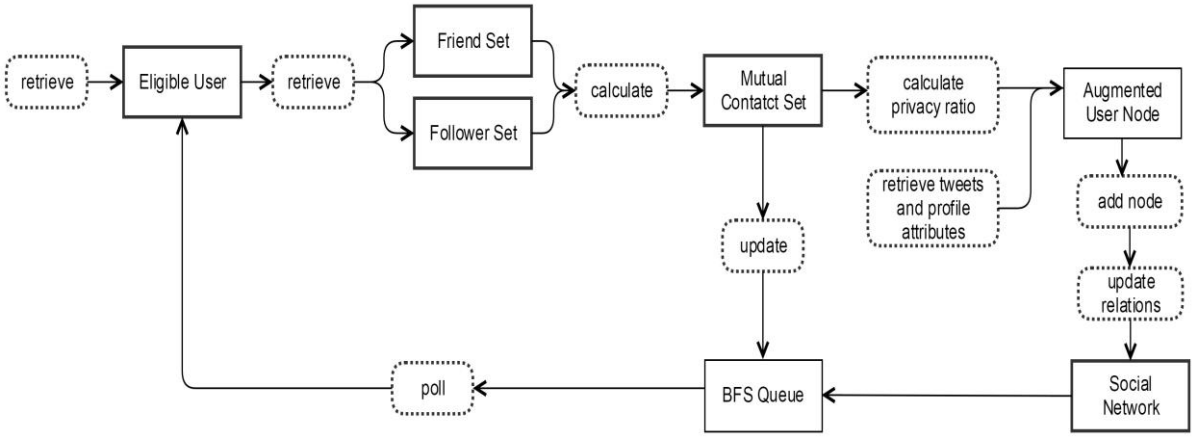
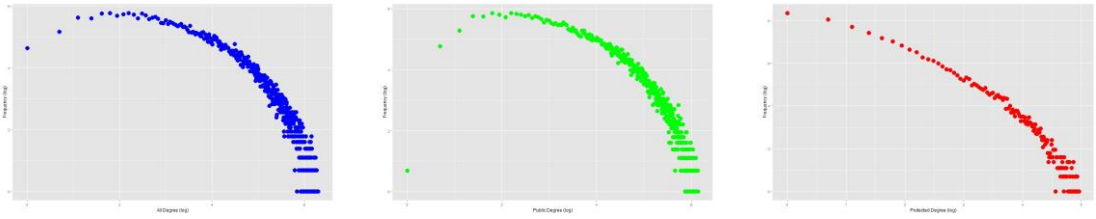


Рисунок 2.1 - Огляд процедури пошуку даних



(a) (b) (c)

Рисунок 2.2 - Розподіл ступенів між усіма користувачами мережі в логарифмічній шкалі. Розподіл показано для всіх соціальних контактів (а), публічних контактів (б) також захищених (в).

Перевірені користувачі та користувачі з понад 1000 підписників/друзів виключені, оскільки вони часто представляють бренди та знаменитостей і не належать до широкої публіки. Дотримуючись цього підходу, ми зібрали в цілому 23 320 загальнодоступних вузлів користувачів і 6 489 419 твітів, опублікованих цими користувачами.

У цьому наборі даних кожен обліковий запис Twitter взаємно пов'язаний із в середньому 86 контактами. Серед цих сусідів у середньому 76 є державними та 10 захищеними. Крім того, кожен користувач пов'язаний з в середньому 339 твітами. Ми можемо отримати деяке уявлення про структуру мережі, вивчивши розподіли ступенів. На рисунку показано розподіл ступенів для всіх взаємних контактів для всіх користувачів у логарифмічній шкалі.

Ми також спробували пристосувати всі три розподіли ступенів до розподілу степеневого закону:

$P(x) \sim x^{-\alpha}$. Під час підбору ми отримали значення α 2,49, 2,98 і 1,79 для всіх публічних і захищених облікових записів відповідно. Для всіх трьох розподілів тест Колмогорова-Смірнова (KS) показує, що розподіл не відмовляється ($P > 0,05$), і степеневий закон дійсно може підходити. Розподіл степеневого закону зазвичай спостерігається в контексті соціальних мереж, хоча цікаво спостерігати ту саму тенденцію навіть після фільтрації користувачів з більш ніж 1000 друзів/послідовників.

На рисунку 2.3 показано взаємозв’язок загальнодоступних і захищених контактів для всіх користувачів. Не дивно, що ці дві метрики позитивно корелюють, що вказує на те, що зі збільшенням кількості публічних контактів зростає кількість захищених контактів. Однак, оскільки лінія лінійної регресії представляє, кількість публічних контактів зростає у більшому масштабі порівняно з

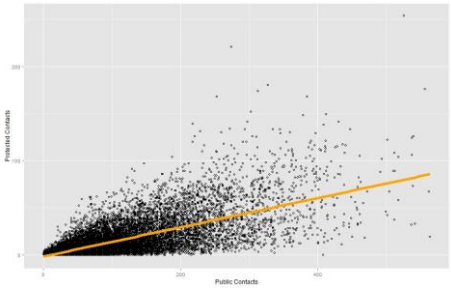


Рисунок 2.3 - Співвідношення кількості публічних контактів і кількості захищених контактів

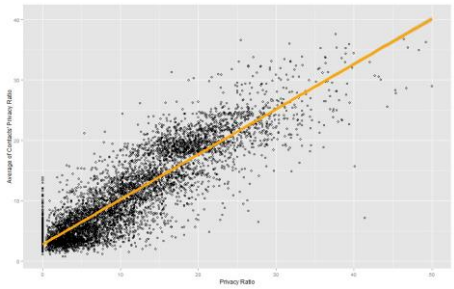


Рисунок 2.4 - Співвідношення коефіцієнта конфіденційності користувачів і середнього коефіцієнта конфіденційності для всіх їхніх контактів

Нарешті, як перший крок для забезпечення локалізації налаштувань конфіденційності в контексті конфіденційності, ми розрахували кореляцію між коефіцієнтом конфіденційності кожного вузла та середнім коефіцієнтом конфіденційності контактів. Аналіз користувачів, які мають щонайменше 10 взаємних контактів у мережі (близько 7000 користувачів), показав сильну позитивну кореляцію між двома змінними (Спірмен $\rho = +0{,}89$). Ця кореляція також очевидна на рисунку 2.4, де помаранчева лінія є лінією лінійної регресії, підігнаною до даних.

Незважаючи на те, що кореляційний аналіз сигналізує про існування переваги місцевості для атрибута конфіденційності, ми провели інший експеримент, заснований на анотаціях людей, щоб додатково підтвердити цей висновок. Ми опублікували 12 тисяч твітів для AMT і попросили працівників AMT оцінити переваги конфіденційності автора твіту на основі цього твіту. Ці 12 тисяч твітів були.

Експеримент Amazon Mechanical Turk обраний відповідно до наступної процедури.

Наш зібраний набір користувачів спочатку було відфільтровано лише для тих користувачів, які твітують англійською. Потім із набору вилучаються 1200 користувачів із найнижчим коефіцієнтом конфіденційності та 1200 користувачів із найвищим коефіцієнтом конфіденційності. Далі ми вибрали п’ять останніх твітів, опублікованих цими користувачами, які не є ретвітами чи відповідями. Ми також переконалися, що ці твіти містять певний текст і не генеруються автоматично (наприклад, не позначаються як [autotweet]). Цей процес призвів до відбору 12 тисяч твітів, опублікованих 2400 різними користувачами з надзвичайно низьким або надзвичайно високим коефіцієнтом конфіденційності. Кожен твіт потім публікується в AMT, щоб бути позначеним анотатором відповідно до його/її рішення щодо конфіденційності автора твіту. Для цього, отримавши твіт, працівники могли вибрати один із трьох запропонованих варіантів:

• Конфіденційність без уваги: його/вона не турбує, якщо його/її дані використовуються компаніями

• Нейтральні/об’єктивні: нейтральні або об’єктивні твіти, які не містять інформації про налаштування конфіденційності користувачів

• Заклопотаність конфіденційності: він/вона не бажає ділитися своїми даними з компаніями

Кожен твіт позначений трьома різними працівниками. В ідеалі мажоритарна етикетка повинна бути предметом подальшого аналізу. Однак результати вкрай незбалансовані. Лише близько 1% твітів позначено як твіти, опубліковані користувачем, який піклується про конфіденційність (тобто, третій варіант), тоді як кожна з інших категорій вибирається більше ніж у 50% випадків. Ця проблема може бути пов'язана з відсутністю досвіду працівників АМТ з даної теми, особливостями вибраних твітів або дизайном завдання анотації. Подальші поглиблені експерименти можуть пролити світло на це питання. Однак вирішення цього питання виходить за рамки цього дослідження і буде включено до нашої майбутньої дослідницької діяльності. Тут, щоб полегшити порівняння та аналіз, ми використали альтернативний підхід.

Твіт вважається віднесеним до категорії твітів, опублікованих користувачем, що піклується про конфіденційність, якщо він анотований як такий принаймні одним працівником. Потім ми підрахували кількість твітів, які класифікуються під цією міткою (тобто, що стосується конфіденційності) для кожного користувача в наборі даних. Порівняння кількості твітів такого роду серед двох груп користувачів показує статистично значущу різницю (p-значення t-тесту < 0,05). Зокрема, твіти, опубліковані користувачами з високим коефіцієнтом конфіденційності, частіше асоціюються з цією категорією, ніж твіти з низьким коефіцієнтом конфіденційності. Хоча ми визнаємо обмеженість аналізу щодо використання анотації одного працівника, цей висновок означає, що користувачі, які знаходяться в приватних районах, вважаються такими, що піклуються про конфіденційність, що є ще одним доказом місцевості уподобань у контексті конфіденційності.

2.3 Аналіз поведінки користувача

Ми проаналізували різні характеристики загальнодоступних користувачів і вивчили їхні потенційні зв'язки з показником конфіденційності.

Ці характеристики випливають з атрибутів профілю користувачів і їх опублікованих твітів, обидва з яких доступні безпосередньо з загальнодоступних профілів користувачів. Крім того, ми перетворили мережу користувачів на двосторонню мережу функцій і користувачів, де функціям надано співвідношення конфіденційності.

Підмножина доступних атрибутів профілю вважається релевантною для нашої мети та вибирається для подальшого аналізу. Цей список наведено нижче разом із коротким описом атрибутів. Зверніть увагу, що описи адаптовані зі специфікації API Twitter.

• Опис: фрагмент тексту, який користувачі надають для опису свого облікового запису.

• URL: URL-адреса, надана користувачем у зв'язку з його профілем.

• Місцезнаходження: визначене користувачем місцезнаходження профілю цього облікового запису.

• Geo-enabled: якщо значення true, це означає, що користувач увімкнув можливість позначення геотегів своїх твітів.

• Зображення за замовчуванням: якщо значення true, це означає, що користувач не завантажив своє зображення, а замість нього використовується аватар за замовчуванням.

• Профіль за замовчуванням: якщо значення true, це означає, що користувач не змінював тему або фон свого профілю користувача.

• Кількість обраних: кількість твітів, доданих цим користувачем у вибране за весь час існування облікового запису

• Кількість твітів: кількість твітів, виданих користувачем.

• Кількість підписників: кількість підписників на даний обліковий запис.

• Кількість друзів: кількість користувачів, за якими слідкує цей обліковий запис.

Ми використовували спрощений двійковий атрибут для вивчення описів користувачів, URL-адрес і місць розташування. Значення цього

атрибута дорівнює «1», якщо відповідна частина інформації надається користувачем, і «0» в іншому випадку. Атрибут із підтримкою геоданих, профіль за замовчуванням і зображення за замовчуванням – це всі двійкові атрибути і вивчаються як двійкові змінні. Аналогічно, числові атрибути кількості улюблених, кількості твітів, кількості підписників/друзів аналізуються як є. Щоб вивчити, як налаштовуються атрибути профілю користувачів із різними коефіцієнтами конфіденційності, ми розрахували співвідношення між атрибутами профілю та співвідношенням конфіденційності для всіх користувачів.

Атрибути профілю Twitter є одними з кількох точок даних, які є видимими та доступними для громадськості та API Twitter незалежно від налаштувань конфіденційності користувача. Тому ми отримали атрибути профілю захищеного та загальнодоступного облікових записів і безпосередньо порівняли їх. Такий додатковий експеримент дозволяє нам зрозуміти, чи по-різному налаштовані атрибути профілю користувачів Twitter із різними налаштуваннями конфіденційності. Крім того, порівнюючи ці висновки з аналізом загальнодоступних користувачів із різними співвідношеннями конфіденційності, ми можемо дослідити, чи захищені облікові записи поводяться так само, як публічні облікові записи, розташовані в більш приватних районах (щодо їх конфігурації атрибутів профілю). Пошук подібності між двома наборами може сприяти підтвердженню того, що атрибути конфіденційності локалізовані.

Щоб достовірно порівняти атрибути профілю користувачів із захищеними та загальнодоступними налаштуваннями, нам потрібен набір користувачів, у яких дихотомія ставлення до конфіденційності зведена до мінімуму. Наш попередній аналіз різноманітних наборів підписників, пов’язаних із відомими акаунтами у Twitter [4], показує, що послідовники «Крайні новини CNN» можуть бути хорошим набором кандидатів. Такий висновок зроблений тому, що відсоток захищених підписників до загальної

кількості підписників на «Пошукні новини CNN» виявився значно вищим, ніж інші набори підписників і середній відсоток у Twitter [4]. Тому ми зібрали атрибути профілів 1 мільйона загальнодоступних і 1 мільйон захищених облікових записів із набору підписників CNN та порівняли їх особливості профілю. У процесі збору ми переконалися, що неактивні облікові записи, бренди та знаменитості були виключені. Деталі процесу збору даних можна знайти в [4].

У таблиці 2.1 представлено аналіз вибраних бінарних ознак. Співвідношення між цими бінарними атрибутами та відношенням конфіденційності користувачів здається невеликим для майже всіх функцій. Однак відмінності між захищеними та публічними рахунками статично значущі для всіх змінних і навіть практично значущі для трьох з них. Цікаво, що за одним винятком (Має опис) результати показують подібну поведінку для користувачів із високим рівнем конфіденційності та тих, хто вибрав захищені облікові записи.

Таблиця 2.1 - Аналіз атрибутів бінарного профілю та функцій, пов’язаних із конфіденційністю.

Binary Attributes	Privacy Ratio (Spearman ρ)	Privacy Setting (Chi-Square P-value)
Has Description	-0.07	+++
Has URL	-0.17	- - -
*Has Location	-0.06	+++
*Is Geo-enabled	+0.11	+++
*Is Default Profile	-0.07	- - -
Is Default Image	-0.03	- - -

Для P < .05 і++ для P < .005, коли значення більше для захищених облікових записів, тоді - і - - - як - коли значення менше для захищених облікових записів. Змінні, для яких спостерігається хоча б невеликий розмір ефекту (V Крамера > 0,1), відомі зірочкою. коефіцієнти конфіденційності і

найчастіше користуються користувачами із захищеними обліковими записами (як статистично, так і практично значущими). Крім того, захищені облікові записи надають зовнішні URL-адреси рідше, ніж загальнодоступні. Аналогічно, негативна кореля має URL-адресу та коефіцієнт конфіденційності, які вказують на те, що користувач використовує більше приватних сусідів, як правило, не повідомляє інформацію про URL-адреси, пов’язані зі своїми аналогами.

У таблиці 2.2 наведено результати експерименту для числових ознак. Позитивна кореляція спостерігається між усіма змінними; однак, коефіцієнт кореляції вважає вищий для людей твітів і фаворитів. Ця позитивна кореляція показує на те, що користувачі з високим коефіцієнтом конфіденційності, як правило, частіше пишуть твіти й уподобають твіти. Результати прямого аналізу захищених та загальнодоступних облікових записів відповідають коефіцієнтам кореляції, оскільки захищені користувачі мають значно більшу кількість твітів, улюблених, друзів і підписників. Різниця для функцій підрахунку твітів також практично важливо.

Наш прямий аналіз захищених та загальнодоступних облікових записів показує, що захищені облікові записи добровільно розкривають більше інформації про себе (наприклад, геотеги) і ак активно беруть участь у мережі (наприклад, кількість твів). Можливі припущення: користувачі із позах обліковими записами, що їхні є конфіденційними, вони відчувають себе в безпеці в цьому середовищі. З іншого боку, користувачі, які свідомо слідкують за публічною обстановкою, використовують іншу стратегію, наприклад, самоцензуру, щоб захистити свою конфіденційність. Серед атрибутів наявних зовнішніх URL-адрес показує інший шаблон. Загальнодоступні облікові записи та користувачі в громадських районах, схоже, надають URL-адреси частіше. Не дивлячись на наші спроби виключити бренди та знаменитості, цю різницю можна віднести до професійних акаунтів у Twitter. Наприклад,

художники, які працюють у Twitter, щоб рекламувати своє мистецтво, часто надають зовнішню URL-адресу свого портфоліо.

Таблиця 2.2 - Аналіз числових атрибутів профілю та функцій, пов’язаних із конфіденційністю

Binary Attributes	Privacy Ratio (Spearman ρ)	Privacy Setting (T-test P-value)
Favorite Count	+0.27	+++
*Tweet Count	+0.30	+++
Follower Count	+0.15	+++
Friend Count	+0.02	+++

Для P < .05 і +++ для P < .005, коли значення більші для захищених облікових записів, тоді як - і - - - використовуються, коли значення менші для захищених облікових записів. Змінні, для яких спостерігається хоча б невеликий розмір ефекту (d Коена > 0,2), позначені зірочкою. захищені контакти. Наслідки цього висновку подвійні. По-перше, це вказує на існування місцевості для налаштувань конфіденційності. По-друге, цей результат означає, що такі користувачі почуваються в безпеці в середовищі з точки зору обміну інформацією, але вони більше піклуються про конфіденційність, ніж інші публічні облікові записи. Таким чином, очікується, що вони, швидше за все, відчують вторгнення цільової реклами та маркетингових повідомлень.

Використання мовного вмісту

Показано, що природна мова є відображенням і посередником внутрішніх станів [6]. Наші слова можуть розкривати особистість, емоційні стани та почуття, моделі уваги, думки та соціальні ситуації [11]. Тому було розроблено різноманітні методи автоматизованого аналізу контенту для вимірювання таких психометричних показників природною мовою. Ці методи варіюються від використання попередньо визначених словників і таксономій,

таких як LIWC, до складних обчислювальних алгоритмів, які часто використовують методи аналізу даних і машинного навчання.

Словники LIWC здатні надати широкий спектр соціальних і психологічних ідей з мови. Тому ми використали LIWC, щоб проаналізувати мову твітів і дослідити зв'язки між набором лінгвістичних показників і поведінкою користувачів щодо конфіденційності. LIWC має компонент обробки, який вивчає текстовий файл слово за словом. Потім кожне слово порівнюється з вбудованими словниками. Враховуючи, що словники LIWC структуровані в ієрархічному форматі, компонент обробки потім визначає, до яких категорій або підкатегорій LIWC належить слово. Після обробки всіх слів LIWC виводить відсоток слів, які належать до певної категорії, до загальної кількості слів у тексті. Крім того, набір змінних LIWC вимірюється незалежно від словників і називається підсумковими змінними. Ці змінні включають чотири непрозорі мовні змінні (аналітичне мислення, вплив, автентичність та емоційний тон) і загальні описові характеристики тексту (слів у реченні та відсоток слів, які довші за шість літер). Визначення та приклади кожної з цих категорій можна знайти на веб-сайті LIWC.

Набори твітів, опубліковані кожним користувачем, спочатку очищаються та попередньо обробляються. Наприклад, elisions обробляються (наприклад, I'm –> I am), URL-адреси та згадки в Twitter замінюються певними маркерами, а смайли замінюються відповідним значенням (наприклад, «:)» –> усмішка). Тоді всі зібрані твіти, опубліковані користувачем, розглядаються як єдиний документ і передаються LIWC для аналізу. Відсотки, розраховані LIWC, потім вивчаються з точки зору їх співвідношення з коефіцієнтом конфіденційності. У таблиці 4.3 узагальнено результати кореляції для категорій LIWC та підсумкових змінних, які ранжуються на основі їхньої сили кореляції. Таблиця містить лише ті змінні, у яких коефіцієнт кореляції перевищує певний поріг ($\rho > 0{,}20$ і $P < 0{,}005$).

Більшість функцій LIWC, які позитивно корелюють із співвідношенням конфіденційності, можуть бути пов’язані з приватним вмістом відповідно до суспільного консенсусу. Прикладами цих особливостей є використання лайливих слів, вираження гніву та тривоги, а також сексуальні теми. Крім того, позитивна кореляція спостерігається для використання «Я» та співвідношення конфіденційності. Особисті займенники в основному з’являються в оповіданнях і твітах, які описують особисті події, почуття, думки тощо. Аналогічно, використання минулого часу часто спостерігається в більш приватних районах.

У LIWC функція аналітичного мислення фіксує ступінь, до якої фрагмент тексту представляє формальне, логічне та ієрархічне мислення. Аналітичне мислення негативно корелює з співвідношенням конфіденційності. Цей висновок можна віднести до професійних облікових записів (наприклад, належать артистам, політикам, спортсменам тощо), які часто розташовані в громадських районах і зазвичай можуть використовувати формальну та логічну мову.

Відповідна категорія серед результатів LIWC називається автентичністю, яка фіксує ступінь більш чесної, особистої та відвертої мови. Незважаючи на те, що кореляція нижче нашого порогу і, таким чином, не включена в таблицю ($\rho = +0{,}22$), показано, що автентичність мови позитивно корелює з коефіцієнтом конфіденційності.

Знову ж таки, цей висновок показує, що люди, які знаходяться в приватних районах, імовірно, будуть стурбовані конфіденційністю, але вони, ймовірно, не знають про конфіденційність і, таким чином, публікують конфіденційну інформацію про себе.

Як обговорювалося раніше, твіти, опубліковані захищеними обліковими записами, недоступні, що унеможливлює порівняння захищеного вмісту з їх загальнодоступними аналогами.

Таблиця 2.3 - Кореляційний аналіз категорій LIWC та коефіцієнтів конфіденційності

LIWC Feature	Privacy Ratio (Spearman ρ)
Swear Words	+0.40
Anger	+0.35
Negative Emotions	+0.34
Body	+0.32
Negations	+0.31
Adverbs	+0.30
Sexual	+0.29
Sad	+0.27
Analytical Thinking	-0.26
FocusPast	+0.26
Interrogative	+0.26
Feel	+0.26
I	+0.26
Pronoun	+0.25
Anxiety	+0.25

Однак існує доступний компонент, який може представляти лінгвістичні характеристики захищених облікових записів: описи профілів. Налаштовуючи атрибути свого профілю, користувачі Twitter можуть вказати до 160 символів у полі опису. У нашій групі підписників CNN є майже 500 тис. захищених облікових записів і приблизно 500 тис. публічних облікових записів, які мають описи. Ми використовували LIWC для аналізу мовних категорій у цих описах. Оскільки описи профілів часто дуже короткі (зазвичай 8-10 слів), відсотки, надані LIWC, дуже малі для більшості категорій. Тому ми зосередилися лише на категоріях вищого рівня, які знаходяться у верхній частині ієрархії LIWC, а також на підсумкових змінних. У таблиці 2.4 показано підсумок результатів.

Таблиця 2.4 - Аналіз лінгвістичних показників та функцій, пов'язаних із конфіденційністю

LIWC Feature	Privacy Ratio (Spearman ρ)	Privacy Setting (T-test P-value)
Analytical Thinking	-0.26	- - -
Authentic	+0.22	- - -
Clout	-0.15	+++
Function Words	+0.23	+++
Affect Words	+0.13	+++

Два базових набори даних представляють дві різні групи користувачів, які різними способами збираються з Twitter. Крім того, мовний зміст твітів та описів надається для різних цілей. Однак ми все ще бачимо ту саму поведінку захищених облікових записів і публічних облікових записів, які підключені до великого відсотка захищених сусідів з точки зору використання ними мови (див. Таблиці 2.3 і 2.4). Наприклад, автентичність, використання функціональних слів, афектних слів, соціальних процесів і когнітивних процесів позитивно корелюють із співвідношенням конфіденційності. Так само вони частіше спостерігаються в описах профілів захищених облікових записів. З іншого боку, аналітичне мислення, вплив, драйвери та потреби негативно корелюють із коефіцієнтом конфіденційності та рідше спостерігаються в захищених облікових записах. Знову ж таки, така подібність сигналізує про наявність місцевості для поведінки користувачів щодо конфіденційності. Тому особливості, характерні для загальнодоступних облікових записів, які підключають велику кількість приватних контактів, можна вважати особливостями, які характеризують користувачів, які піклуються про конфіденційність.

LIWC фіксує відсоток слів, які належать до різних категорій, пов’язаних із настроями (наприклад, позитивні та негативні слова, гнів та тривога) у наборах твітів користувачів. На додаток до цих категорій LIWC ми

скористалися перевагами лексичного ресурсу під назвою SentiWordNet [3], щоб проаналізувати настрої твітів і функції конфіденційності з іншої точки зору. У SentiWordNet кожне слово отримує три оцінки: позитивність, негативність та об'єктивність. Спочатку ми очистили та токенізували кожен твіт. Потім токени позначаються POS-тегами та об'єднуються. Отримана пара-тег токен-POS потім порівнюється з SentiWordNet. Нарешті, оцінки, отримані з SentiWordNet, потім агрегуються для всіх маркерів у твіті, щоб створити загальну оцінку настроїв твіту. Слід зазначити, що щоразу, коли спостерігається заперечення, враховується обернена оцінка наступного маркера. Після того, як були визначені оцінки твітів, ми розрахували співвідношення позитивних і негативних твітів до загальної кількості твітів. Виходячи з аналізу, співвідношення позитивних твітів має невелику негативну кореляцію з коефіцієнтом конфіденційності ($\rho = -0{,}12$), тоді як співвідношення негативних твітів позитивно корелює з коефіцієнтом конфіденційності ($\rho = +0{,}26$). Кількість заперечень, помічених у твітах, також позитивно корелює з коефіцієнтом конфіденційності ($\rho = 0{,}28$). Через недоступність твітів, опублікованих захищеним обліковим записом, прямий аналіз захищених і публічних облікових записів з точки зору співвідношення твітів з різними мітками настроїв неможливий.

Комунікаційна поведінка

Ми використали набір простих змінних, щоб охарактеризувати комунікаційну поведінку користувачів за їхніми часовими шкалами. У розділі 2 ми спостерігали позитивну кореляцію між кількістю твітів користувачів і їх співвідношенням конфіденційності. Крім частоти твітів, ми дослідили середню довжину твітів і виявили негативну кореляцію з коефіцієнтом конфіденційності ($\rho = -0{,}27$). Опубліковані твіти можуть бути ретвітами з інших облікових записів або новими твітами, які надходять з облікового запису у фокусі. Зауважте, що ретвіти виключаються з аналізу довжини твітів. Кореляція між відношенням ретвітів до загальної кількості твітів і

коефіцієнтом конфіденційності дуже мала (ρ = +0,04), а співвідношення нових твітів також позитивно корелює з коефіцієнтом конфіденційності (ρ = +0,15). Крім того, користувачі твітують, щоб взаємодіяти один з одним і брати участь у розмовах. Співвідношення цих розмовних твітів також показує позитивну кореляцію з коефіцієнтом конфіденційності (ρ = +0,22). Іншою потенційно цікавою змінною для дослідження є використання URL-адрес під час твітів. Ми очікуємо, що професійні та неособисті облікові записи будуть частіше публікувати новини та події, які часто пов’язані з URL-адресами. Як і очікувалося, виявляється, що співвідношення твітів із URL-адресами негативно корелює з коефіцієнтом конфіденційності (ρ = -0,22). Ми виявили незначну кореляцію між моделями використання хештегу та співвідношенням конфіденційності. Подібно до настроїв твітів, прямий аналіз захищених і публічних облікових записів неможливий через обмеження конфіденційності, пов’язані із захищеними обліковими записами.

Приховані атрибути

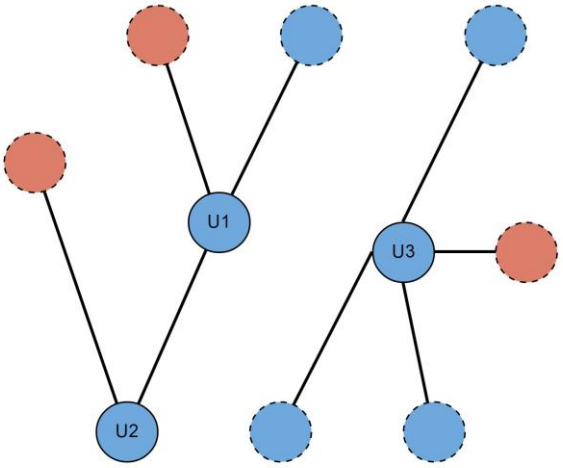
Щоб виявити набір прихованих атрибутів, які цікавлять користувачів, які піклуються про конфіденційність, ми перетворили кожен вузол користувача в мережі в набір атрибутів. Потім для кожного вузла атрибутів ми розрахували відношення захищених контактів до загальної кількості контактів. Отримана мережа дозволяє нам зрозуміти, які функції приваблюють користувачів, які піклуються про конфіденційність, а які з них частіше спостерігаються в громадських районах. На рисунку показано цю процедуру перетворення для зразкової мережі, в якій загальнодоступні вузли закодовані синім, а захищені — червоним. Припустимо, що в мережі є три користувача: U1, U2 і U3. Вузли з пунктирними межами — це соціальні контакти користувачів, які ще не додані до мережі, але враховуються в процесі обчислення метаданих і додаються до черги BFS. Як показано на рисунку , U1 пов'язаний з трьома соціальними контактами, серед яких два публічні, а один

захищений; отже, йому буде надано коефіцієнт конфіденційності θu1 = 33%. Аналогічно, U2 і U3 отримують значення θu2 = 50% і θu3 = 25% відповідно.

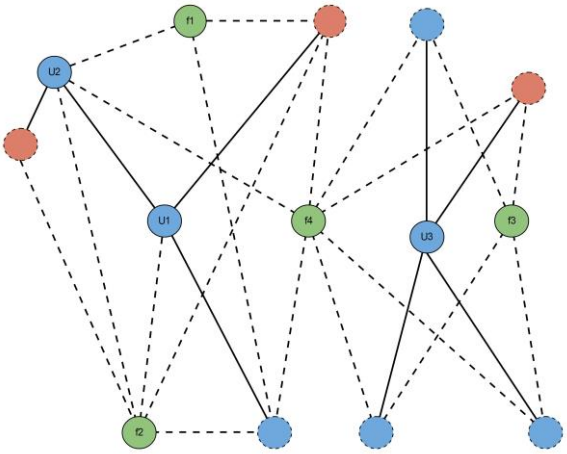
Тепер припустимо, що U1 можна охарактеризувати трьома ознаками такими, що U1 = { f1, f2, f4}. Потім ці три особливості пов’язуються з усіма трьома соціальними контактами U1. Тоді, якщо U2 = { f2}, f2 також буде пов’язано з сусідами U2. Нарешті, припустимо, що ми маємо U3 = { f1, f3, f4} у нашому прикладі мережі. На рисунку показано процес зв'язування ознак з контактами. Отримана мережа тоді буде двосторонньою мережею контактів і функцій, у якій можна розрахувати коефіцієнти конфіденційності для функцій. Наприклад, f1 є ознакою, яка характеризує користувачів U1 і U3 і не пов’язана з U2. Тому на малюнках він пов’язаний із сусідами U1 і U3. Таким чином, f1 пов’язаний із п’ятьма загальнодоступними користувачами та двома захищеними, що призводить до коефіцієнта конфіденційності θf1 = 40%.

Для опису облікових записів із хронології Twitter можна витягти різноманітні функції. Ми провели експерименти з чотирма основними функціями: 1) уніграми твітів, 2) хештеги, 3) акаунти в Twitter, з яких користувачі ретвітують, і 4) теми. Кожен з цих наборів функцій може охопити та виявити певний аспект поведінки та інтересу користувачів. Уніграми витягуються за допомогою Apache Lucene, добре відомої бібліотеки текстового пошукового механізму, розробленої на Java. Хештеги та джерела ретвітів просто витягуються за допомогою регулярних виразів. Нарешті, реалізація прихованого розподілу Діріхле (LDA) у Mallet використовується для вилучення та аналізу тем, які обговорюються у твітах.

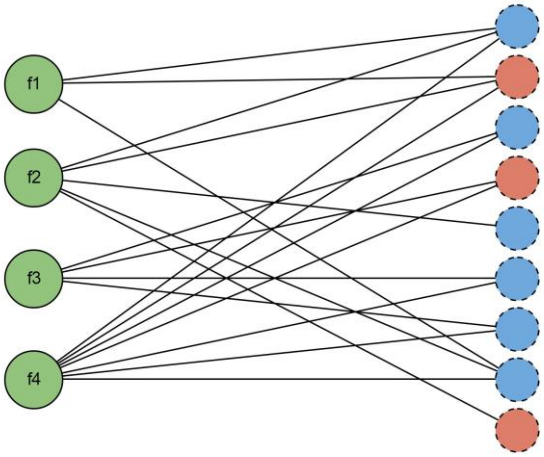
У таблиці 2.5 наведено огляд статистики функцій у нашому наборі даних. Щоб отримати надійні коефіцієнти конфіденційності для функцій, їх має використовувати значна кількість користувачів. Тому ми відфільтрували списки лише для включення тих функцій, якими користуються щонайменше 100 користувачів із базового набору. Ця фільтрація привела до 1228 хештегів, 906 джерел ретвітів,



(a)



(b)



(c)

Рисунок 2.5 - Огляд процесу трансформації мережі

Таблиця 2.5 - Описова статистика функцій, які представляють часові шкали користувачів

Feature	Unique Instances	Filtered Instances	Privacy Ratio (Average)	Privacy Ratio (Variance)
Unigrams	1,094,642	23470	9.15	6.10
Hashtags	514,186	1228	9.43	10.29
Retweet Source	655,384	906	11.69	16.74
Topics	150 (pre-specified)	32	9.09	1.30

Останні два стовпці таблиці представляють середнє значення та дисперсію коефіцієнтів конфіденційності для цих відфільтрованих списків.

Таблиця показує набір прикладів функцій, які розташовані на двох кінцях спектру відношення конфіденційності. Як бачимо, певні функції в списку є специфічними для набору даних у фокусі. Наприклад, @CityOfHamilton і @Kathleen Wayne показують, що значна кількість користувачів, вибраних у процесі збору даних, перебуває в Канаді. Це обмеження, однак, є загальною проблемою, від якої страждають підходи, керовані даними, і буде зменшуватися в міру збору більше даних. Крім того, використання онтологій для відображення цих функцій у просторі вищого рівня може вирішити цю проблему і буде розглянуто в наших майбутніх планах дослідження. Тим не менш, ми зосередимося на тому, щоб отримати уявлення про загальні закономірності цих функцій та їх співвідношення конфіденційності. Погляд на таблицю прикладів показує, що функції, чутливі до конфіденційності, мають високі показники конфіденційності, тоді як приклади на іншому кінці часто можна побачити у професійному та неособистому контенті. Ці результати узгоджуються з нашими попередніми висновками з аналізу характеристик профілю та атрибутів LIWC. Враховуючи, що цей підхід зосереджений лише на стосунках із користувачами та їх налаштуваннях конфіденційності, наш аналіз уніграм і хештегів можна проводити незалежно від мови твітів. Наприклад, usluge, що означає

обслуговування боснійською мовою, є уніграмою з дуже низьким показником конфіденційності (θ = 2,59) у нашому списку.

Як пояснювалося раніше, ми використовували LDA для створення тем із хронології користувачів. Для побудови тематичної моделі LDA вимагає трьох вхідних параметрів, які необхідно вказати: кількість тем, які будуть створені (часто позначається k), значення α і значення β. α представляє щільність тем документа. Таким чином, вищий α означає, що документи містять більше тем, а нижчий α означає, що документи містять менше тем. β представляє щільність тематичних слів. Тому високе значення β означає, що теми складаються з більшості слів у корпусі, а низьке значення β означає, що вони складаються з кількох слів. Ми побудували тематичні моделі з різною кількістю тем (k = 50, 100, 150 і 200). Для кожної моделі ми обрали значення α=50/k і β=0,01, як запропоновано.

Таблиця 2.6 - Приклад функцій, витягнутих із часових шкал користувачів разом із відповідними показниками конфіденційності

Feature Type	Top Features	Privacy Ratio	Bottom Features	Privacy Ratio
Unigrams	hooka	17.60	consultation	3.43
	shittiest	16.69	workshops	3.33
	hungover	16.63	catering	2.94
Hashtags	#TakeMeBack	19.76	#Adventure	3.88
	#SoTired	16.87	#ShopLocal	3.30
	#MissYou	16.64	#Entrepreneur	1.87
Retweet Source	@ColiegeStudent	17.5	@Kathleen Wayne	4.60
	@FemalePains	17.10	@CityOfHamilton	4.05
	@MensHumour	16.91	@DanceMoms	4.02

Наш кореляційний аналіз співвідношення конфіденційності в різних районах показує, що параметри конфіденційності локалізовані. Крім того, анотовані людиною дані, зібрані від AMT, надали додаткову підтримку, щоб виправдати уподобання щодо конфіденційності місцевості. Нарешті, аналіз загальнодоступних і захищених облікових записів та їх порівняння з

користувачами, розташованими в різних районах із різними налаштуваннями конфіденційності, дали додаткові підказки для підтвердження місцевості в цьому контексті. Цей висновок зроблено насамперед на основі подібності, що спостерігається між захищеними рахунками та державними рахунками, які пов'язані з більшою часткою захищених сусідів. Наскільки нам відомо, дослідження, проведене CaliskanIslam et al. [4] — єдине дослідження, яке вивчало потенційні зв'язки між функціями конфіденційності користувачів і атрибутами конфіденційності їхніх сусідів у соціальних мережах онлайн. Згідно з нашими висновками, їхній аналіз 45 користувачів Twitter показав, що кількість приватної інформації, якою поділилися користувачі, позитивно корелює з кількістю приватної інформації, якою поділилися їхні сусіди.

Висновок, що параметри конфіденційності є локалізованими, може вказувати на те, що функції конфіденційності є гомофільними в контексті Twitter. Крім того, на переваги користувачів Twitter можуть впливати їхні близькі сусіди. Необхідно провести подальші поглиблені дослідження, щоб виявити потенційний вплив кожного з цих двох процесів на налаштування конфіденційності. Більш ранні соціологічні дослідження виділяють два типи : статусну гомофілію та гомофілію цінностей. Статусна ґрунтується на неформальних, формальних або приписаних статусах людей і включає соціально-демографічні виміри, такі як раса, етнічна приналежність, вік, стать та місце розташування. Набуті характеристики, такі як релігія, освіта та професія, також класифікуються як статусна. Статусна широко спостерігається в контексті соціальних мереж в Інтернеті. Крім того, існує велика кількість досліджень, які пов'язують соціально-демографічну інформацію з уподобаннями щодо конфіденційності. Наприклад, різні опитування встановили позитивне співвідношення віку, освіти та доходу, пов'язане з проблемами приватності. Таким чином, схожість між уподобаннями соціальних контактів щодо приватності може випливати з їх схожості за цими гомофільними характеристиками. З іншого боку, цінностей

включає широкий спектр внутрішніх станів, які, як передбачається, формують нашу орієнтацію на майбутню поведінку. Таким чином, подібність конфіденційності може бути прямим результатом цінностей, яка випливає з цінностей користувачів і турботи про конфіденційність. У результаті нашого дослідження можуть виникнути такі цікаві дослідницькі питання: чи це цінність та турбота про конфіденційність керує та впливає на зв'язки? Або подібність конфіденційності є лише похідною від статусної ? А може, обидва? Для вирішення цих питань необхідні подальші поглиблені дослідження конфіденційності.

Наші базові набори даних можуть страждати від проблеми дихотомії конфіденційності, і ми визнаємо, що ця проблема призведе до деяких помилок. Однак ми спробували мінімізувати цю помилку, вибравши набір даних CNN для прямого аналізу захищених і публічних облікових записів, і виявили статистично значущі відмінності між двома групами. Наш мережевий підхід також залежить від поточних налаштувань конфіденційності контактів, тому він схильний до помилок, спричинених розбіжністю між ставленнями користувачів до конфіденційності та налаштування. Проте, об'єднавши кілька міток конфіденційності від усіх контактів, ми сподіваємося зменшити наслідки цієї проблеми та ефективно охопити загальні функції конфіденційності. Крім того, наші дослідження кореляції засновані на відносних співвідношеннях конфіденційності. Тому, припускаючи, що проблема дихотомії конфіденційності рівномірно розподілена по мережі, наш мережевий підхід повинен мати можливість охопити загальні переваги конфіденційності в різних районах. Відмінності, виявлені між атрибутами користувачів у більш приватних районах, можуть бути ознакою успіху підходу. Крім того, незважаючи на відмінності в двох наборах користувачів, розумні результати, отримані в результаті порівняння захищених облікових записів у даних CNN, і користувачів із більшим коефіцієнтом конфіденційності можуть підтвердити нашу заяву.

Одним з основних обмежень цього дослідження є вивчення переваги конфіденційності як двійкової змінної. Проте попередні дослідження конфіденційності показали, що налаштування та рішення щодо конфіденційності набагато складніші. Фактично, користувачі можуть використовувати різноманітні стратегії захисту конфіденційності, такі як самоцензура або вибіркове поширення на веб-сайтах соціальних мереж [9]. Вивчення таких стратегій захисту конфіденційності та їх зв'язку з характеристиками користувачів у мережі є виправданим. Крім того, наше дослідження зосереджено лише на Twitter, а дизайн веб-сайту соціальної мережі, як відомо, формує та впливає на поведінку користувачів [2]. Вивчення інших платформ з більшим набором користувачів може дати уявлення про узагальнюваність підходу і включено в наші майбутні плани дослідження.

2.4 Визначення параметрів конфіденційності за допомогою спільної фільтрації

Надання персоналізованого вмісту може принести взаємну користь як компаніям, так і клієнтам. Однак ефективне профілювання користувачів залежить від збору великої кількості даних від користувачів. Отже, конфіденційність є наріжним каменем будь-якої діяльності з персоналізації, яка, якщо її не враховувати, вплине на задоволення, отримане від персоналізації, і може призвести до непоправних проблем з довірою. Це питання може бути особливо проблематичним у контексті соціальних мереж, де сьогодні відбувається величезна кількість щоденних соціальних взаємодій.

Незважаючи на те, що користувачі мають можливість захищати свої дані за допомогою засобів контролю конфіденційності, такі рішення щодо конфіденційності часто є складними і вимагають ретельного вивчення компромісів між потенційною соціальною вигодою та можливими ризиками конфіденційності. Тому багато користувачів уникають клопоту з

налаштуванням конфіденційності та дотримуються налаштувань за замовчуванням [2]. Однак вони зазвичай не знають, що налаштування за замовчуванням часто відкриті та дозволені. Навіть серед тих, хто докладає зусиль, щоб керувати своєю конфіденційністю, багато хто все ще не усвідомлюють наслідків своїх рішень [7].

Незважаючи на обмежену кількість, попередні дослідження поведінки щодо конфіденційності мали скромний успіх у прогнозуванні уподобань користувачів щодо конфіденційності, покладаючись на їхні соціальні сліди [4, 2]. Тут ми пропонуємо підхід спільної фільтрації (CF), який поєднує методи на основі сусідства з моделлю латентного фактора, виведеною з соціального графіка користувачів. Наші

Версія цієї глави була опублікована в матеріалах Міжнародної конференції з Інтернету та соціальних медіа (ICWSM) як стендовий документ.

Підхід прогнозування конфіденційності є новим і спрямований на виявлення користувачів, які піклуються про конфіденційність, з загальнодоступними профілями. Внесок цієї статті полягає в наступному: 1) адаптація гібридного методу CF для визначення переваг соціальної конфіденційності, 2) вивчення корисності атрибутів профілю для визначення переваг конфіденційності, 3) встановлення та аналіз властивостей соціального графіка з покращеною конфіденційністю, і 4) виявлення набору прихованих факторів, пов'язаних з атрибутами конфіденційності.

Наше дослідження проведено в Twitter, де контроль конфіденційності відповідає бінарній специфікації. У Twitter користувачі можуть або слідувати загальнодоступним налаштуванням за замовчуванням, які вказують на те, що їхні твіти та контакти є загальнодоступними, або вони можуть змінити налаштування на захищені, що робить їхні твіти та контакти доступними лише для їх затверджених підписників.

Методи CF показали великі перспективи при розробці

рекомендаційних систем, де невідомі переваги користувачів визначаються за допомогою відомих переваг інших користувачів. Такі підходи можуть бути особливо цінними в контексті соціальних медіа через наявність додаткових соціальних відносин. Проте було зроблено обмежені спроби адаптувати методи CF для прогнозування переваг конфіденційності.

Наприклад, Squicciarini et al. спочатку формують соціальні кола на основі характеристик користувачів (наприклад, стать та хобі). Коли основний користувач завантажує новий об’єкт, система шукає соціальні кола, які, швидше за все, мають справу з об’єктом так само, як і користувач. Тоді політика конфіденційності, яку використовує вибране коло, є основою для прогнозування політики для доданого об’єкта.

У активне навчання та властивості соціального графа вперше використовуються для виявлення набору інформаційних контактів, які мають бути позначені як навчальні зразки. У процесі маркування користувач вказує, чи бажає він/вона поділитися певним елементом із вибраним контактом. Потім мітки поширюються від позначених екземплярів до немічених на графіку. Це поширення керується показником подібності користувачів, який обчислюється на основі інформації профілю контактів, а також показників їх мережі та спільноти.

CF також дотримується в [3], де набір функцій профілю, інтересів користувачів і конфігурації конфіденційності використовується для пошуку набору користувачів, подібних до основного користувача. У своїй роботі користувачів спочатку характеризують відповідно до їхніх уподобань щодо конфіденційності як фундаменталістів конфіденційності, прагматиків конфіденційності або байдужих до конфіденційності. Користувачів призначають до цих категорій на основі кількості їхніх загальнодоступних, налаштованих та приватних фотоальбомів. Потім використовується алгоритм K-найближчого сусіда, щоб визначити, до якої категоризації конфіденційності належить основний користувач.

Методи прогнозування конфіденційності, засновані на CF, в основному використовують методи на основі сусідства, де основою прогнозу є спостереження за конфіденційністю від сусідніх або подібних користувачів. Однак моделі латентного фактора, які є сучасним станом для МВ [1], ще не досліджені. У попередніх роботах також бракує вивчення гібридних методів, які показали значно кращі результати порівняно з моделями чистого сусідства та чисто латентного фактора в інших областях [6].

2.5 Модель латентного фактора на основі сусідства

Двома популярними методами CF є методи на основі сусідства та моделі латентних факторів. У методах, заснованих на сусідстві, спостереження сусідніх та/або подібних користувачів використовуються для виявлення атрибутів і переваг основного користувача. Виявлення налаштувань конфіденційності за допомогою таких методів, орієнтованих на користувача, створює унікальні проблеми через відсутність доступної інформації, пов’язаної з приватними обліковими записами. У Twitter обмежена кількість атрибутів профілю доступна як для загальнодоступних, так і для захищених облікових записів. Отже, такі атрибути профілю можна використовувати для вимірювання схожості користувачів.

Виходячи з цієї формули, більше значення уі вказує на більш високий рівень занепокоєння конфіденційністю користувача і. Щоб отримати уявлення про можливості та обмеження атрибутів профілю Twitter для нашого завдання, ми провели ряд експериментів. Ці експерименти, які коротко представлені в наступному розділі, припускають значення атрибутів профілю у висновку про переваги конфіденційності та вказують на їх потенціал для вимірювання подібності в підході на основі сусідства.

Тим часом моделі латентного фактора здобули величезний успіх у контексті рекомендаційних систем. Такі методи мають на меті виявити набір

інформаційних прихованих факторів щодо користувачів і пізніше використовувати такі атрибути для визначення уподобань. Ми можемо обчислити ймовірність того, що публічний користувач і піклується про конфіденційність, використовуючи наступне:

$$p(yi) = p(yi|\theta 1,...,\theta k) = \Pi j=1,..,kp(yi|\theta j)$$

2.6 Атрибути профілю для передбачення конфіденційності

Щоб проаналізувати зв'язок між атрибутами профілю та поведінкою конфіденційності, ми спочатку створили каталог користувачів Twitter, зібравши підписників кількох відомих акаунтів у Twitter (наприклад, «Facebook», «Katy Perry», «Obama»). Потім для кожного облікового запису ми обчислили відсоток захищених підписників до загальної кількості підписників. Результати показують, що цей відсоток значно вищий для «Крайні новини CNN» (11%) у порівнянні з іншими наборами підписників (від 5% до 7%) та середнім відсотком у Twitter (4,8%). Таким чином, ми вибрали цей набір для нашого дослідження, оскільки дихотомія ставлення до конфіденційності та поведінки, здається, зведена до мінімуму.

Потім ми проаналізували цей набір користувачів, який включає збалансований набір із майже 1 мільйона користувачів, щоб отримати уявлення про потенційні відмінності в тому, як налаштовуються атрибути профілю загальнодоступних і захищених облікових записів. Ми зосередили свій аналіз на наборі функцій профілю, які легко доступні з облікових записів Twitter, а також додаткових функцій, розроблених на основі наявних атрибутів профілю. Наприклад, ми використали каталог англійських імен, щоб проаналізувати, чи оголошене ім'я містить справжнє ім'я людини. Крім того, ми дослідили лінгвістичні атрибути описів профілів за допомогою LIWC та аналізу частоти ключових слів.

В результаті розроблено набір ознак розміру, зведення якого наведено в таблиці 2.8. Перший стовпець таблиці показує функції Twitter API. У другому стовпці представлені лінгвістичні атрибути, отримані з описів профілів. На додаток до набору категорій LIWC2, цей список містить чотири функції на основі ключових слів, які вказують на наявність відповідного ключового слова в описі. Відмінності між захищеними та публічними обліковими записами є статистично значущими для всіх цих функцій (визначено за результатами хі-квадрат або t-тест залежно від типу функції). Для чотирьох із цих функцій відмінності також є практично суттєвими (що визначають за V Крамера або d Коена залежно від типу ознаки).

Таблиця 2.8 - Функції профілю для виявлення захищених облікових записів

Атрибути Twitter	Лінгвістичні атрибути
Кількість твітів*	Слова з шістьма буквами
Друг граф	Функціональні слова
Улюблений граф	Clout
Кількість списків	Емоційний тон
Фактична кількість імен	Автентичний
Гео-ввімкнено*	Аналітичне мислення
Має справжню назву	Вплив на слова
Ім'я користувача має ім'я	Соціальні процеси
Має URL-адресу	Когнітивні процеси
Має місцезнаходження*	Відносність
Має зображення за замовчуванням	Водії та потреби
Має профіль за замовчуванням*	«слідувати»

Ці чотири ознаки позначені зірочкою в таблиці. Таким чином, атрибути профілю є відмінними для захищених і загальнодоступних користувачів і виявилися корисними для нашого завдання визначення переваг конфіденційності в підході на основі сусідства. Використання цього набору функцій у контрольованому алгоритмі на основі регресії призвело до F-показу

0,72, що перевершує випадковий базовий рівень більш ніж на 20%. Деталі набору функцій та експериментів з машинним навчанням можна знайти в [5]

2.7 Графік конфіденційності та приховані атрибути

Побудова графіка та властивості Наш підхід спрямований на експлуатацію [8] і заснований на тому, що люди подібних інтересів мають тенденцію до спілкування один з одним. Тому замість того, щоб використовувати асиметричне відношення підписки чи друзів у Twitter, ми побудували ненаправлений взаємний граф користувачів, який включає лише ребра, які є взаємними. Очікується, що взаємні відносини вказують на міцніші стосунки між двома користувачами, і вони відрізняють розділ соціальної мережі сфери Twitter від її інформаційної мережі [9].

Починаючи з випадкового загальнодоступного користувача, ми ітераційно побудували взаємний графік користувачів методом пошуку в ширину (BFS). Для кожного загальнодоступного користувача ми спочатку підрахували кількість захищених спільних сусідів, а також відношення захищених до всіх спільних сусідів.

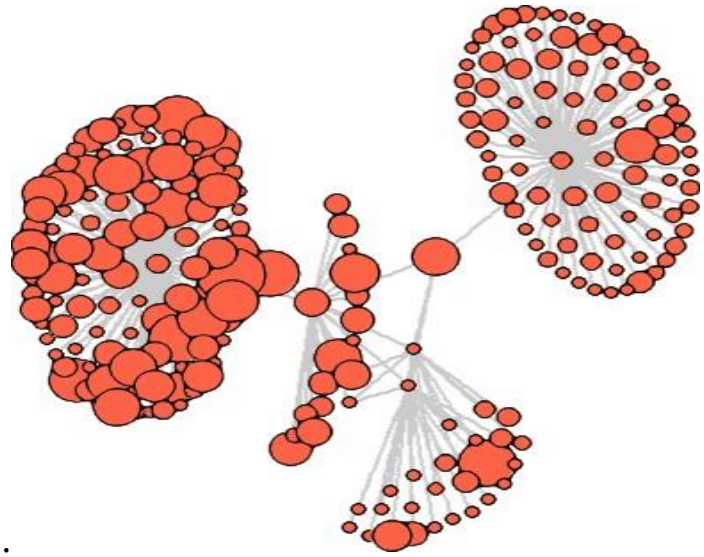


Рисунок 2.6 - Знімок графіка конфіденційності

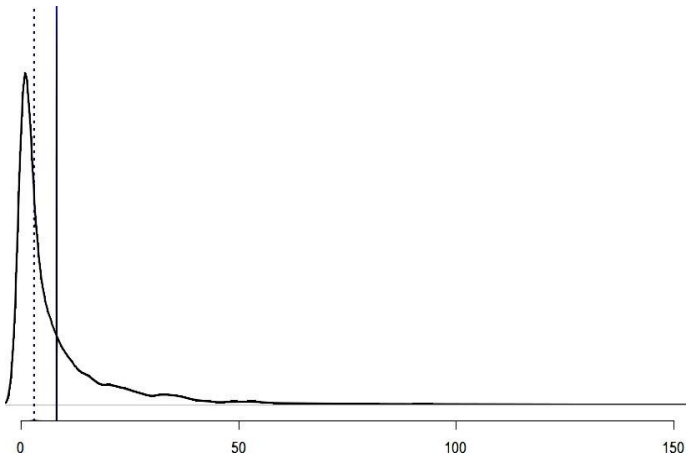


Рисунок 2.7 - Розподіл захищених контактів

Цей користувач потім анотується цими показниками та додається до графіка. Потім ми перевіряємо, чи має новий вузол взаємні зв'язки з будь-яким існуючим вузлом в графі, і додаємо відповідні ребра. Цей процес повторюється з новим відкритим користувачем, вилученим із черги BFS. Користувачі з менше ніж 10 твітів або менше ніж 30 підписників/друзів вважаються неактивними і, таким чином, не додаються до графіка. Крім того, не включаються перевірені користувачі та користувачі з понад 1000 підписників/друзів, оскільки вони часто представляють бренди та знаменитостей і не є представниками широкого загалу. Ми зібрали 3К загальнодоступних вузлів, які анотовані на основі їх показника коефіцієнта конфіденційності. На рисунку показано знімок невеликої частини графіка, візуалізованого за допомогою примусово орієнтованого макета, де показник коефіцієнта конфіденційності зіставляється з розміром вузла. У цьому наборі даних кожен обліковий запис Twitter взаємно пов'язаний із в середньому 77 контактами. Серед цих сусідів у середньому 69 є державними та 8 захищеними. На рисунку показано розподіл захищених сусідів, де середнє та медіана відзначені суцільною та пунктирною лінією відповідно. Як видно, розподіл перекошений вправо, що вказує на те, що, незважаючи на меншу кількість користувачів із великою кількістю захищених сусідів, ці числа є значно великими, тому середнє значення перетягується вправо.

Щоб переконатися, що застосовується в контексті конфіденційності, ми розрахували кореляцію між коефіцієнтом конфіденційності кожного вузла та середнім коефіцієнтом конфіденційності сусідів. Цей аналіз близько 700 користувачів із принаймні 10 взаємними контактами на графіку показав сильну позитивну кореляцію між двома змінними (коефіцієнт кореляції Пірсона r = 0,88). Цей результат вказує на те, що на поведінку користувачів щодо конфіденційності впливають або їхні близькі соціальні контакти, або люди з подібною поведінкою щодо конфіденційності, як правило, об'єднуються в соціальні мережі. У будь-якому випадку, цей висновок передбачає великий потенціал CF для прогнозування переваг конфіденційності.

Щоб виявити набір прихованих атрибутів, які цікавлять користувачів, які піклуються про конфіденційність, ми перетворили кожен вузол у набір атрибутів. Потім для кожного вузла атрибутів ми обчислили кількість і співвідношення захищених сусідів. Очікується, що отриманий графік дозволить обчислити p(yi|θj), що вказує на ймовірність того, що користувач і буде приватним, якщо він/вона позначений атрибутом θj. Наприклад, таким значенням ймовірності може служити коефіцієнт конфіденційності кожного латентного фактора. Ми проводили експерименти, використовуючи уніграми та хештеги, витягнуті з твітів користувачів, як приховані змінні. Ці атрибути вибираються з 500 останніх твітів, опублікованих кожним користувачем. У таблиці наведено 10 найкращих хештегів та уніграм, які пов'язані щонайменше з 10% (n = 30) користувачів. Ці функції виділяються на основі показника коефіцієнта конфіденційності, який також наведено в таблиці.

Верхній хештег у таблиці — «#neverforget», який є пам'ятним політичним гаслом, який заохочує пам'ятати про національні та міжнародні трагедії. Високий коефіцієнт конфіденційності цього хештегу, а також найпопулярніші хештеги «#USA» та «#respect» можуть свідчити про те, що спільноти, зацікавлені в політичних темах, як правило, більш приватні. Цей

висновок є цікавим, і його можна розглядати разом із високим відсотком захищених послідовників CNN. Однак, враховуючи те, що графік було зібрано за певний проміжок часу, ці результати можуть залежати від часу. Ще один орієнтований на час хештег у списку – «#aquarius». Збір даних протягом тривалого часу та групування хештегів у теми високого рівня може пролити світло на потенційні зв’язки між інтересами користувачів у політиці та знаках зодіаку та їхніми перевагами щодо
конфіденційності.

Ми також дослідили зібрані твіти, щоб зрозуміти контекст, у якому використовується верхня уніграма «окремий». Більшість цих твітів пов’язані з любовним життям та стосунками користувачів. Цікаво, що «#relationshipgoals» і «#lastrelationshipnaughtme» мають високий рівень конфіденційності в нашому списку. Таким чином, можна зробити висновок, що користувачі, які розповідають про свої стосунки в соціальних мережах, швидше за все, піклуються про конфіденційність. Знову ж таки, групування хештегів у загальну концепцію та теми може дати більш точні результати.

3 РІШЕННЯ ЗАДАЧІ ОБЧИСЛЮВАЛЬНОГО АНАЛІЗУ КОЛЕКТИВНОЇ ІНТЕЛЕКТУАЛІЗАЦІЇ WEB

3.1. Концепція інтелектуалізації

Концепція інтелекту в людському мозку вже багато років активно досліджується в психологічних дослідженнях. За останні десятиліття в дослідженнях інтелекту відбувся зсув від розгляду пізнання лише як чогось у мозку однієї людини до соціального та колективного інтелекту, що виникає внаслідок локальних взаємодій у групах індивідів [8]. Добре вивченою формою інтелекту в соціальних структурах є феномен «ройового інтелекту», який широко спостерігається в природі. Дослідження ройового інтелекту пояснюють, як групи тварин можуть виконувати надзвичайно складні й інтелектуальні завдання, які виходять далеко за межі когнітивних здібностей окремих тварин.

Наприклад, колонії мурашок давно захоплювали вчених, демонструючи комплексне вирішення проблем та інтелектуальні функції, такі як пошук найкоротшого шляху до джерела їжі або найкраще гніздо серед багатьох варіантів [7]. На додаток до колоній мурах, зграї птахів, випасання тварин, ріст бактерій та зграї риб є іншими прикладами самоорганізованих живих груп, які демонструють ознаки інтелекту [8].

Незважаючи на те, що колективний інтелект у людей та ройовий інтелект у тварин мають спільну ознаку «емерджент», вони відрізняються двома основними ознаками [6]. По-перше, рої тварин включають когнітивно простих агентів, тоді як колективний інтелект включає агентів зі складними когнітивними здібностями, тобто людей. По-друге, агенти кожного типу використовують різні

Стигмергія — це форма непрямої взаємодії, яка відбувається через подальші модифікації спільного середовища. Наприклад, мурахи відкладають феромони, повертаючись до гнізда, коли знайшли джерело їжі. Потім феромон

стимулює інших мурах і йдуть по слідах із сильнішим запахом. Оскільки мурахи проходять коротші шляхи за менший час, їх спільні дії, нарешті, призводять до пошуку найкоротшого шляху до їжі.

На відміну від роїв тварин, взаємодія в людських колективах може бути заснована на стигмергі (тобто непряма) або пряма. Крім того, взаємодія обох форм може відбуватися через два різні канали: вербальне спілкування або невербальні дії. Дотримуючись цих двох бінарних відмінностей, взаємодії між людьми можна розділити на чотири різні групи: пряме вербальне спілкування, пряма невербальна дія, непряма вербальна комунікація та непряма невербальна дія. Обговорення та обговорення, пішоходи, які утворюють двоспрямовані смуги на переповненому тротуарі [4], спільне авторство та пошук оптимізованого маршруту через активне формування стежки [5] є традиційними прикладами кожної форми відповідно.

Тим часом прогрес у комп'ютерно-опосередкованій комунікації та поява Web 2.0 уможливили масовий обмін даними у глобальному масштабі, що призвело до нових і передових форм колективного інтелекту. Вікіпедія, Google, форуми запитань і відповідей та соціальні мережі – все це приклади веб-платформ колективної розвідки. Наприклад, Вікіпедія використовує словесні взаємодії на основі стигмергії, коли користувачі можуть змінювати локальні частини свого спільного віртуального середовища. У проектах з відкритим кодом дії учасників накладаються один на одного для створення складних продуктів, таких як операційна система Linux. З іншого боку, у таких середовищах, як форуми питань і відповідей, соціальні мережі та системи обговорення, колективний інтелект може бути викликаний прямим спілкуванням та онлайн-діалогом між користувачами. Рисунок ілюструє різні типи взаємодії в колективному інтелекті разом із відповідними прикладами на основі Інтернету.

Платформи колективного інтелекту з підтримкою Інтернет дозволяють натовпу збиратися і вносити свій внесок у вирішення проблем, які може бути важко або неможливо вирішити навіть найрозумнішим людям і найшвидшим

комп’ютерам. Різні аналітичні методи розроблені, щоб зрозуміти, як інтелект виникає в соціальних взаємодіях, і визначити різні фактори мікро- та макрорівня, які можуть позитивно чи негативно впливати на феномен колективного інтелекту.

Отже, основним напрямком цієї роботи є виконання емпіричного аналізу різноманітних і великомасштабних наборів даних, створених користувачами, а також розробка обчислювальних інструментів і моделей для кращого розуміння колективного інтелекту. Отримана аналітика буде потім використовуватися разом із візуалізацією інформації та принципами взаємодії людини з комп’ютером для проектування та розробляти веб-інструменти, які ефективно використовують інтелектуальну силу людських колективів.

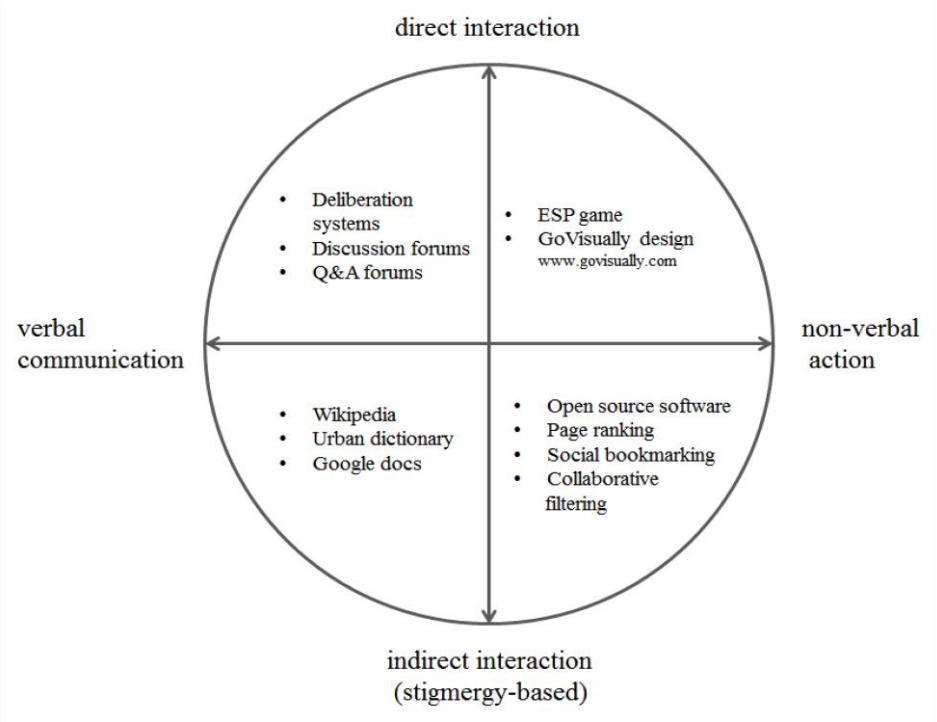


Рисунок 3.1 - Колективний інтелект можна розділити на чотири різні групи на основі типів взаємодії між людьми

Це дослідження в основному зосереджено на даних, орієнтованих на онлайн-дискурс, які включають принаймні деякі прямі взаємозв’язки між учасниками (лівий верхній квадрант на рисунку 3.1). Незважаючи на те, що обчислювальні інструменти були розроблені для вилучення різних аспектів

інтелекту з колекцій оглядів продуктів і фільмів, мікроблогів і заголовків новин, вони є незалежними творами, створеними окремими особами, і тому виключені з цього дослідження.

Щоб визначити програму досліджень у напрямку розвитку та аналізу колективного інтелекту в Інтернеті, ми переглянули існуючі обчислювальні підходи, які використовували автоматизовані методи для вивчення великомасштабних веб-платформ колективного інтелекту. Ми спеціально зосередилися на колективному інтелекті, який виникає внаслідок вербальних взаємодій. Тому наш огляд виключив платформи для онлайн-оглядів, мікроблогів та заголовків новин, оскільки ці текстові блоки є незалежними творами, створеними окремими особами. Крім того, ми зосереджуємось на потенційно повторюваних взаємодіях у обмежених середовищах, тому дослідження про Інтернет та соціальні мережі загалом, не включені. Інструменти чату та електронної пошти також виключені, оскільки вони в першу чергу призначені для підтримки індивідуальних і невеликих взаємодій і зазвичай не розглядаються як великомасштабні платформи.

У наступному розділі ми описуємо структуру групового дослідження, яка використовується для класифікації існуючої літератури. Потім ми надаємо більш детальні пояснення досліджень у кожній категорії.

3.2. Розробка методології

Хоча існують різні моделі обробки малих груп (наприклад, модель групового процесу Макграта, модель усвідомлення ситуації Ендслі і модель етапів Такманса, ми не можемо знайти модель, яка б специфічна для масштабна групова діяльність. Макграт визначає групи як «соціальні агрегати, які передбачають взаємне усвідомлення та потенційну взаємну взаємодію». Спочатку призначена для вирішення відносно невеликих і структурованих груп, модель Макгратса – це не просто модель чи теорія діяльності малих груп, а

основа для систематичного вивчення груп, враховуючи всі фактори, які потенційно впливають на процес і результат групової діяльності, і те, як вони взаємодія. Ми припускаємо, що ці фактори також існують у більших онлайн-колективах і мають подібні структури взаємодії, як описано в моделі Макгратса. Тому ми обрали цю модель для нашого систематичного огляду та застосували її до дослідження більших онлайн-колективів, які передбачають взаємну взаємодію між людьми.

Спрощена версія фреймворку McGrath показана на рисунку. У цьому контексті процес групової взаємодії розглядається як центральна ознака та сутність групи, і він відноситься до «шаблонних відносин між поведінкою індивідів». На додаток до цього основного компонента, структура охоплює кілька елементів, які якимось чином можуть бути пов’язані з процесами групової взаємодії. Наприклад, учасники впливають і можуть на них впливати процеси взаємодії; таким чином, властивості члена включені в структуру. Зразок стосунків між учасниками є важливим фактором у розумінні взаємодій всередині індивідів, і тому він виглядає як постійна група чи групова структура в рамках. Ще два важливі аспекти включають середовище, де відбувається групова взаємодія, і завдання (тобто неформально передбачувані цілі), які переслідують члени групи.

На додаток до загальної рамки групових досліджень, Макграт описує процеси взаємодії на мікрорівні, розглядаючи їх у термінах трьох різних етапів. Перший етап розглядається зі структурою і розподілом комунікацій між взаємодіючими людьми, який називається комунікаційним шаблоном. На другому етапі кожне спілкування розглядається з огляду на його зміст. Сам зміст потім розглядається з точки зору двох різних аспектів: міжособистісних стосунків і виконання завдань. Нарешті, третій етап пов'язаний з впливом різних процесів групової взаємодії один на одного і на учасників.

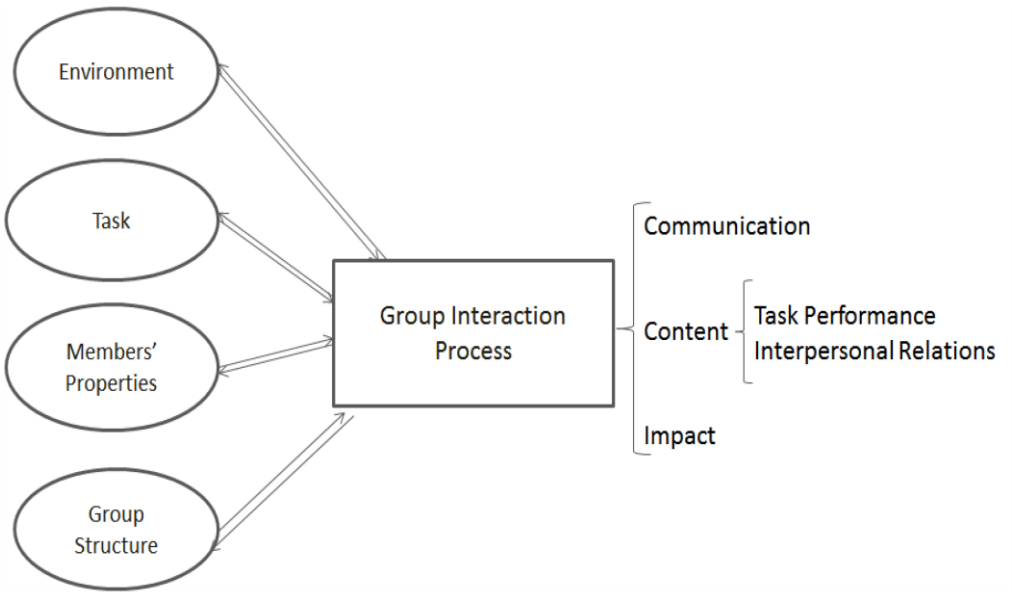


Рисунок 3.2 - Структура Макграта може використовуватися для класифікації літератури про онлайн-колективи.

Сучасний напрям дослідження аналізу колективного інтелекту зазвичай розглядає властивості членів (наприклад, стать), структуру групи (наприклад, ієрархічну), середовище (наприклад, дискусійний форум) і завдання (наприклад, прийняття рішень) як контекст і незалежні змінні дослідження, а потім полегшує різні обчислювальні методи для вивчення компонента процесу групової взаємодії. Тому пов'язані роботи з аналізу колективного інтелекту поділяються на три основні групи: аналіз моделі комунікації, аналіз контенту та аналіз впливу. Контент-аналіз потім поділяється на дві категорії досліджень. Через різну природу віртуальних просторів компонент міжособистісних стосунків розглядається на вищому рівні, зосереджено на дослідженнях, які аналізують емоційні та сентиментальні аспекти змісту. Крім того, компонент виконання завдання аналізу вмісту називається виконанням завдання/цілі, оскільки такі параметри можуть бути менш орієнтованими на завдання.

Оскільки цей огляд зосереджений на обчислювальних методах та алгоритмах для аналізу колективного інтелекту, ми почали процес відбору статті, провівши кілька пошуків за ключовими словами в цифровій бібліотеці ACM.

Запити були сформовані шляхом запозичення ключових слів із системи групових досліджень McGrath, таких як «аналіз моделі спілкування» та «аналіз настроїв». Початковий аналіз 100 найкращих документів, отриманих за кожним запитом виконується шляхом читання заголовків, рефератів та вступів, щоб відфільтрувати набір, щоб включати лише ті статті, які входять до сфери цього дослідження. Таким чином, щоб визначити відповідні статті, ми використовували наступні критерії для кожного дослідження: 1) запропоновані автоматизовані підходи з мінімальним кодуванням і зусиллями людини та 2) провели аналіз на великомасштабних веб-платформах, які передбачають обмін інформацією та діалог серед кількох учасників.

Після того, як початковий набір документів був обраний, для охоплення відповідних статей в інших місцях використовувалися методи зворотного та прямого ланцюга. Метою було не охоплення кожного окремого дослідження, а надання репрезентативної вибірки статей, щоб отримати достатнє розуміння тем дослідження та прогалин у сучасному стані літератури.

Шаблон спілкування відноситься до форми або структури серії інтерактивних форм поведінки або спілкування, які мають місце серед людей. Недавнє дослідження Woodley показав, що моделі повороту учасників співвідносяться з колективним інтелектом групи. Цей висновок свідчить про важливість аналізу комунікаційних моделей для розуміння та розвитку інтелекту між людськими колективами.

Незважаючи на те, що попередні роботи в таких контекстах спиралися на основні статистичні показники, більш складні обчислювальні методи аналізу комунікаційних моделей дотримувалися двох основних напрямків: аналіз соціальних мереж та аналіз ниток розмов як деревоподібних структур. Крім того, замість того, щоб зосередитися на моделях групового рівня, ряд досліджень досліджував моделі спілкування як поведінку індивіда.

Велика кількість досліджень намагається створити соціальну мережу учасників, а потім використовувати різні методи аналізу графіків, запропоновані

аналітиками соціальних мереж, щоб краще зрозуміти основні комунікаційні моделі. У соціальних мережах кожному учаснику відповідає вузол i ∈ V на графіку G = hV, Ei, де ребро (i, j) ∈ E на графіку вказує на якийсь соціальний зв'язок між двома учасниками i та j. При вивченні моделей спілкування краї зазвичай представляють повідомлення та онлайн-діалоги, якими обмінюються учасники. Різні типи соціальних мереж можуть бути сформовані на основі різних інтерпретацій того, як повинні бути пов'язані два учасники, наприклад, чи мають бути спрямовані краї, на основі відправника та одержувача повідомлення, чи краї мають бути зважені на основі кількості повідомлень. обмінялися.

Наприклад, Gomez et al. [2] досліджували три різні типи соціальних мереж, отримані з відповідей на коментарі на веб-сайті обміну новинами Slashdot, які відповідають різним стратегіям визначення межі між парою вузлів. Далі вони проаналізували різні атрибути цих мереж, такі як максимальна кількість кластерів, середня довжина шляху, максимальна відстань між двома учасниками, взаємність, розподіл ступенів та асортативне змішування за ступенем і балом. Їхні висновки означають, що, незважаючи на те, що Slashdot поділяє деякі особливості моделі спілкування, такі як невелика світова власність і висока кластеризація, з традиційними соціальними мережами, її учасники продемонстрували нейтральне змішування за ступенем, ідентичне розподілення та просто помірну взаємність, що спостерігається рідко. в інших соціальних

мережах.

В іншому дослідженні аналізуються моделі взаємодії на сторінках обговорення Вікіпедії [3], де виділяються три різні мережі відповідно до різних типів взаємодій: мережа відповідей на статті, яка включає прямі відповіді на сторінках обговорення статей; мережа розмов користувачів, яка включає прямі відповіді на сторінках обговорення користувачів; та настінна мережа, яка містить особисті повідомлення, розміщені на сторінках обговорень іншого користувача. Ці три мережі порівнюються за допомогою метрики перекриття країв. Результати показують відносно велике перекриття між двома мережами, витягнутими зі

сторінок обговорення, в той час як мережа з особистих комунікацій було виявлено, що істотно відрізняється від двох інших. Крім того, аналіз спрямованих асортативно за ступенем показав, що користувачі, які відповідають багатьом іншим користувачам, як правило, відповідають недосвідченим користувачам, тоді як ті, хто отримує коментарі від багатьох користувачів, частіше взаємодіють один з одним.

Замість того, щоб формувати соціальні мережі з розмов, деякі дослідники досліджували комунікаційні моделі як інформаційні каскади або потоки, де дія відбувається внаслідок впливу діяльності інших. У цих дослідженнях каскади моделюються як дерево T = hV, Ei, де кожен вузол i ∈ V представляє дію учасника, а між i і j є ребро (i, j) ∈ E, якщо одне з цих дії є відповіддю на іншого або під впливом іншого.

У дослідженні Gomez обговореннях Slashdot потоки обговорень відображаються на радіальні дерева, де повідомлення є центральним вузлом, а його коментарі додаються на різних рівнях вкладення залежно від їх рівня до центрального повідомлення. Аналізуючи коефіцієнт розгалуження b, тобто кількість коментарів, викликаних даним коментарем, вони виявили високий рівень неоднорідності в дискусійних ланцюгах. Крім того, результати аналізу вказують на різницю в процесі генерації коментарів початкового рівня та коментарів наступних рівнів. Крім того, метрика h-індексу, що використовується

в науковій спільноті, прийнята для визначення рівня дискусії.

Проаналізовано структуру ланцюжкових розмов у трьох різних наборах даних, а саме групах Usenet, Yahoo! Групи та Twitter. Досліджуються взаємозв’язки між розміром і глибиною, розподілом ступенів і рівнем потоку, а також розміром і кількістю авторів у потоках. На основі базового процесу розгалуження також розробляється математична модель. У процесі розгалуження кожен вузол генерує k дітей з ймовірністю p(k), де p — функція розподілу ймовірностей. Беручи до уваги початковий аналіз, автори включили в свою модель ступені вузлів, їх сучасність та ідентичність автора.

У своїй пізнішій роботі Gomez проаналізували чотири різні набори соціальних даних і показали, що модель преференційної прихильності з ухилом до початкового повідомлення здатна охопити більшість структурних властивостей, а також моделі еволюції каскадів розмов. У моделі переваги приєднання ймовірність того, що існуючий вузол буде пов'язаний з новим вузлом, пропорційна його розподілу ступенів, що означає, що вузли з вищими ступенями, швидше за все, будуть пов'язані з новими. Для забезпечення більш повної моделі використовується інформація про останній час разом із упередженням щодо початкового повідомлення та факторів розподілу ступеня.

Ланіадо вивчали форму та розмір дерев обговорення на рівні статті у Вікіпедії та запропонували різні показники для характеристики сторінок обговорення, такі як ланцюжки прямих відповідей, максимальна глибина та h-індекс дерева. Після класифікації статей Вікіпедії та проведення їх аналізу на різні теми вони виявили значну різницю в структурах обговорення з різних семантичних областей.

Хоча вищезгадані дослідження були зосереджені на моделях спілкування на груповому рівні, були зроблені деякі спроби охарактеризувати комунікативну поведінку індивідів у широкомасштабних соціальних умовах. Наприклад, проведено велику кількість досліджень, присвячених моделюванню та аналізу часових інтервалів комунікативної діяльності індивіда. Крім того, его-мережі використовувалися для розуміння комунікаційних моделей індивідів, таких як ідентифікація «осіб, які обговорюють» та «відповідаючих» або ідентифікація лідерів. Окрім мереж его, Адамік запровадив показник ентропії, який фіксує концентрацію шаблонів відповідей учасника на різні теми на форумі питань і відповідей.

Була проведена велика робота з аналізу настроїв, дослідники досліджували різні методи для визначення суб'єктивних речень, оцінки рівня настроїв за двійковою (наприклад, позитивний/негативний) або порядковою шкалою (наприклад, від 1 до 5), щоб виявляти настрої та емоції (наприклад, гнівні чи

радісні) і, нарешті, зрозуміти джерело, ціль та складні типи ставлення. Позначаючи ту саму область дослідження, використовується багато інших термінів, серед яких найбільш поширеними є аналіз суб'єктивності та вивчення думки.

Для автоматичного аналізу настроїв було використано два основні методи. Методи, засновані на лексиці, спиралися на анотовані словники, які включають полярність настроїв та оцінки, щоб визначити настрої текстової одиниці на основі слів і фраз, які вона містить. Замість використання анотованих словників, підходи до навчання з учителем створюють класифікатори, використовуючи мічені дані, щоб передбачити настрої речень, фраз або документів. Різні форми та комбінації цих методів використовувалися та вивчалися в загальних текстах та онлайн-оглядах, тоді як онлайн-дискусіям на соціальних платформах приділялося відносно менше уваги.

У попередніх роботах, присвячених методам на основі лексики, використовувалися як словники, створені людьми-анотаторами, так і автоматично створені, створені за допомогою методів навчання без нагляду. Традиційний підхід до використання таких словників полягає у визначенні оцінки настрою текстової одиниці шляхом усереднення значень настрою окремих слів, витягнутих зі словника. Ваннер наприклад, використовував словник, що містить настрої, для оцінки оцінок настроїв.

Через важливість контексту в аналізі настроїв, пізніші підходи вийшли за

межі заперечень і врахували структуру дискурсу та інші контекстуальні особливості в процесі обчислення настроїв. Наприклад, Лі і Ву використовували китайський словник із настроями, що включає списки позитивних і негативних слів, приватів і сентиментально зважених модифікаторів, щоб призначити рівень настроїв для постів для обговорення. Для кожного ключового слова в даній публікації рейтинг настроїв спочатку оцінюється на основі словника. Потім алгоритм перевіряє, чи присутній привілей у словах перед ключовим словом. Якщо це так, оцінка сентименту слова заперечується; інакше алгоритм

переходить до останнього кроку. На цьому кроці, якщо в словах, що оточують ключове слово, присутній модифікатор, рейтинг настрою ключового слова множиться на вагу модифікатора, витягнутий зі словника. Результати аналізу настроїв потім включають у вектори ознак для кластеризації k-середніх, за допомогою яких можна виявити гарячі точки. Крім того, класифікатор Support Vector Machine (SVM) навчається на векторах ознак настроїв і результатах k-середніх для прогнозування наступного гарячого форуму.

Зростання доступності позначених екземплярів в Інтернеті разом із появою платформ краудсорінгу (наприклад, Amazon Mechanical Turk), де мічені дані можна ефективно та ефективно збирати, призвело до використання різноманітних контрольованих методів класифікації в дослідженні настроїв. Набори функцій, які зазвичай використовуються в техніках вивчення настроїв, включають присутність і частоту термінів, позицію терміна, частину мови (POS), синтаксис і заперечення. Наприклад, Розенталь і МакКевон адаптували існуючий метод, який використовує лексичні оцінки слів, аналіз N-грам і полярність оточуючих синтаксичних складових, до даних соціальних мереж, беручи до уваги емоції, акроніми та орфографічні помилки.

Хоча методи навчання під наглядом були доведені як корисні та точні в завданнях класифікації настроїв, коли вони навчаються в іншій області, ніж область інтересу, їх ефективність значно падає. Незалежний від домену алгоритм, який називається аргументацією через пошук, використовується для класифікації настрою (тобто сердитий, сумний, радісний) коментарів новин на Yahoo!Buzz на основі позначених екземплярів, отриманих з LiveJournal. набір даних коментарів.

Щоб автоматично виявляти речення зі ставленням до інших учасників онлайн-дискусій, Hassan застосував дещо інший підхід, ніж основні дослідження аналізу настроїв. У їхньому підході мовна модель будується з використанням позначених даних. Однак замість того, щоб присвоювати мітки текстовим одиницям, мовна модель присвоює значення ймовірності. Найрелевантнішою

частиною речень зі ставленням є те, де включається займенник другої особи. Таким чином, цей фрагмент спочатку витягується. Потім за допомогою лексичних елементів, тегів POS, тегів полярності слів і відносин залежностей витягується набір шаблонів, пов'язаних з кожним фрагментом. Шаблони потім використовуються для побудови моделей Маркової для речень з ставленням і без нього. Тоді нове речення можна позначити як речення з ставленням або без нього.

Цінність груп полягає в їхній потенційній здатності ефективно виконувати широкий спектр завдань. Тому було проведено широке класичне дослідження соціальної психології за допомогою польових досліджень, опитувань та лабораторних експериментів із застосуванням ручних методів для дослідження того, як і в яких умовах різні групи можуть бути ефективними. Проте останніми роками поширення соціальних платформ на основі Інтернету вимагало автоматизованих підходів до обробки та оцінки великомасштабного текстового вмісту, створеного натовпом. Низка досліджень, присвячених онлайн-соціальним середовищам, намагається вирішити проблему перевантаження інформацією шляхом організації та опису особливостей основного інформаційного простору. Такі дослідження можна розглядати як попередні кроки, які можуть привести до легшого та більш точного аналізу виконання завдань машинами або користувачами. Інший набір робіт, з іншого боку, розробив нові обчислювальні методи для безпосередньої оцінки якості та релевантності розмовного тексту в Інтернеті.

Одне з ранніх досліджень з опису онлайн-діалогів було проведено Саком, в якому основні методи обробки природної мови використовуються для вилучення та представлення списку найбільш часто використовуваних термінів у дискусіях разом із їх семантичною мережею. Методи кластеризації також були використані для опису змістових особливостей комунікаційного простору. Кластеризація відноситься до групування об'єктів на основі деяких показників подібності. Загалом, методи кластеризації тексту кодують вміст документа у

відповідних структурах даних, таких як векторна модель простору, а потім використовують різні метрики відстані для обчислення міри подібності та формування кластерів документів. Основний принцип, що лежить в основі методів кластеризації на основі вмісту, полягає в тому, що чим більше ключових слів містить документи, тим більше вони схожі.

Кластеризація документів – добре вивчена тема; однак відсутність структури та редакційного контролю в розмовному тексті ставить нові виклики та складнощі. Таким чином, прийняття та розширення методів, що використовуються для вирішення загальних текстів, часто пропонуються для групування розмовного тексту. Наприклад, Саїд і Ваннас запропонували метод кластеризації повідомлень на основі лідера (LPC), модифікацію алгоритму Leader, для кластерування повідомлень у ланцюжках обговорень. Їх підхід починається з головного стовпа нитки як першого ведучого. Використовуючи запропоновану ними метрику відстані, яка включає функції як вмісту, так і моделі комунікації (тобто тегування між дописами), розраховується схожість публікації з наявними кластерами. Якщо результуюча міра нижча за певний поріг, посада відноситься до відповідного кластеру; інакше він стає новим лідером. Цей процес продовжується до тих пір, поки всі дописи не будуть об’єднані в групу, що призведе до виявлення публікацій, які не відповідають темі, і дописів, що виходять за межі.

Близько пов’язаним підходом до кластеризації документів є тематичне моделювання, коли дослідники використовували різноманітні методи видобутку під наглядом і без нагляду для сегментації, позначення та відстеження різних тем. Незалежно від проблем моделювання теми для розмовного тексту, деякі вчені застосовували нові методи в невеликих дискусіях; однак, подібно до кластеризації, тематичне моделювання менш досліджується для масштабних веб-розмов. Один приклад наведено Zhue et al., який розширює основний алгоритм виявлення та відстеження тем (TDT), щоб відповідати цьому методу для даних, згенерованих в онлайн-спільнотах.

Замість кластеризації та тематичного моделювання, інший напрямок роботи з описового лінгвістичного аналізу онлайн-дискусій зосереджено на завданні пошуку та маркування висловлювань та одиниць тексту, які є цінними в аналізі ефективності, наприклад, ідентифікації претензій, угоди чи розбіжності [1], обґрунтування чи аргументи. Підходи, що використовуються в цьому напрямі дослідження, мають певну спільну мову з дослідженнями аналізу настроїв, оскільки обидва стосуються зіставлення фрагмента тексту з міткою з попередньо визначеного списку. Крім того, деякі підходи в цьому напрямку використовують суб'єктивний аналіз для визначення певних ознак (наприклад, претензій та угод/розбіжностей). Таким чином, обчислення настроїв іноді включається в підходи.

У роботі, представленій Розенталем і Маккевоном [8], претензії автоматично виявляються за допомогою регресійного аналізу за різними ознаками, включаючи настрої, слова віри, POS та N-грами. Вони виявили, що настрої та POS-теги були важливішими в LiveJournal, тоді як віра та настрої були важливішими для форумів Вікіпедії. Ця відмінність свідчить про наявність відмінних мовних характеристик у різних соціальних контекстах.

Їхній підхід витягує список показників відношень теорії риторичної структури (RST), які вони вважали відповідними для виконання завдання. Для кожного показника список пар слів, що зустрічаються разом, витягується з англійської Вікіпедії та становить основу функцій, які вони використовували у своєму підході до навчання під керівництвом.

Розроблено два інструменти визначення сенсу для систем управління ідеями. Перший інструмент, який називається ідеєю, використовує теорію мовленнєвих актів, щоб автоматично позначати ядро ідеї в неструктурованій текстовій інформації. Другий інструмент, званий інтерпретатором коментарів, шукає конкретні мовні характеристики в коментарях, створених для ідеї, а потім призначає коментар до однієї з трьох попередньо визначених категорій коментарів: реакції, пов'язані зі змістом, значенням або статусом ідеї.

Використовуючи ці мітки, система проактивно пропонує реакції на ці коментарі на основі категорії, до якої він належить.

Інший напрям дослідження аналізу виконання завдання/цілі використовує обчислювальні методи для безпосередньої оцінки контенту, створеного користувачами. Наприклад, щоб підрахувати релевантність коментарів до фактичного повідомлення, [5] використовує TF-IDF, у якому коментарі розглядаються як запити, а дописи розглядаються як документи. Оскільки вони застосовують свою техніку на сайті коментування новин, корпус новин Reuters використовується як базова колекція для обчислення показника IDF. Аналогічно, [4] представляє новий метод виявлення повідомлень не по темі на форумах. У цій техніці набір термінів витягується за допомогою методу інформативності слів BLTR для представлення кожного потоку. Кожна публікація також представлена набором термінів. Міра подібності цих двох векторів поєднується з трьома іншими мірами, які обчислюють подібність публікації з головною, з попередніми повідомленнями та з усіма попередніми повідомленнями, щоб обчислити релевантність кожного
повідомлення для ланцюжка.

Запозичуючи алгоритм HITS з інформаційно-пошукових досліджень запропоновано метод на основі графіків для виявлення найважливіших дописів у ланцюжковому обговоренні. Щоб сформувати графік із розмовного тексту, кожне повідомлення представлено вузлом на графіку. Запропоновано метод генерування посилань на основі ознак для розміщення зв’язків між парою вузлів на основі трьох ознак лексичної подібності, достовірності плакату та відносин мовленнєвого акту. Потім для кожного вузла ітераційно обчислюються два бали концентратора (тобто якість вузла як вказівника на інші) і авторитет (тобто якість вузла як ресурсу), доки алгоритм не збіжиться. Weimer et al. [6] застосував класифікатор SVM, заснований на різних внутрішніх характеристиках, таких як лексична, синтаксична та схожість вмісту, щоб оцінити якість публікацій (тобто хороших чи поганих) на дискусійних форумах.

Аналіз впливу має на меті зрозуміти потенційний вплив процесів групової взаємодії, тобто моделей спілкування, процесу виконання та сентиментальних аспектів, один на одного. Крім того, дуже цікавим є розуміння того, як ці процеси впливають на поведінку учасників, як-от їхнє ставлення, сприйняття, судження та навчання. Використовуючи обчислювальні методи, було дотримано дві різні процедури для вивчення впливу в онлайн-середовищі. Низка робіт була зосереджена на застосуванні існуючих методів для незалежного обчислення змінних, а потім використання основних статистичних методів для оцінки їхньої кореляції. Кілька останніх підходів використовували більш складні методи для безпосередньої оцінки потенційного впливу однієї змінної на іншу.

Аналізуючи сайт коментування новин, дослідили зв’язок між актуальністю, тимчасовістю, настроєм і якістю коментарів новин. Під час аналізу актуальності були визначені теми, які викликали негатив і, таким чином, потребували більшої модерації. Була виявлена кореляція між негативністю та кількістю видалених коментарів, що свідчить про те, що настрої можуть бути показником якості коментарів. Крім того, у своєму часовому аналізі вони виявили позитивну кореляцію між частотою коментарів та негативністю коментарів користувачів, що також свідчить про потенційний зв’язок між настроями та якістю коментарів. У відповідному дослідженні на форумах BBC [10] досліджується зв’язок між активністю учасників та їхніми настроями [10], виявляючи, що негатив підвищує активність учасників, а більш активні користувачі схильні висловлювати негативні емоції.

Chmiel та ін. провели відносно складний аналіз онлайн-повідомлень на різних платформах, щоб зрозуміти вплив поточних емоцій та настроїв, висловлених в онлайн-спільноті, на емоції наступних публікацій. У їхньому підході ланцюжки дописів спочатку об’єднуються в групи послідовних дописів з однаковим значенням настроїв (тобто позитивні, нейтральні чи негативні). Порівняння сформованих кластерів з кластерами, сформованими на основі випадкових даних, показало значну різницю, причому кластери фактичних даних

мають більший розмір. Їх статистичний аналіз показує, що умовна ймовірність для послідовних постів зростає як степеневий закон із розміром кластера, що схоже на преференційні процеси. Загалом вони дійшли висновку, що публікації в Інтернеті, як правило, ініціюють публікації подібних

3.3. Представлення результатів тестування

Крім того, зв’язок між розміром потоку та емоціями, вираженими в ланцюжку, показавши, що коротші нитки, як правило, починаються з меншого негативного настрою, а довші нитки мають більші
варіації настроїв.

Щоб виявити прогалини в дослідженні з наявної літератури, ми розглянули основне та низькорівневе призначення робіт, використані обчислювальні методи, середовище їхнього фокусування та вторинні переваги, які вони можуть надати технологіям Web 2.0. Результати підсумовані в таблиці 3.1.

Як показано в таблиці, незважаючи на цінність і можливості, які полягають у середовищі, орієнтованому на завдання, більшість попередніх робіт були зосереджені на середовищах, в першу чергу, призначених для полегшення соціальних взаємодій (наприклад, форуми для дискусій та набори коментарів у новинах/медіа). Було докладено небагато зусиль, щоб зрозуміти орієнтовані на завдання онлайн-середовища, такі як інструменти обговорення та системи управління ідеями. Можливо, було багато досліджень щодо зручності використання та користувацького досвіду (UX), які досліджували вплив середовища за допомогою традиційних методів UX. Новим і недостатньо вивченим напрямком у літературі є розробка обчислювальних методів для вивчення цих середовищ шляхом виявлення та аналізу «слідів» колективного інтелекту. Очікується, що на «сліди» впливають різні фактори, такі як характеристики учасників, види розвідувальних завдань та дизайн середовища.

Відсутність складних автоматизованих методів аналізу впливу також видно з таблиці. У той час як велика група робіт була зосереджена на застосуванні існуючих методів для незалежного обчислення змінних і використанні основних статистичних методів для оцінки їх кореляції; лише кілька останніх підходів використовували більш складні методи для безпосередньої оцінки потенційного впливу одного процесу взаємодії на інший або на учасників. Методи навчання під наглядом і без нагляду можуть бути цінними та ефективними інструментами для моделювання та потенційного прогнозування таких впливів.

Крім того, як видно з таблиці, були зроблені деякі спроби проаналізувати вплив процесів взаємодії, тобто моделі спілкування, виконання завдання/мети та настроїв, один на одного. Однак менше уваги було приділено розумінню того, як ці процеси впливають на індивідуальну та колективну поведінку учасників з часом, наприклад на їхнє сприйняття, навчання та судження.

ВИСНОВКИ

В дипломній роботі прелставлені техніка та методологія інтелектуалізації web. Протягом останніх кількох десятиліть Інтернет перетворився з платформи для простого споживання вмісту до місця, де люди можуть створювати, ділитися, змінювати й обговорювати вміст різних форм. Цей прогрес призвів до явища соціальної мережі, яка сьогодні значно впливає на наше повсякденне життя. Безпрецедентний рівень інтелекту(знань), отриманий внаслідок таких масштабних онлайн-взаємодій, є одним із помітних ефектів соціальної мережі. З метою зробити внесок у галузі веб-колективної та бізнес-аналітики, ми вперше визначили два соціальних феномена конфіденційності та переконання, які мають неймовірний вплив на парадигму видобування даних.

Для подальшого вивчення концепції соціальної конфіденційності ми спочатку провели огляд літератури. Потім ми вивчили зв'язок різних типів даних соціальних медіа з перевагами конфіденційності користувачів і запропонували модель для визначення профілів соціальних мереж, для яких діє дихотомія конфіденційності. Потім ми зосередилися на феномені онлайн-міркування, проглянувши відповідні дослідження. Потім були проаналізовані й обговорені механізми та стратегії онлайн-переконання. Враховуючи значення риторичних відносин у вивченні переконання в тексті, було запропоновано новий підхід для виявлення та уникнення неоднозначності явних риторичних зв'язків із текстом. Нарешті, було запропоновано інтерактивний інструмент візуалізації, щоб відобразити значення моделей інтелекту та структур для подальшого розвитку інтелекту.

СПИСОК ПОСИЛАНЬ НА ДЖЕРЕЛА

[1] A. Hanneman and M. Riddle, “Introduction to social network methods,” online at http://www.faculty.ucr.edu/ hanneman/nettext/, 2005.

[2] S. Wasserman and K. Faust, Social Network Analysis: Methods and Applications. Cambridge University Press, Novemver 1994.

[3] R. L. Breiger, The Analysis of Social Networks. London: Sage Publications Ltd, 2004, pages 505-526 in Handbook of Data Analysis, edited by Melissa Hardy and Alan Bryman.

[4] T. Berners-Lee, “Semantic web road map,” online at http://www.w3.org/DesignIssues/Semantic.html, 1998.

[5] A. Swartz and J. Hendler, “The semantic web: A network of content for the digital city,” in Proceedings of Second Annual Digital Cities Workshop, Kyoto, Japan, October 2001.

[6] J. Hendler, T. Berners-Lee, and E. Miller, “Integrating applications on the semantic web,” Journal of the Institute of Electrical Engineers of Japan, vol. 122, no. 10, pp. 676–680, 2002.

[7] L. Freeman, Graphic techniques for exploring social network data. New York: Cambridge University Press, 2005, pages 248-269 in ’Models and methods in social network analysis’ edited by P.J. Carrington, J. Scott, and S. Wasserman.

[8] O. Frank and D. Strauss, “Markov graphs,” Journal of the American Statistical Association, vol. 81, pp. 832–842, 1986.

[9] S. Wasserman and P. Pattison, “Logit models and logistic regression for social networks: I. an introduction to markov graphs and p ∗,” Psychometrika, vol. 61, pp. 401–425, 1996.

[10] C. Anderson, S. Wasserman, and B. Crouch, “A p∗ primer: logit models for social networks,” Social Networks, vol. 21, pp. 37–66, 1999.

[11] T. A. Snijders, P. E. Pattison, G. L. Robins, and M. S. Handcock, “New specifications for exponential random graph models,” 2004.

[12] G. Robins, P. Pattison, and S. Wasserman, “Logit models and logistic regressions for social networks,” Psychometrika, vol. 64, pp. 371–394, November 1999.

[13] P. Hoff, A. Raftery, and M. Handcock, “Latent space approaches to social network analysis,” Journal of the American Statistical Association, vol. 97, pp. 1090–1098, 2002.

[14] C. Butts, “Network inference, error, and informant (in)accuracy: a bayesian approach,” Social Networks, vol. 25, no. 2, pp. 103–140, 2003.

[15] T. Snijders, “Markov chain monte carlo estimation of exponential random graph models,” Journal of Social Structure, 2002.

[16] P. Hoff, “Random effects models for network data,” Irvine, California , November 2003.

[17] Y. Wang and G. Wong, “Stochastic blockmodels for directed graphs,” vol. 82, no. 8-19, 1987.

[18] M. Handcock, “Assessing degeneracy in statistical models of social networks,” December 2003, working paper, University of Washington.

[19] P. Smyth, “Statistical modeling of graph and network data,” in Proceedings of IJCAI Workshop on Learning Statistical Models from Relational Data, Acapulco, Mexico, August 2003.

[20] D. S. Johnson and M. A. Trick, Cliques, Coloring, and Satisfiability: Second Dimacs Implementation Challenge. American Mathematical Society, October 1996.

[21] “Connectivity and generalized cliques in sociometric group structure,” Psychometrika, vol. 15, pp. 169–190, 1950.

[22] M. Everett, “Graph theoretic blockings, k−plexes and k−cutpoints,” Journal of Mathematical Sociology, vol. 9, pp. 75–84, 1982.

[23] P. Mika, “Flink: Semantic web technology for the extraction and analysis of social networks,” Journal of Web Semantics, vol. 3, pp. 211–223, October 2005.

[24] L. Kahney, “Making friendsters in high places,” July 2003, the Wired News.

[25] S. Brin and L. Page, “The anatomy of a large-scale hypertextual web search engine,” in Proceedings of the 7th World-Wide Web Conference (WWW7), BrisBane, Australia, April 1998.

[26] J. Kleinberg, “Authoritative sources in a hyperlinked environment,” in Proceedings of Ninth Annual ACM-SIAM Symposium on Discrete Algorithms, San Francisco, California, January 1998.

[27] J. Dean and M. R. Henzinger, “Finding related pages in the world wide web,” in In Proceedings of the 8th World Wide Web Conference, Toronto, Canada, May 1999.

[28] K. Bharat, A. Broder, M. Henzinger, P. Kumar, and S. Venkatasubramanian, “The connectivity server: Fast access to linkage information on the web,” Brisbane, Australia, April 1998.

[29] P. Reddy and M. Kitsuregawa, “Inferring web communities through relaxed cocitation and dense bipartite graphs,” in Proceedings of the Second International Conference on Web Information Systems Engineering (WISE2001), Kyoto, Japan, December 2001.

[30] R. A. S. Chakrabarti, B. Dom and P. Raghavan, “Using taxonomy, discriminants, and signatures to navigate in text databases,” in Proceedings of 23rd International Conference on Very Large Data Bases, Athens, Greece, August 1997.

[31] J. Allan, “Automatic hypertext link typing,” in Proceedings of 7th ACM Conference on Hypertext, Hypertext ’96, Washington DC, March 1996, pp. 42–51.

[32] R. Agrawal, J. Gehrke, D. Gunopulos, and P. Raghavan, “Automatic subspace clustering of high dimensional data for data mining applications,” in In SIGMOD Conference on Management of Data, Seattle, USA, June 1998.

[33] S. Chakrabarti, B. Dom, S. R. Kumar, P. Raghavan, S. Rajagopalan, A. Tomkins, D. Gibson, and J. Kleinberg, “Mining the web’s link structure,” IEEE Computer, vol. 32, no. 8, pp. 60–67, August 1999.

БІБЛІОГРАФІЧНА ДОВІДКА

Тема дипломної роботи бакалавра: " Адаптивні алгоритмічні стратегії інтелектуалізації Web "

Обсяг пояснювальної записки: 109 аркушів

Дата закінчення дипломної роботи 13 червня 2021р.

Підпис студента _____