

## Лабораторная работа №2

### Задача кластеризации (методом k-средних)

Постановка задачи: на вход программе подаются целые числа  $N$  и  $K$ . Далее, на следующих  $N$  строчках, вводятся вещественные координаты точек в двухмерном пространстве. Вам необходимо реализовать программу, которая используя метод k-средних ([https://en.wikipedia.org/wiki/K-means\\_clustering](https://en.wikipedia.org/wiki/K-means_clustering)), распределит заданные точки на  $K$  кластеров. В качестве ответа программа должна выводить  $K$  пар координат центроид каждого кластера, а на последующих  $N$  строчках - номер кластера для каждой точки.

В файлах \*.in представлены данные, на которых необходимо будет провести исследование работы Вашей реализации.

Тестирование и отчёт:

- 1) Реализовать многопоточную версию алгоритма k-means с использованием OpenMP. Решения, в которых присутствуют неоправданные критические секции, race condition, undefined behavior и false sharing будут штрафовать. Проверить корректность ответов программы.
- 2) Посчитать среднее ускорение (отношение времени выполнения параллельной реализации к времени выполнения последовательной реализации) на предоставленных данных для различного кол-ва потоков ( $[1, N]$ , где  $N$  – максимальное кол-во потоков вашей системы), построить график для 4 и 5 набора данных.
- 3) Отчётный материал должен содержать код параллельной реализации задачи, описание реализованного параллельного алгоритма и графики по результатам тестирования. Отчёт можно представлять в текстовом файле, либо в ipynb.

Срок сдачи: до 12.11.2020 до 00:00

4 курс присылает на проверку лабораторную мне (@Cyberogg, evkonova10v@yandex.ru)

3 курс присылает на проверку лабораторную ассистенту Кочергину Владиславу (metrotram102@gmail.com)