

ОРГАНІЗАЦІЯ СЕМАНТИЧНОГО АНАЛІЗУ КОНТЕНТУ СОЦІАЛЬНИХ МЕРЕЖ

Актуальність обраної теми пов'язана з постійно зростаючою кількістю задач обробки природної мови людини у сучасному цифровому світі, що характеризується стрімким розвитком цифрових технологій та необхідністю пошуку певної інформації і її аналізу за певний, невеликий, проміжок часу. Сучасний стан комп'ютерної техніки та програмного забезпечення дозволяє проводити подібний пошук за час від кількох хвилин до кількох днів, тижнів, а то й місяців, що є незручним, а часто й недопустимим з огляду на потребу у подібних задачах. Серед факторів, що впливають на тривалість згаданих задач пошуку інформації, є як фактори, пов'язані з потужністю сучасних комп'ютерних систем, так і фактори, пов'язані з задіяним програмним забезпеченням, на які впливає, зокрема, спосіб організації даних у оперативній пам'яті комп'ютерних систем та алгоритми обробки таких даних.

1. Поняття «соціальна мережа» та групи соціальних мереж

Соціальна мережа – це структура, яка створена людьми чи компаніями, організаціями, і у якій відображені різні зв'язки між учасниками соціальної мережі. Саме поняття «соціальна мережа» вперше було представлено Дж. А. Барнесом у 1953 році.

Соціальні мережі можна розділити на наступні групи [1]:

- для спілкування (Relationship networks);
- для обміну медіа-контентом (Media sharing networks);
- для відгуків і оглядів (Online reviews);
- для колективних обговорень (Discussion forums);
- для авторських записів (Social publishing platforms);
- сервіси соціальних закладок (Bookmarking sites);
- за інтересами (Interest-based networks);

2. База даних та види БД

База даних – це така сукупність даних, яка організована відповідно до певних правил і має певну структуру. Вона редагується за допомогою системи управління базами даних (СУБД).

Види:

Ієрархічна модель даних – логічна модель даних у вигляді дерева, яка представляє собою серію елементів, розташованих в порядку їх підпорядкування від загальнодоступного до приватного, утворюючи перевернуте дерево (граф). Ця модель характеризується такими параметрами, як рівні, вузли, зв'язку. Принцип роботи моделі полягає в тому, що кілька вузлів нижчого рівня зв'язані за допомогою зв'язку з вузлом вищого рівня. Інформаційна модель вузла елемента на заданому рівні ієрархії [4].

Реляційна модель даних була запропонована співробітником IBM Едгаром Коддом і базується на концепції відносин. Реляція (від англ. **Relationship** - відношення) - це математичний термін у теорії множин, який ліг в основу реляційної моделі даних. У реляційних базах даних вся інформація узагальнена в таблицях, рядки яких називаються записами, а стовпці - полями. Реляційна модель даних для даного домену складається з ряду взаємозв'язків, які змінюються з часом [3].

Мережевий підхід до організації даних є розширенням ієрархічного підходу. В ієрархічних структурах запис нащадка повинен мати рівно одного предка. У мережевій структурі даних нащадок може мати будь-яку кількість предків. Окрім вертикальних посилань, горизонтальні посилання також реалізовані в мережевих базах даних. Однак багато недоліків ієрархії успадковуються, головним недоліком є чітке визначення зв'язків даних на фізичному рівні та дотримання цієї структури зв'язків з однаковою чіткістю під час пошуку в базі даних [3].

3. Контент-аналіз

Контент-аналіз (від англ. *contents* зміст) - метод якісно-кількісного аналізу змісту документів з метою виявлення або вимірювання різних фактів і тенденцій, відображених у цих документах. Особливість контент-аналізу полягає в тому, що він вивчає документи в їх соціальному контексті. Може використовуватися як основний метод дослідження (наприклад, контент-аналіз тексту при дослідженні політичної спрямованості газети), паралельний, тобто у поєднанні з іншими методами (напр., у дослідженні ефективності функціонування засобів масової

інформації), допоміжний або контрольний (напр., при класифікації відповідей на відкриті питання анкет) [2].

Важливим аспектом у користуванні соціальними мережами в Інтернеті є потреба у фільтрації матеріалу, яку, у першу чергу, можна проводити за допомогою семантичного аналізу контенту соціальних мереж для полегшення задачі пошуку та обробки інформації. Для задачі фільтрації надзвичайно важливими стають алгоритми обробки інформації та збереження результатів обробки у базах даних та знань. Цілком зрозуміло, що швидкодія алгоритмів обробки інформації значною мірою залежить від ефективної організації даних у пам'яті інформаційних систем та від оптимального розміщення даних у інформаційних базах та оптимізації цих баз.

Список використаних джерел

1. Види социальных сетей: классификация и представители [Електронний ресурс] – Режим доступу до ресурсу: <https://darksiteofmarketing.com/stati/vidy-socialnyh-setei-klassifikacija-i-predstaviteli.html>.
2. Манаев О. Т. Контент-анализ — описание метода [Електронний ресурс] / О. Т. Манаев – Режим доступу до ресурсу: [https://psyfactor.org/lib/kontent.htm#:~:text=%D0%9A%D0%BE%D0%BD%D1%82%D0%B5%D0%BD%D1%82%2D%D0%B0%D0%BD%D0%B0%D0%BB%D0%B8%D0%B7%20\(%D0%BE%D1%82%20%D0%B0%D0%BD%D0%B3%D0%BB.,%D1%82%D0%B5%D0%BD%D0%B4%D0%B5%D0%BD%D1%86%D0%B8%D0%B9%2C%20%D0%BE%D1%82%D1%80%D0%B0%D0%B6%D0%B5%D0%BD%D0%BD%D1%8B%D1%85%20%D0%B2%20%D1%8D%D1%82%D0%B8%D1%85%20%D0%B4%D0%BE%D0%BA%D1%83%D0%BC%D0%B5%D0%BD%D1%82%D0%B0%D1%85..](https://psyfactor.org/lib/kontent.htm#:~:text=%D0%9A%D0%BE%D0%BD%D1%82%D0%B5%D0%BD%D1%82%2D%D0%B0%D0%BD%D0%B0%D0%BB%D0%B8%D0%B7%20(%D0%BE%D1%82%20%D0%B0%D0%BD%D0%B3%D0%BB.,%D1%82%D0%B5%D0%BD%D0%B4%D0%B5%D0%BD%D1%86%D0%B8%D0%B9%2C%20%D0%BE%D1%82%D1%80%D0%B0%D0%B6%D0%B5%D0%BD%D0%BD%D1%8B%D1%85%20%D0%B2%20%D1%8D%D1%82%D0%B8%D1%85%20%D0%B4%D0%BE%D0%BA%D1%83%D0%BC%D0%B5%D0%BD%D1%82%D0%B0%D1%85..)
3. Модели организации баз данных [Електронний ресурс] – Режим доступу до ресурсу: <https://vshot.ru/it-ebook/b20.html>.
4. Модели организации данных [Електронний ресурс]. – 2016. – Режим доступу до ресурсу: <https://studepedia.org/index.php?vol=1&post=75225>.