

Лабораторна робота №2

Завдання на підготовку до комп'ютерного практикуму «Робота з графовими структурами»

- 1) Використовуючи Spark Graphx / GraphFrame та дані про авіасполучення вирішити, описаний нижче ряд задач.

Опис даних - <https://openflights.org/data.html>

- 2) Використати наступні дані:

- airports-extended.dat,

<https://github.com/fcerbell/sampleddata/tree/master/OpenFlights.org> (в файлі є координати аеропортів)

- airlines.dat

- routes.dat

- 3) Вирішити наступні задачі

- Знайти авіакомпанію з найбільшою сумою відстаней всіх рейсів. Теж саме з найменшою сумою.

- Знайти всі можливі рейси між Польщею та Францією (при не більше ніж 2-х стиковках). Використайте motif та інші варіанти рішень, порівняйте їх результати по швидкодії та складності реалізації.

- Знайти найкоротший та найдовший маршрут між 2 заданими аеропортами. За довжину рейсу беремо відстань між аеропортами. Необхідно врахувати маршрути з пересадками (не більше 3).

- Знайти авіакомпанію з найбільшою сумою відстаней всіх рейсів.

- Виділити великі кластери аеропортів (не менше 5, на основі інформації про рейси). Проаналізувати результати створення цих кластерів.

- Знайти 5 найбільш впливовіших аеропортів методом PageRank.

- Знайти підграфи, з кількістю вузлів не менше 2 зв'язаних між собою, які не мають зв'язків з іншими аеропортами. Використайте методи зв'язних компонентів та Label Propagation.

Вимоги до звіту комп'ютерного практикуму

Звіт до комп'ютерного практикуму виконується у Microsoft Word та повинен містити наступні розділи:

1. Відомості про виконавця.
2. Екранні форми та опис виконаних завдань.
3. Додаток. Лістинг коду.
4. Файл виконаного ноутбука, прикріплений окремо в класрум.

Література

1. Spark GraphX in Action

https://drive.google.com/drive/u/3/folders/1yzqhXB017s8yS0-Bm9oK9C_rVYeJREz9

2. Graph Algorithms. Practical Examples in Apache Spark & Neo4j

<https://drive.google.com/file/d/1SSI-l3E2ZW7MZauaXMxcVhAEN87W8e2W/view?usp=sharing>

ng

3. <https://www.edureka.co/blog/spark-graphx/>

4. Презентації Apache Spark Graphx

<https://drive.google.com/drive/u/3/folders/11jarDXInkZKXaeNOS6WAgkKdPVSMDoD6>