Project Name

Large scale graph processing

Group name : Tomodachi

20CS10034 - Sushanth Marada
20CS10059 - Siripuram Bhanu Teja
20CS10028 - Garlapati Sathvik Reddy
20CS10021 - Chukka Jaswanth Kumar
20CS30055 - Talabattula Sai Sahan

Question (Problem number 3)

The goal of the project is process large graphs in a database

i. Install any graph processing systems e.g., ApacheGraph, Pregel (GoldenOrb), Giraph, or Stanford GPS,

ii. Load a large graph from Stanford SNAP large graph repository

iii. Provide interface to run simple graph queries. Bonus for computing PageRank.

iv. Profile performance

The steps :

1. Install a graph processing system: We can install any graph processing system such as ApacheGraph, Pregel (GoldenOrb), Giraph, or Stanford GPS. These systems provide APIs and tools to perform graph processing on large datasets.
2. Load a large graph: We then download a large graph dataset from the Stanford SNAP large graph repository or any other graph repository. The graph dataset can be loaded into the graph processing system.
3. Provide interface to run simple graph queries: In this step we create an interface that allows users to run simple graph queries on the loaded graph. The interface can be a web-based interface or a command-line interface. We intend to use Python(Django) for the backend and Bootstrap, Html, CSS to build the interface.
4. (Bonus part)Computing PageRank: We will add a bonus feature to the project by implementing the PageRank algorithm. PageRank is an algorithm that measures the importance of nodes in a graph. You can use the graph processing system to implement the PageRank algorithm and compute the PageRank of nodes in the loaded graph.At this point we intend to use the page ranking algorithm of google.
5. Profile performance: You can profile the performance of the graph processing system by measuring the execution time and resource usage of the system when processing large graphs. You can use profiling tools such as JProfiler, VisualVM, or Valgrind to measure the performance of the system.