Fundamentals of Database Management System

Term – 3

Wikipedia Article Networks

Dataset information	
Directed	No.
Node features	Yes.
Edge features	No.
Node labels	Yes. Continuous target.
Temporal	No.

Dataset statistics	
	Chameleon
Nodes	2,277
Edges	31,421
Density	0.012
Transitvity	0.314

Submitted to:

Prof. Ashok Harnal

Submitted by:

Group 8

Pranjal Mishra - 025026

Sarthak Jain - 025029

Tanya Goel - 025034

Graph processing:

➢ FuchtermanReingold

Select from Layout dropdown, option Fruchterman Reingold option with parameters as mentioned in the figure below. We set 'Area' to 30000 and 'Speed' to 30 as the number of nodes is very large. Large speed sacrifices precision. Click 'Run'. Wait for at least for three minutes.
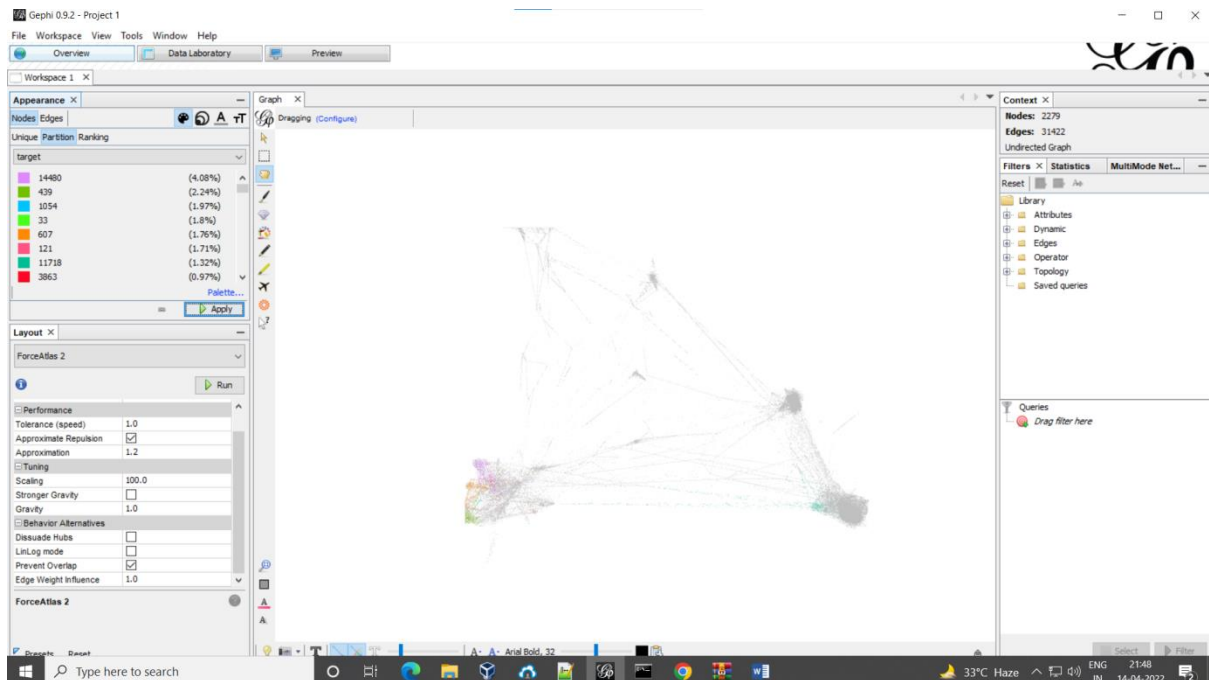
Here is the graph after 3 minutes.

➢ Force Atlas 2

select Layout ForceAtlas2. Set Scaling to 100. Scaling decides how much repulsion one wants. More scaling makes a more sparse graph. Also check Prevent Overlap
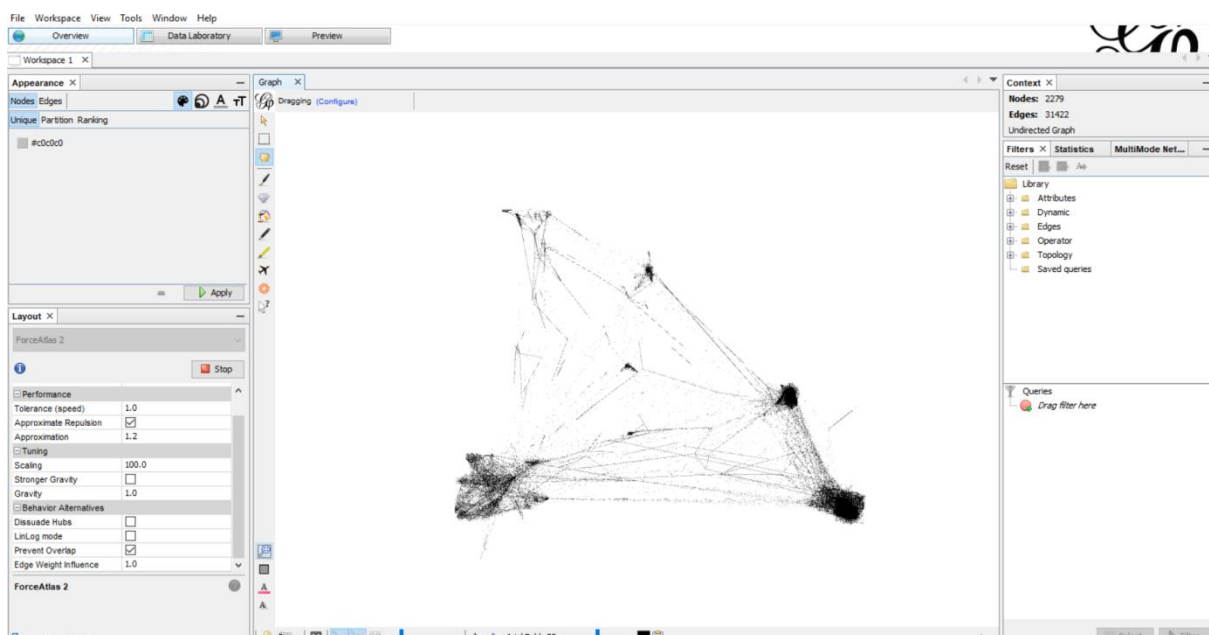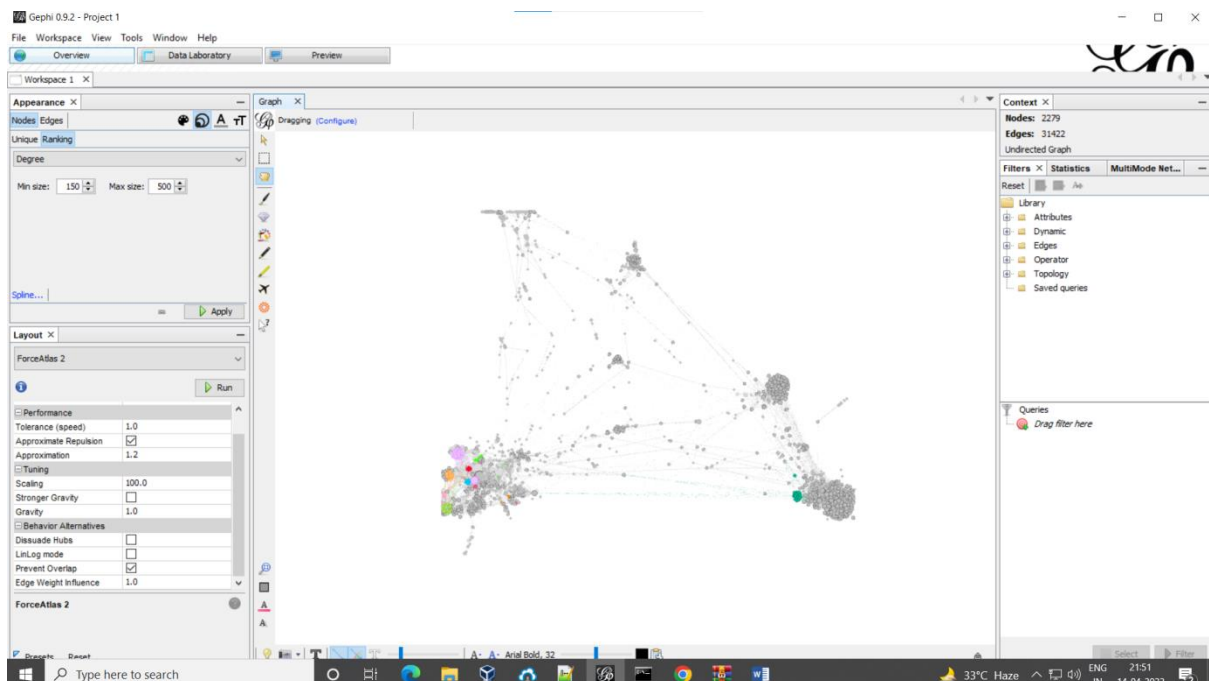
Differentiating nodes:

Differentiating nodes on the basis of target and adding colors to them.


Increasing the size of nodes

# Analysing network with filters and communities:

➢ Modularity Report

The structure of networks is measured by modularity. It assesses the effectiveness of a network's segmentation into modules, groups, or communities. High-modularity networks exhibit extensive connections between nodes inside modules, but sparse connections between nodes in different modules. In networks, modularity is frequently used to detect community structure.

In the Statistics tab, click on Modularity. Repeated clicking on Modularity gave us a different number of communities. At 0.702. we got 15 communities. Keep communities to a minimum.
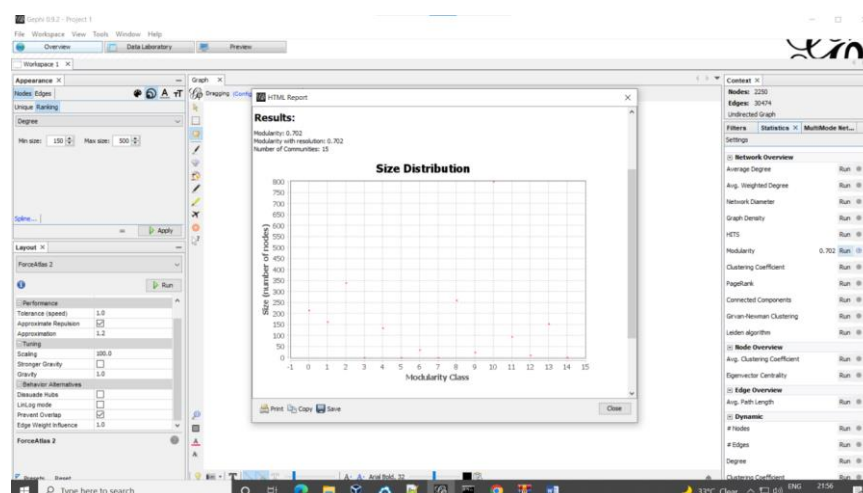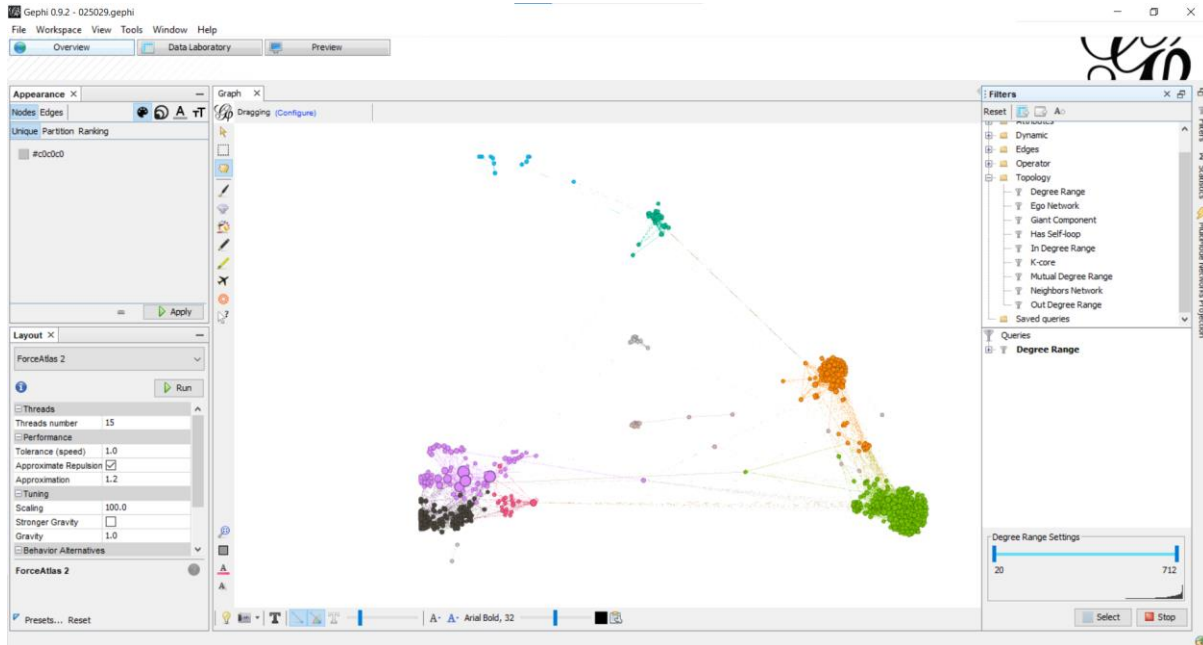

Colour nodes as per modularity class

Coloring nodes as per communities. Mainly 5 communities out of total 15 communities dominate. These 5 communities are coloured as Magenta, Green, Black, Blue and Orange.
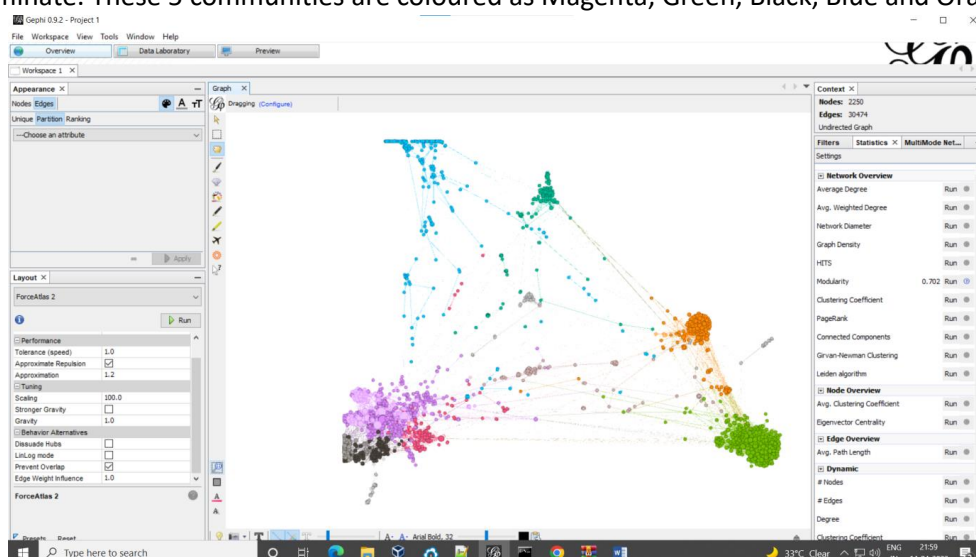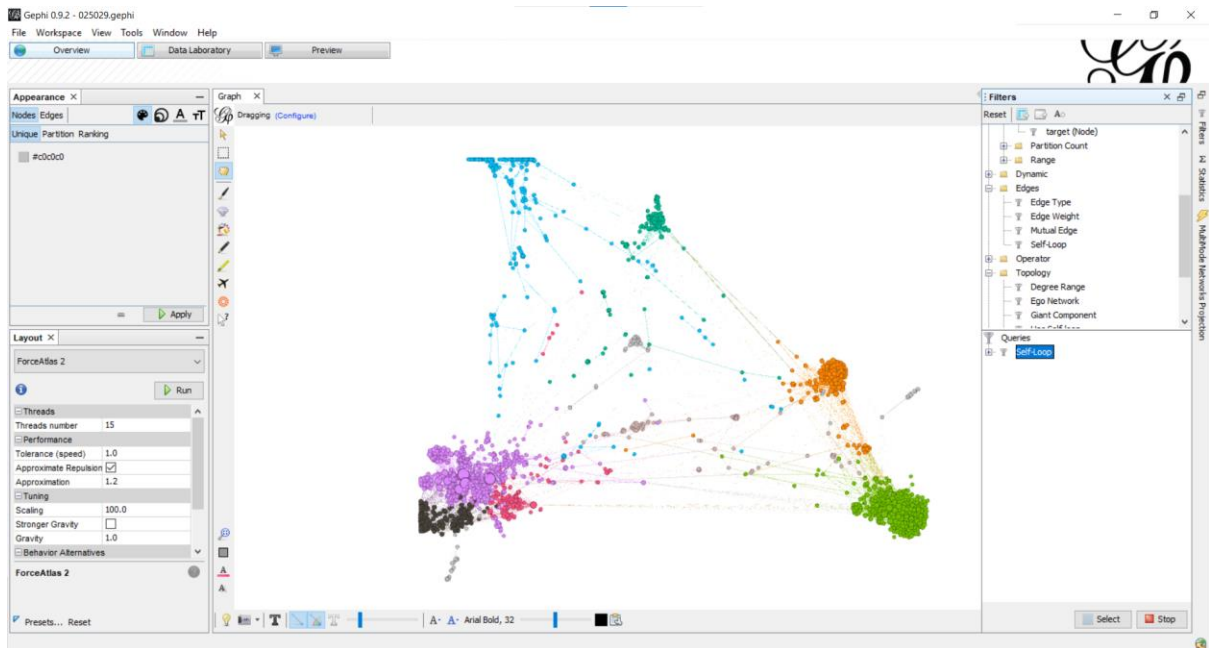
Filtration:

Topology degree range

Partition Modularity class

Self-Loop Filter

Mask Filter