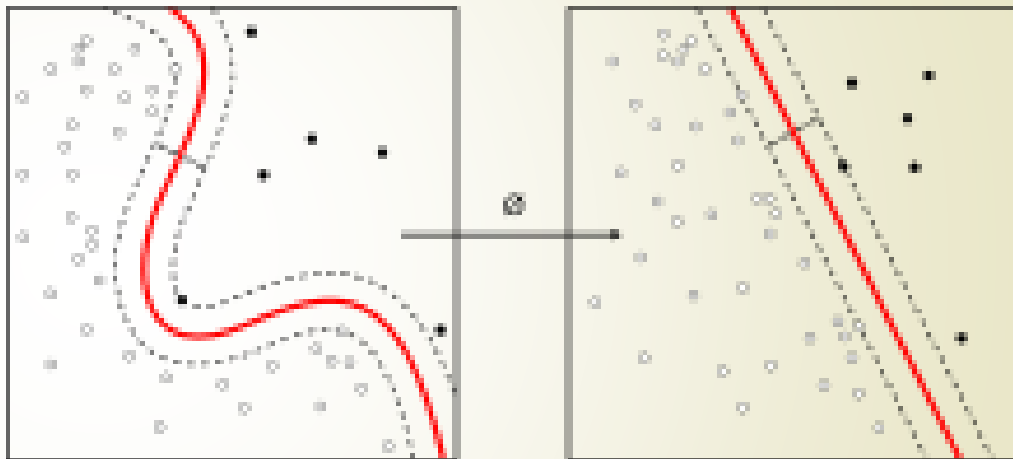


Групування об'єктів на мапі модифікованим методом Ллойда

Виконав: Галька Дмитро, КА-45
Науковий керівник: ас. Кухарев С.О.

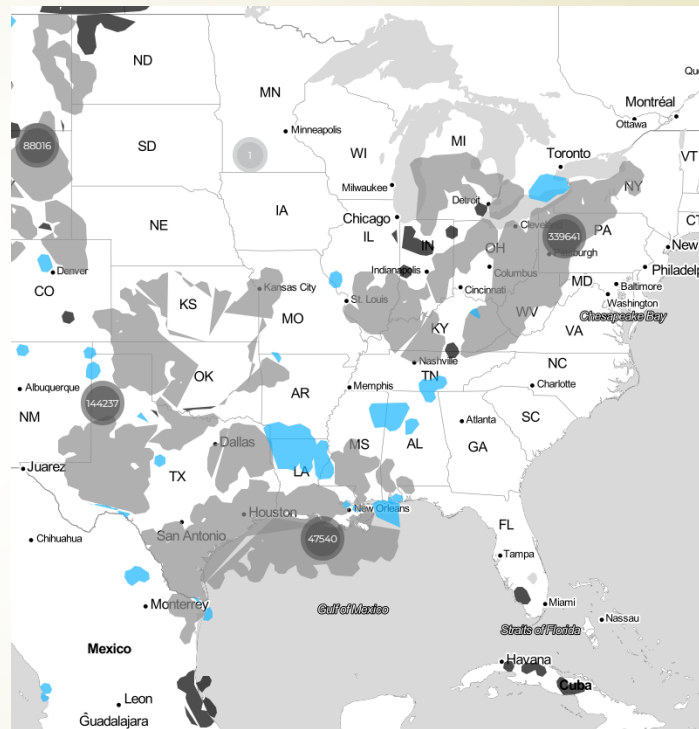
Актуальність

- ← Більшість алгоритмів кластеризації активно використовуються в таких прогресивних напрямках, як розпізнавання об'єктів та машинне навчання



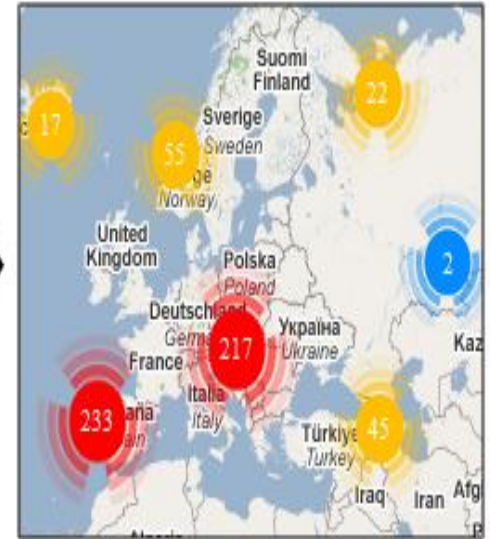
Об'єкт дослідження

← Об'єктом дослідження обрано координати кількості точок на всесвітній мапі.



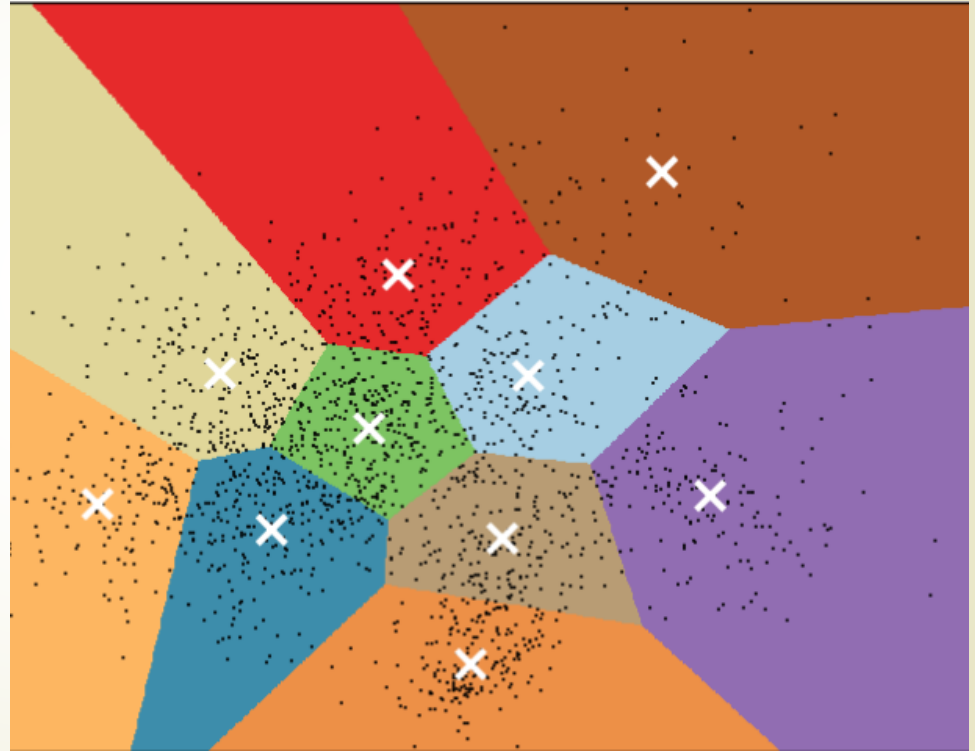
Предмет дослідження

- ← Предметом дослідження є метод групування маркерів на мапі



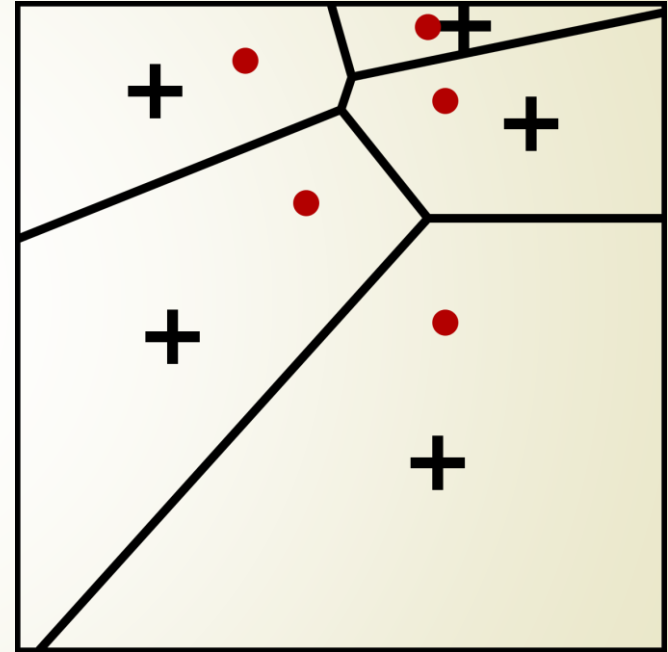
Мета роботи

- ← Побудувати ефективний алгоритм для групування великої кількості об'єктів



Постановка задачі

- ← Ознайомитися з методами кластеризації
- ← Обрати та запрограмувати алгоритм для вирішення задачі групування об'єктів на мапі
- ← Знайти шляхи прискорення обраного алгоритму



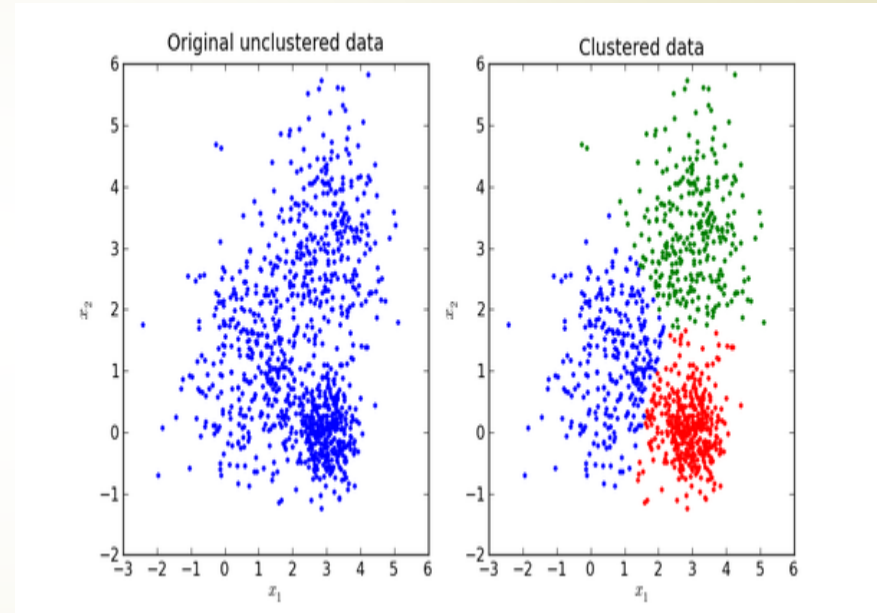
Групування об'єктів на мапі

← Подібне групування дозволяє швидко та зручно орієнтуватися серед великої кількості об'єктів на мапі



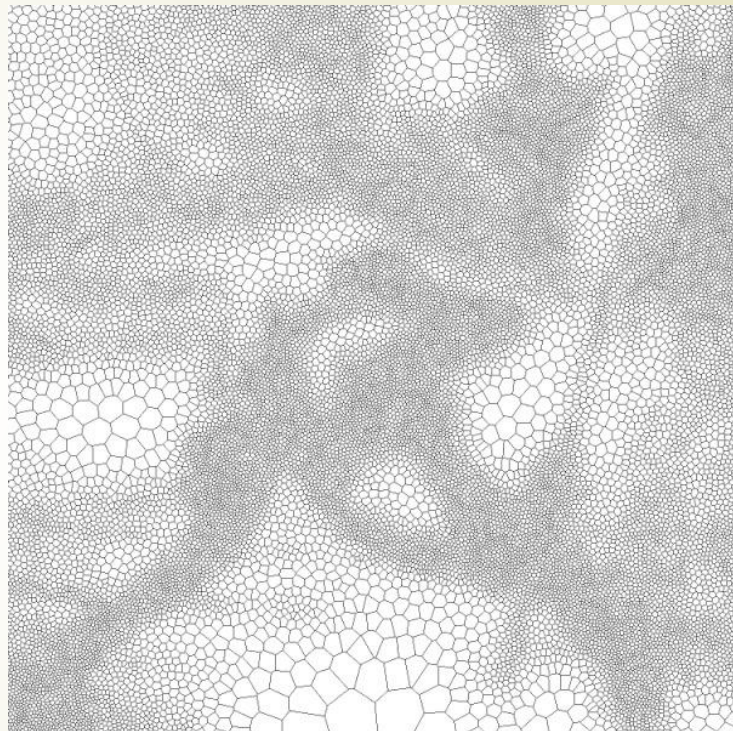
Основні методи кластеризації

- ← Ієрархічні:
 - ← Метод простого зв'язку
 - ← Метод повного зв'язку
 - ← UPGMA
- ← Центроїдні:
 - ← Кластеризація методом к-середніх
 - ← Алгоритм Ллойда
- ← Алгоритми, що базуються на розподіленні
 - ← EM-алгоритм (англ. Expectation-maximization (EM) algorithm)
- ← Алгоритми, що базуються на щільності об'єктів



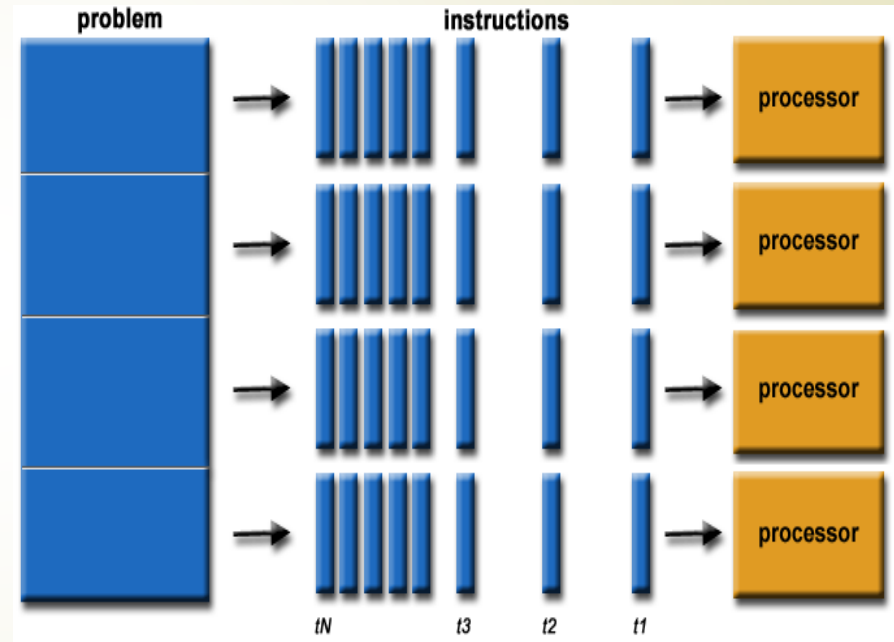
Алгоритм Ллойда

- ✓ Швидкий
- ✓ Простий для реалізації
- ✓ Центроїдний
- ✓ Не потребує заданої кількості кластерів



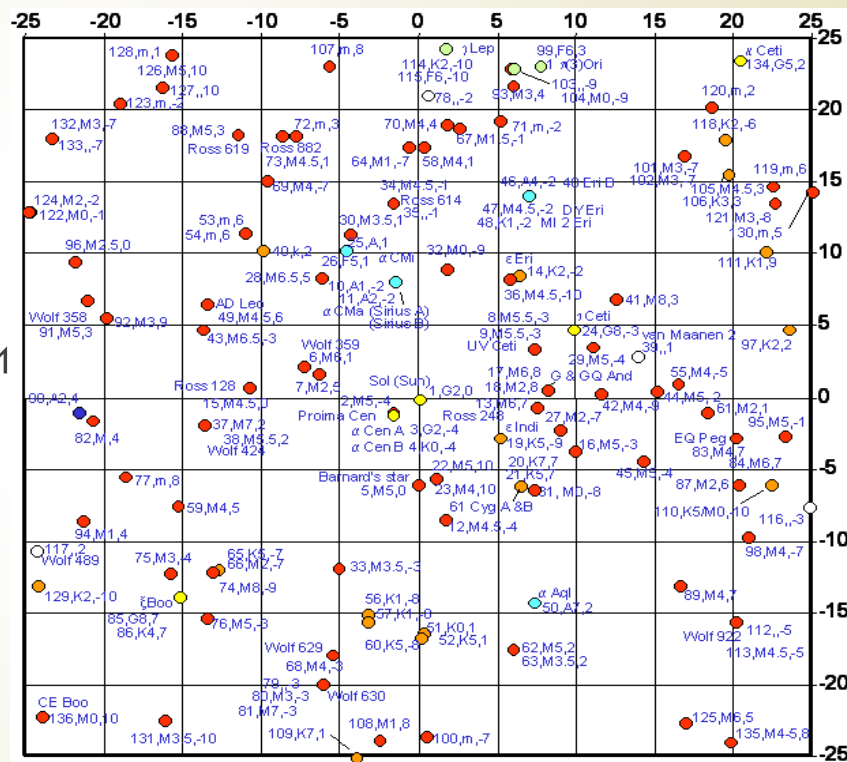
Паралельне обчислення

В випадку групування об'єктів на мапі, найбільше часу витрачається при обчисленні відстані між точками під час пошуку найближчого кластеру. Тож, потрібно розділити цю процедуру пошуку на різні потоки. Це дало приріст продуктивності приблизно в два рази.

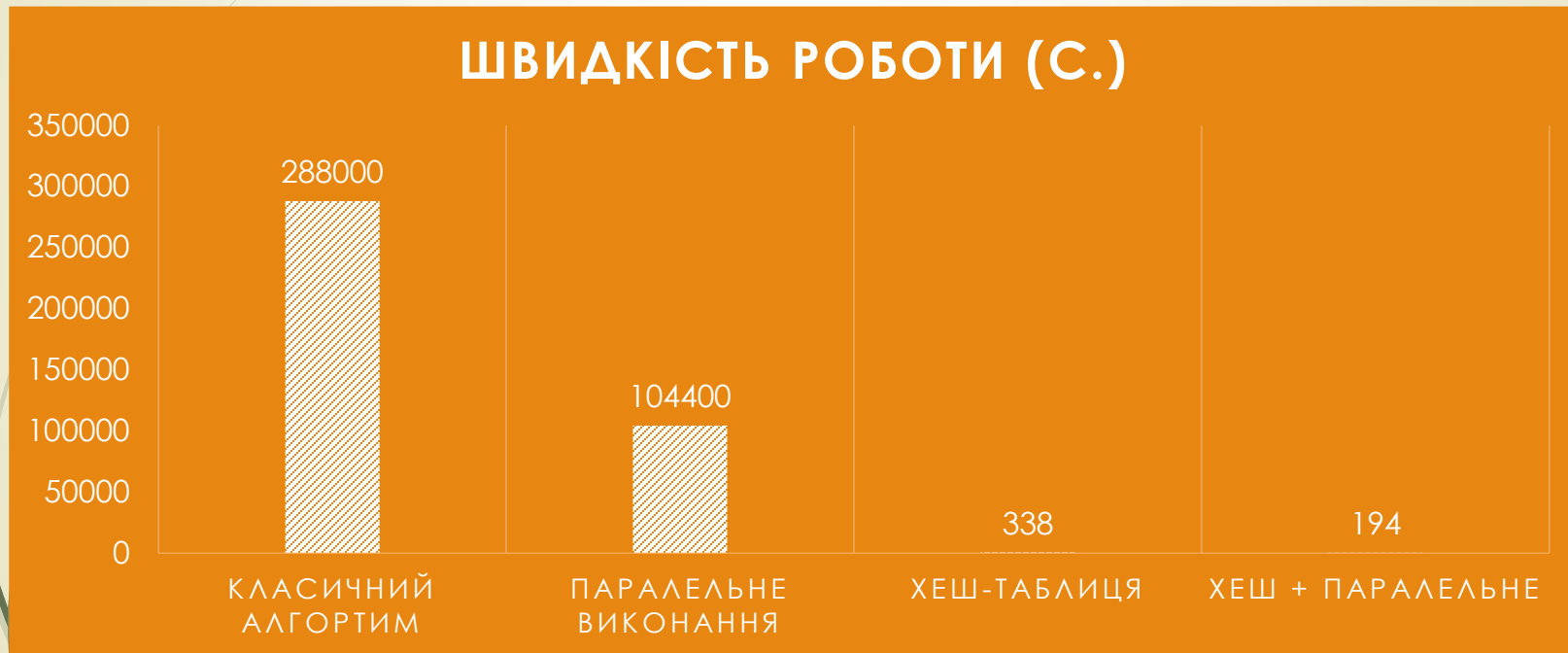


Хеш-таблиці

Розділимо мапу на квадранти розміром $l \times l$, де l – максимальна відстань до кластеру. Таким чином, можна шукати найближчий кластер лише в сусідніх квадрантах, а не по всій мапі. Для цього необхідно створити хеш-таблицю, в якій ключ – це номер квадранта, а значення – список всіх кластерів в цьому квадранті.



Аналіз результатів



Зручність користування програмою



Висновки

- Було проаналізовано існуючі методи кластеризації.
- Було запрограмовано класичний алгоритм Ллойда, який показав погані результати на задачі групування великої кількості об'єктів на мапі.
- Було побудовано оптимізований алгоритм з використанням паралельних обчислень та розподілення точок в хеш таблиці.

Дякую за
увагу

