Universite Paris Dauphine

Big Data

Single-source shortest path - Djikstra Algorithm

Students
Elie Abi Hanna Daher
Bilal El Chami
Badr Erraji

Professor
Mr Dario Colazzo

January 30, 2018

Contents

1 Project goal

The goal of the project is to find the shortest paths from a source node to all other nodes in the graph using the Dijkstra's algorithm. The algorithm should be implemented in both Python-Hadoop and Spark.

A long side the implementation, a scalability experiments is needed to check the performance of the algorithm implemented.

2 Dijkstra Algorithm

The Dijkstra's algorithm finds the shortest path from source to all other nodes. Technically the dijkstra algortihm is very similar to the BFS, but instead of having

3 Implementation

3.1 Input

3.1.1 Data

3.1.2 Prepare

3.2 Mapper

3.3 Reduce

3.4 Job Chaining

4 Results

4.1 Hadoop

4.2 Spark

5 Performance

References

[1] *Cloud Computing Lecture 4 - Graph Algorithms with MapReduce.* Jimmy Lin, The iSchool, University of Maryland, February 6, 2008.

A Appendix example

This an example of appendix