

DB Project

Amartya Sanyal :13089

Sanjana Garg :13617

March 5, 2016

1 Project Idea

A graph visualization system: Most graph visualization tools end up displaying a giant mess of links. We plan to develop an app that interactively allows us to zoom in and zoom out by automatically clustering or sampling nodes and edges so that the user is not overwhelmed with the visualization. We plan to implement some other graph analysis options too and provide scalable techniques for implementing the above idea.

2 Application

Usually visualizing a very large graph is difficult both in terms of the complexity in the huge number of edges as well as storing the entire graph in memory. We plan to use database to store the graphs and then define a wrapper around it in python that can provide various labels of abstraction such that at a time only a few nodes are stored in memory.

3 Access levels

There will be an admin level access and a user level access. An user can view the graph in different levels while the admin can add and remove edges and recalculate the properties.

4 Methodology

We plan to implement hierarchial clustering in the graph using SQL and then store the cluster tree information in an SQL database. While displaying, the required cluster can be selected at the required level of abstraction and displayed using python libraries like networkx. The main work of the project would be in

- Deciding and implementing the correct *relations* to store this data .

- Implementing *clustering* using SQL.
- Implementing other algorithms to display other graph properties. Some examples :-
 1. Centrality
 2. Density
 3. Clustering Coefficient
- *Displaying* the graph using python libraries like networkx in an aesthetic manner and applying a certain level of *interactivity*.