

## РЕФЕРАТ

Магістерська дисертація: 100 с., 15 рис., 34 табл., 1 додаток, 31 джерел.

**Актуальність.** Сьогодні Інтернет є основним джерелом отримання або поширення інформації. Більшість людей використовує Інтернет для того, щоб проводити час у соціальних мережах. Кожна людина може спробувати себе у ролі репортера, і поділитись важливими новинами, або поширити свою думку серед певної аудиторії і знайти однодумців/послідовників.

На початок 2018 року його аудиторія становила 336 мільйонів людей [1]. Користувачі спілкуються короткими повідомленнями (до 280 символів) які називаються твітами. Це дуже зручно, так як можна отримувати дуже великий об'єм інформації, витрачаючи на це не дуже велику кількість часу.

Саме тому є можливою розробка системи, яка зможе групувати користувачів схожих за інтересами беручи за основу текстову складову їх поведінки у соціальній мережі Twitter.

**Зв'язок роботи з науковими програмами, планами, темами.** Робота виконувалась на кафедрі автоматизованих систем обробки інформації та управління Національного технічного університету України «Київський політехнічний інститут ім. Ігоря Сікорського» в рамках теми «Методи та технології високопродуктивних обчислень та обробки надвеликих масивів даних». Державний реєстраційний номер 0117U000924.

**Мета дослідження** – підвищення якості поділу користувачів соціальної мережі на групи за рахунок аналізу текстової складової їх поведінки.

Для досягнення мети необхідно виконати наступні **завдання**:

- проаналізувати дані, які можна отримати про користувача в соціальній мережі Twitter;
- обрати критерії за якими порівнюються користувачі;
- проаналізувати методи та засоби кластеризації великих даних;

- дослідити способи аналізу поведінки користувача за текстовою складовою;
- реалізувати програмно аналітичний комплекс поділу користувачів соціальної мережі на групи за інтересами.

**Об’єкт дослідження** – процес поділу користувачів на групи, в залежності від схожості їх інтересів.

**Предмет дослідження** – методи кластеризацій користувачів соціальної мережі на основі їх текстової складової поведінки.

**Наукова новизна отриманих результатів** полягає в аналізі існуючих методів кластеризації. Також було проаналізовано способи аналізу текстової складової, так як одним із критеріїв порівняння схожості було взято тематика блогінгу.

**Публікації.** Булгар М.М. Кластеризація користувачів за їх інтересами / М.М. Булгар // МОДС. 2018. С. 28-29.

Булгар М.М. Спосіб кластеризації користувачів соціальної мережі Twitter / М.М. Булгар // ІСТУ. 2018. С. 28-32.

КЛАСТЕРИЗАЦІЯ, СОЦІАЛЬНІ МЕРЕЖІ, ВЕЛИКІ ДАНІ, TWITTER, ТВІТ, АНАЛІЗ