Fundamentals of Database Management System

Social Network Analysis

Submitted to:

Prof. Ashok Harnal

Submitted By:

Alakshendra Sharma- 025004

Trisha Goyal- 025035

Kriti Suri- 025047

C. Elegan’s Neural Network

Introduction:

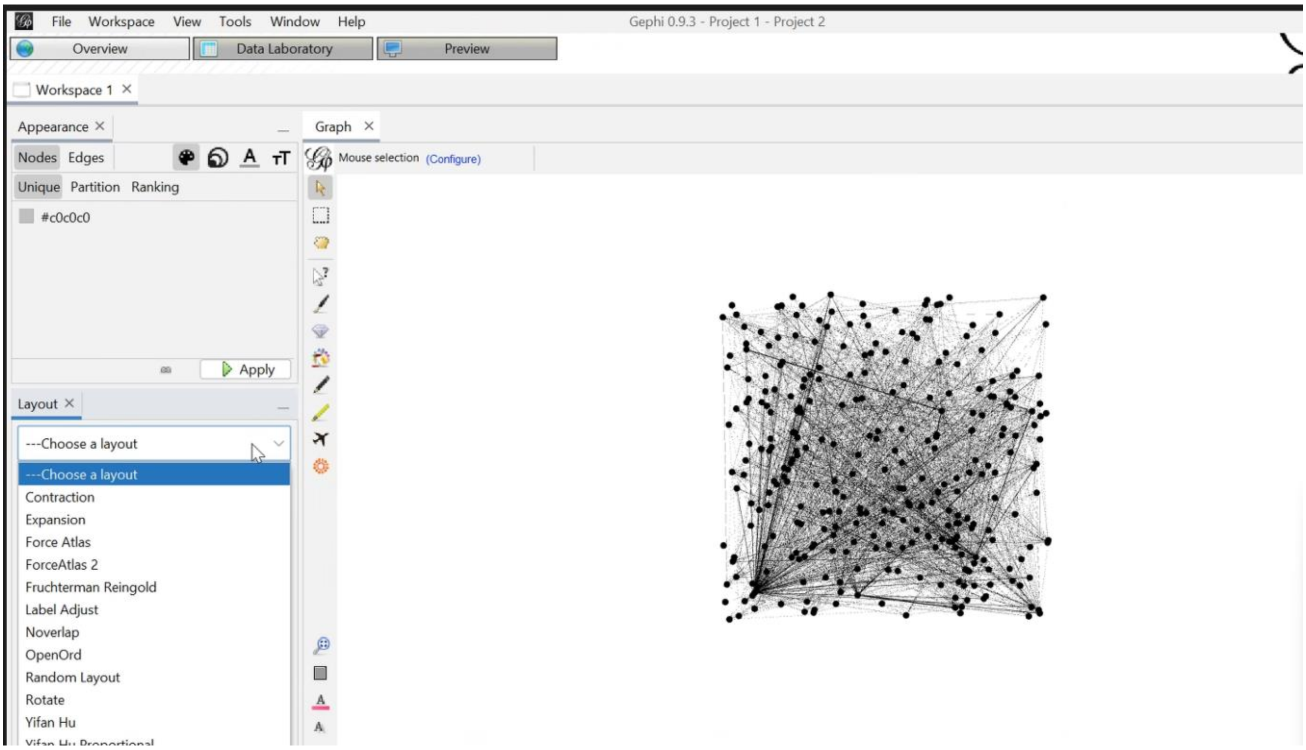
An undirected network of neural network of C. Elegan, Compiled by Duncan Watts and Steven Strogatz from original experimental data by White et al. This is a neural network of neurons and synapses in C. elegans, a type of worm.

Nodes: 297

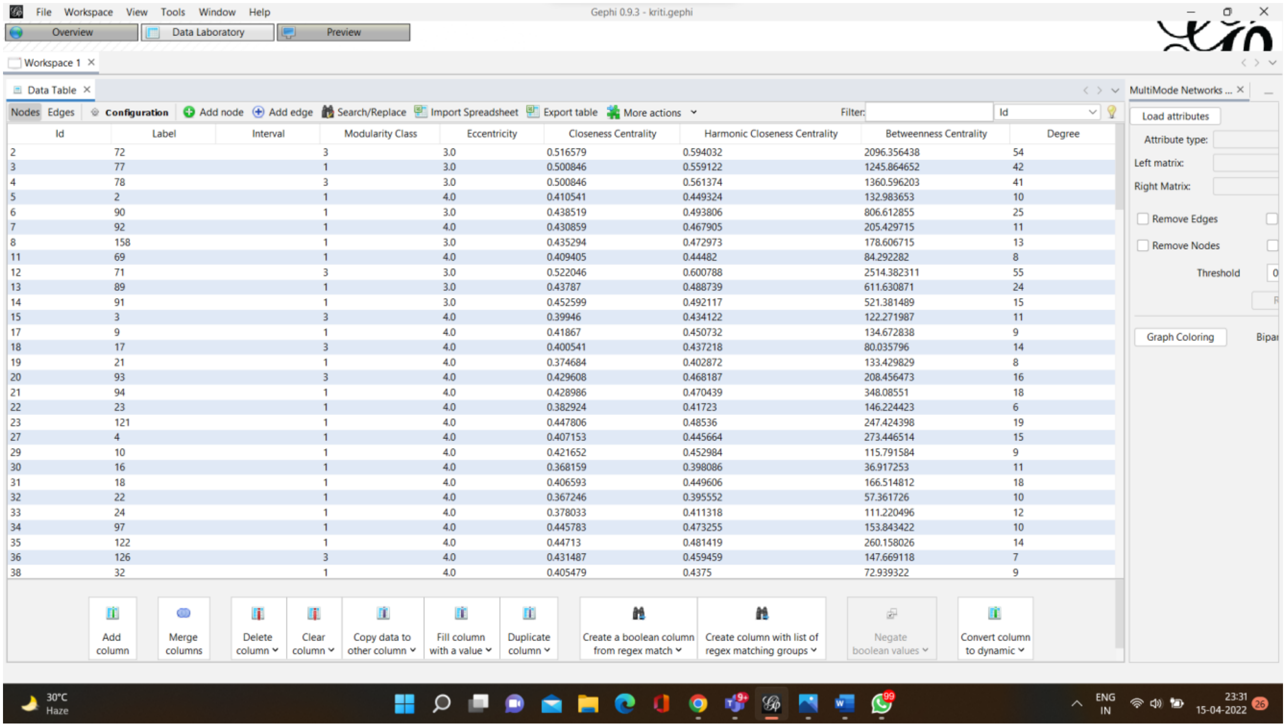
Edges: 2148

Analysis of human neural network has been a great challenge, since humans have ~ 10° neurons and ~1012 synaptic connections. To understand how the brain works, we can start from some simple models. Here we want to first build up an undirected, weighted graph for the neural network of Caenorhabditis elegans worm, it's a relatively simple graph with only ~300 nodes. Then we will use different approaches of network analysis to get more information of this network, for example, the betweenness centrality, the degree distribution, and so on. We show the close connection between structure and function. Finally, we try more complex analysis of the dynamic neural network.

The following procedure has been followed for social media analysis: The files were imported as follows →

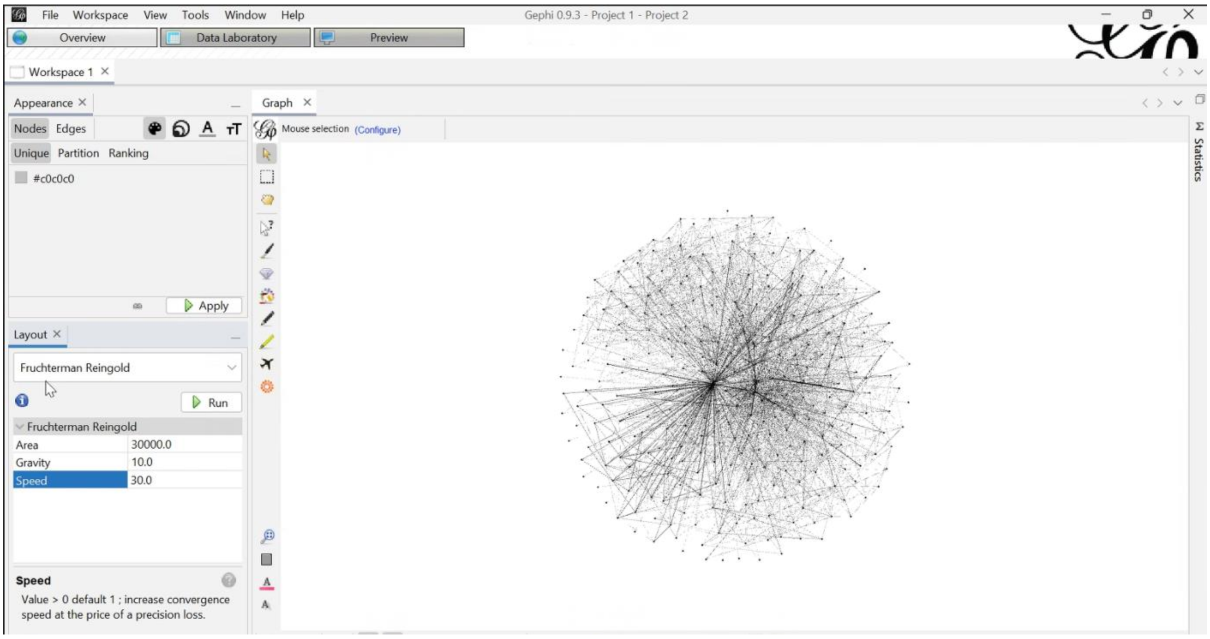


DATA LABORATORY REPORT:



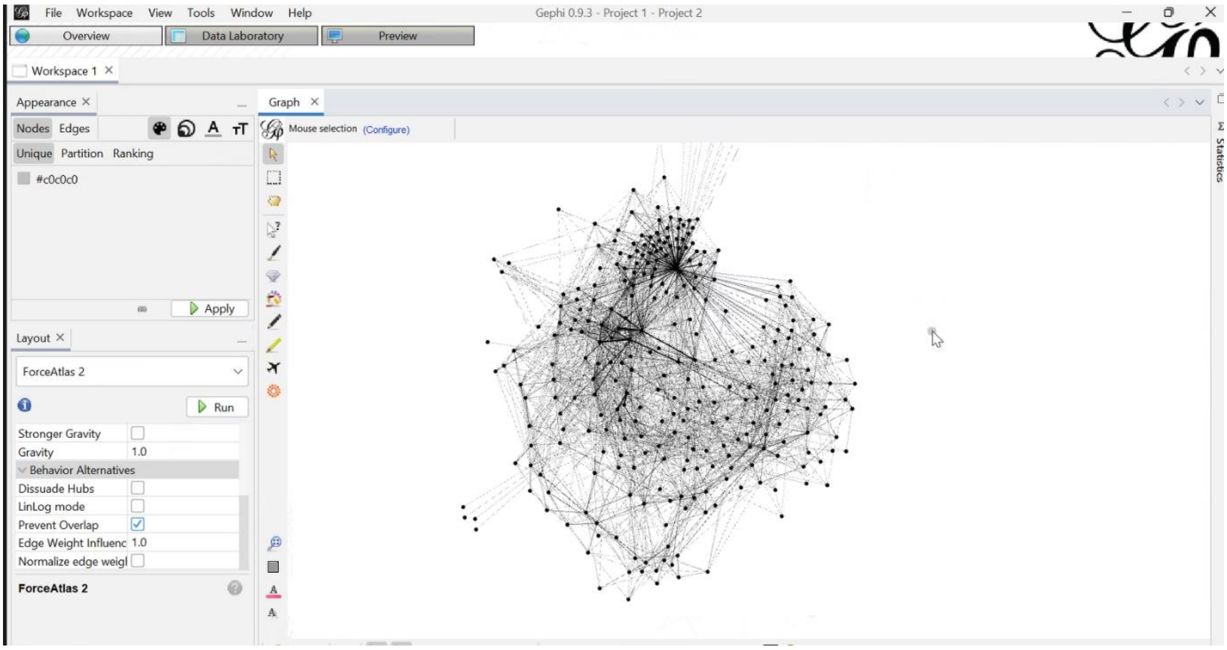
Id	Label	Interval	Modularity Class	Eccentricity	Closeness Centrality	Harmonic Closeness Centrality	Betweenness Centrality	Degree
2	72		3	3.0	0.516579	0.594032	2096.356438	54
3	77		1	3.0	0.500846	0.559122	1245.864652	42
4	78		3	3.0	0.500846	0.561374	1360.596203	41
5	2		1	4.0	0.410541	0.449324	132.983653	10
6	90		1	3.0	0.438519	0.493806	806.612855	25
7	92		1	4.0	0.430859	0.467905	205.429715	11
8	158		1	3.0	0.435294	0.472973	178.606715	13
11	69		1	4.0	0.409405	0.44482	84.292282	8
12	71		3	3.0	0.522046	0.600788	2514.382311	55
13	89		1	3.0	0.43787	0.488739	611.630871	24
14	91		1	3.0	0.452599	0.492117	521.381489	15
15	3		3	4.0	0.39946	0.434122	122.271987	11
17	9		1	4.0	0.41867	0.450732	134.672838	9
18	17		3	4.0	0.400541	0.437218	80.035796	14
19	21		1	4.0	0.374684	0.402872	133.429829	8
20	93		3	4.0	0.429608	0.468187	208.456473	16
21	94		1	4.0	0.428986	0.470439	348.08551	18
22	23		1	4.0	0.382924	0.41723	146.224423	6
23	121		1	4.0	0.447806	0.48536	247.424398	19
27	4		1	4.0	0.407153	0.445664	273.446514	15
29	10		1	4.0	0.421652	0.452984	115.791584	9
30	16		1	4.0	0.368159	0.398086	36.917253	11
31	18		1	4.0	0.406593	0.449606	166.514812	18
32	22		1	4.0	0.367246	0.395552	57.361726	10
33	24		1	4.0	0.378033	0.411318	111.220496	12
34	97		1	4.0	0.445783	0.473255	153.843422	10
35	122		1	4.0	0.44713	0.481419	260.158026	14
36	126		3	4.0	0.431487	0.459459	147.669118	7
38	32		1	4.0	0.405479	0.4375	72.939322	9

Fruchterman Reingold: The Fruchterman-Reingold Layout works well for many large social networks, though it may require some adjustment. It's an example of a force-directed algorithm, which uses physical springs as edges to pull connected vertices toward each other and a competing repulsive force to push all vertices away from each other, whether connected or not [5, 7]. **Area:30,000, Gravity:10, Speed:30.**

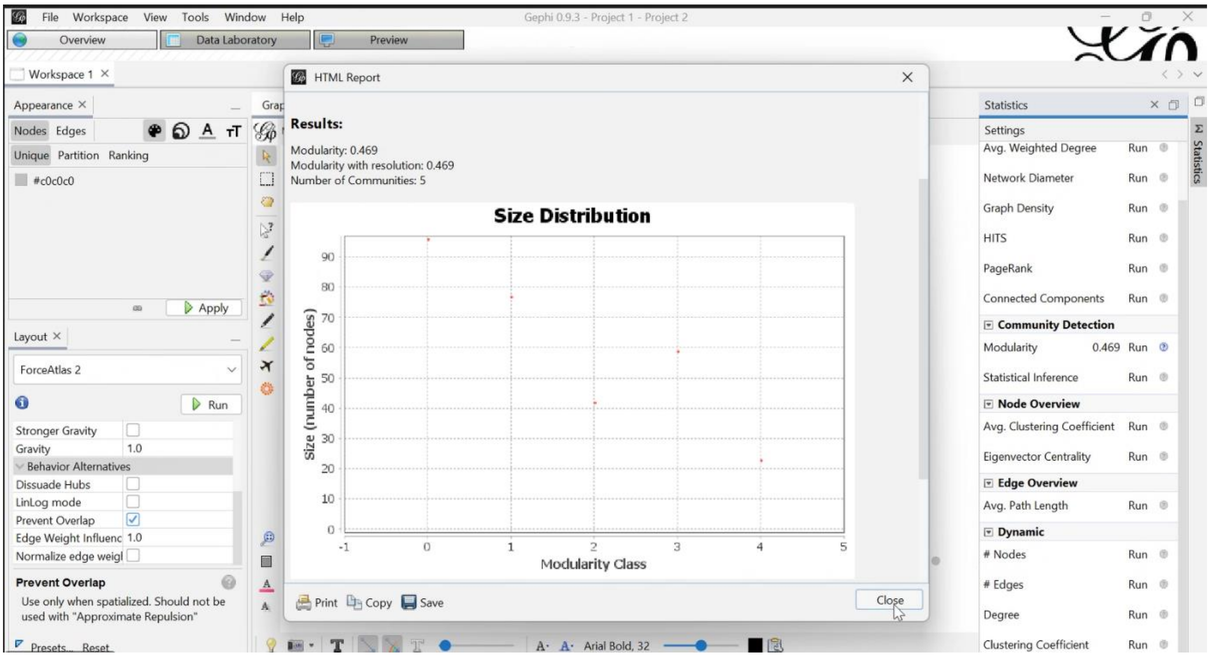


<u>Force Atlas 2:</u> ForceAtlas2 is a force-directed layout that spatializes a network by simulating a physical system. Nodes repel each other like charged particles, while edges, like springs, attract their nodes. These pressures produce a convergent movement toward a balanced state. The data interpretation should be aided by this final arrangement.

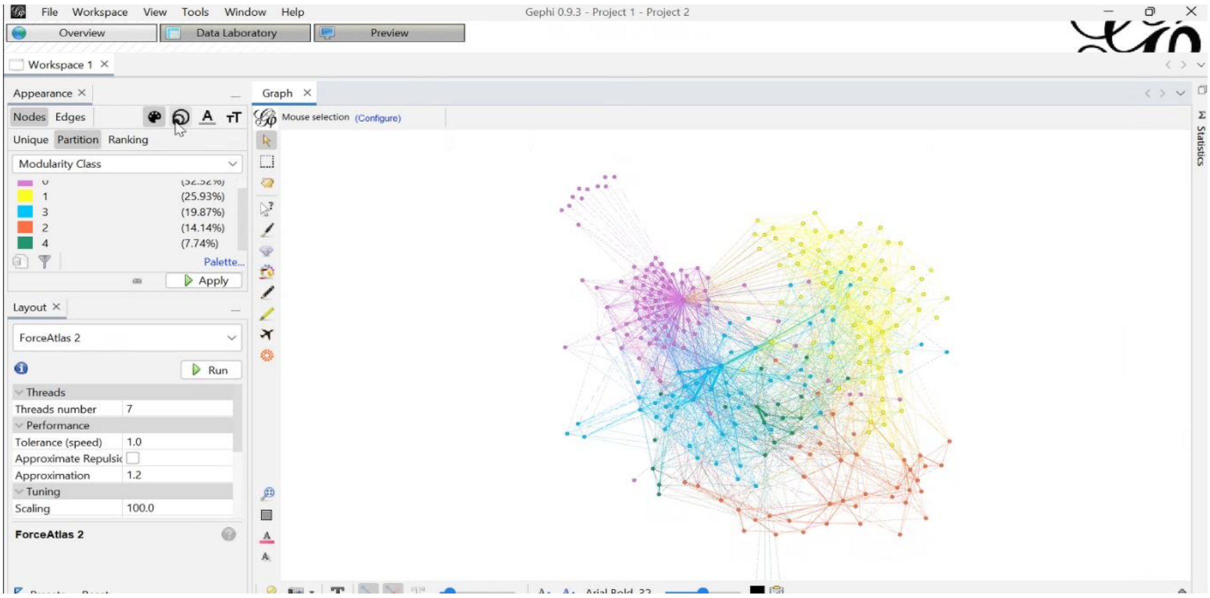
Scaling: 100 → Prevent overlap. Prevent nodes from overlapping.



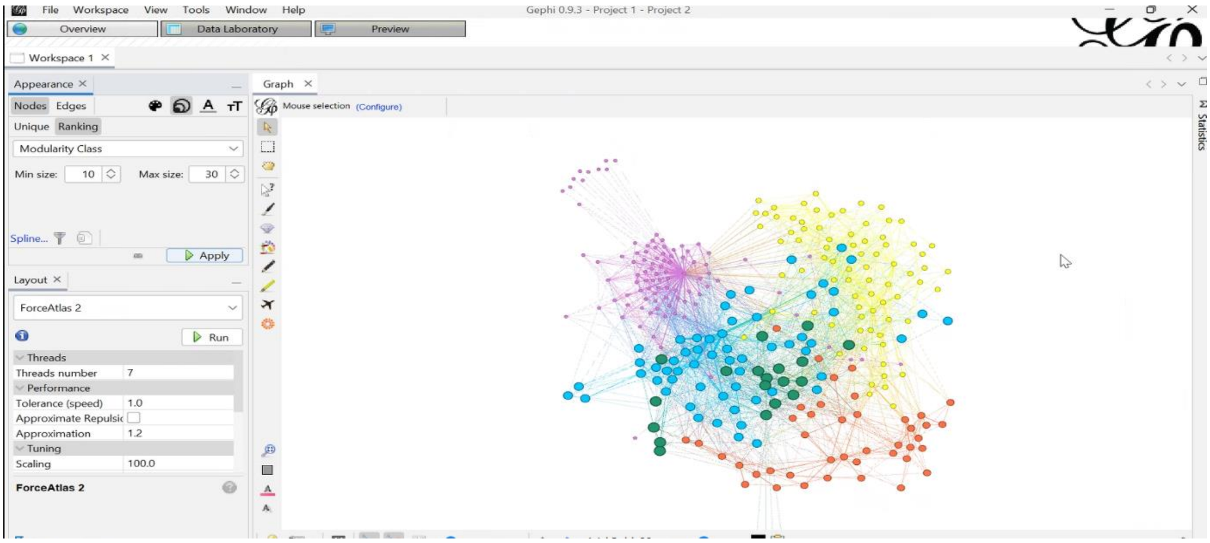
<u>Modularity:</u> The structure of networks or graphs can be measured in terms of modularity. It was created to assess the strength of a network's module division (also called groups, clusters or communities). High-modularity networks exhibit extensive connections between nodes inside modules, but sparse connections between nodes in different modules. In optimization methods for discovering community structure in networks, modularity is frequently used. **Modularity: 0.469**



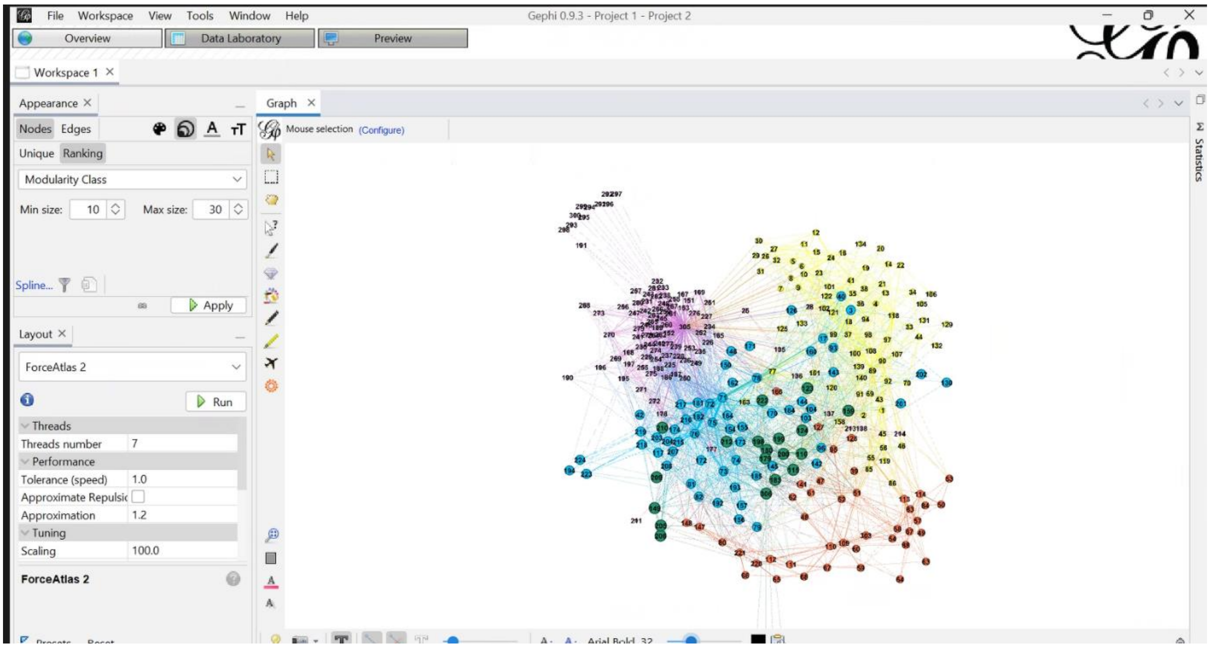
NODES → PARTITION



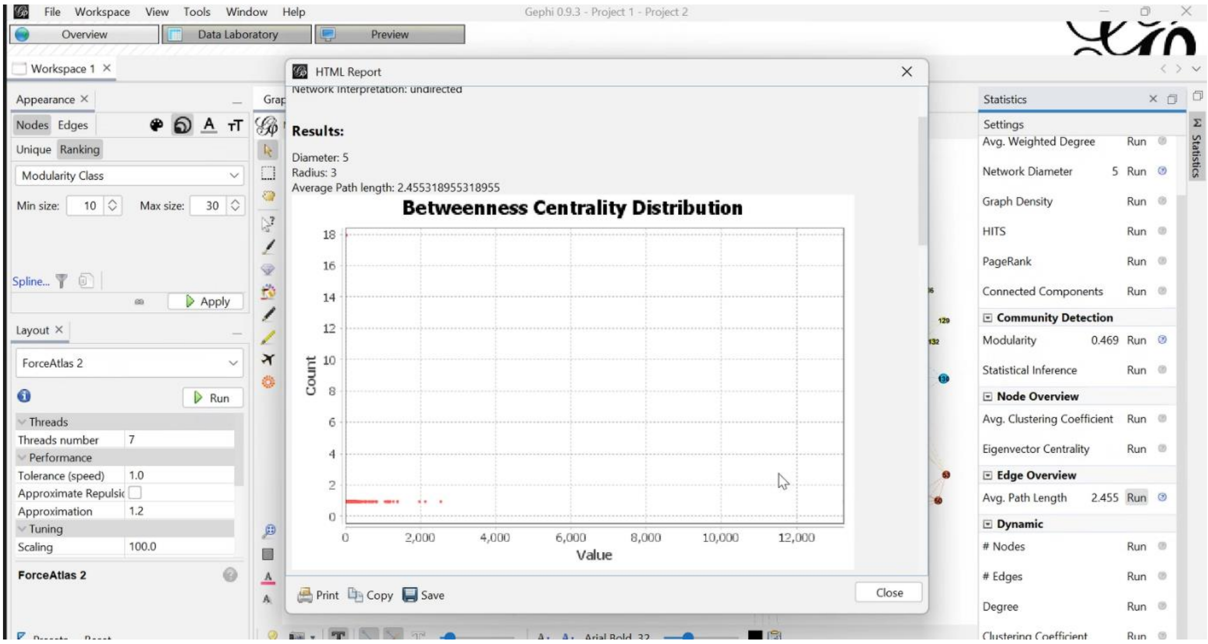
NODES → RANKING → DEGREE



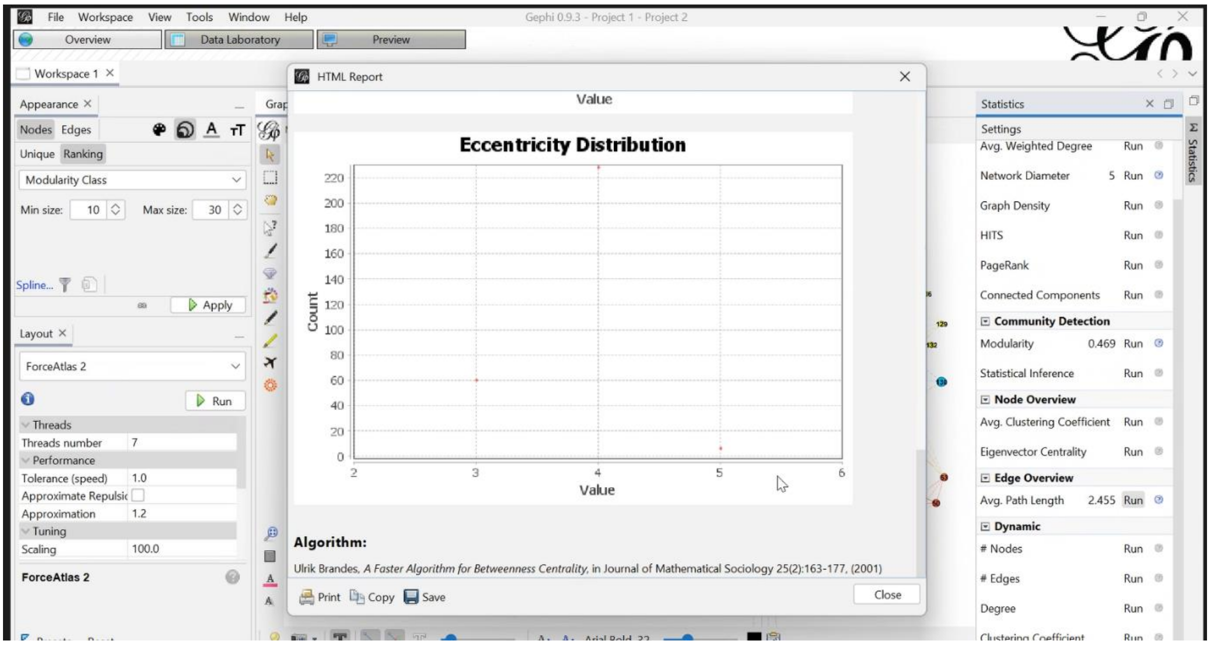
PUTTING LABELS



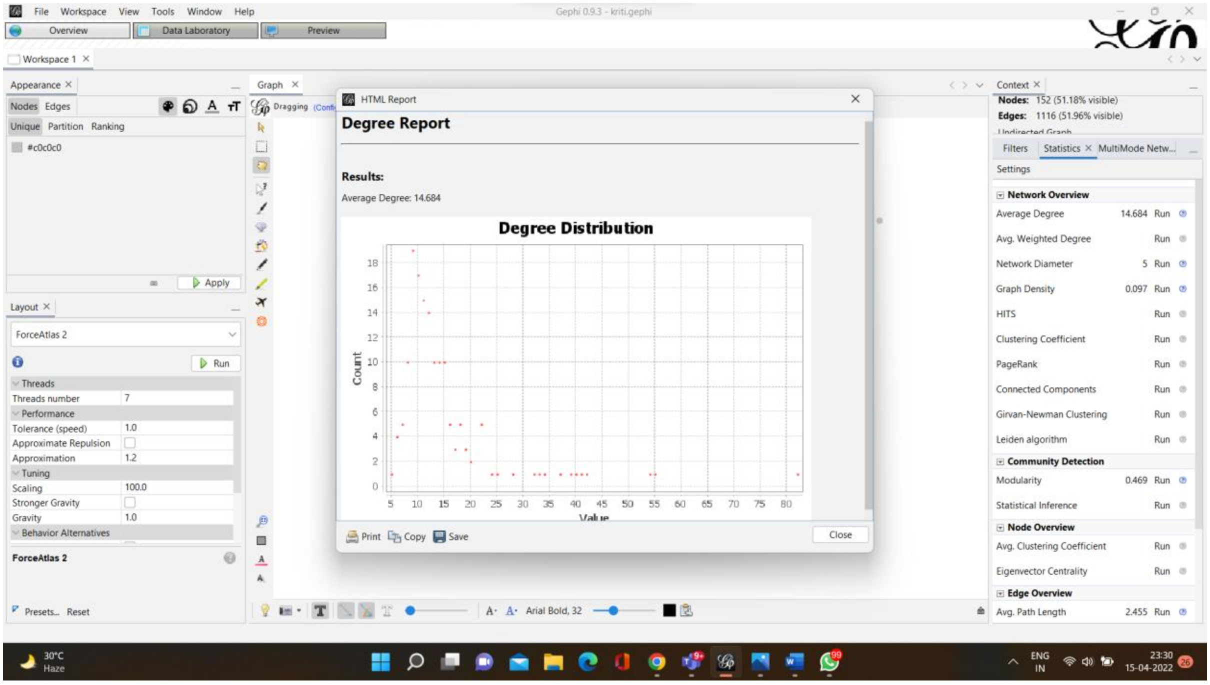
Betweenness Centrality: Betweenness centrality counts how many times a node is on a shortest path between two others after measuring all the shortest paths between every pair of nodes in the network. **Average Path Length: 2.455318955318955**



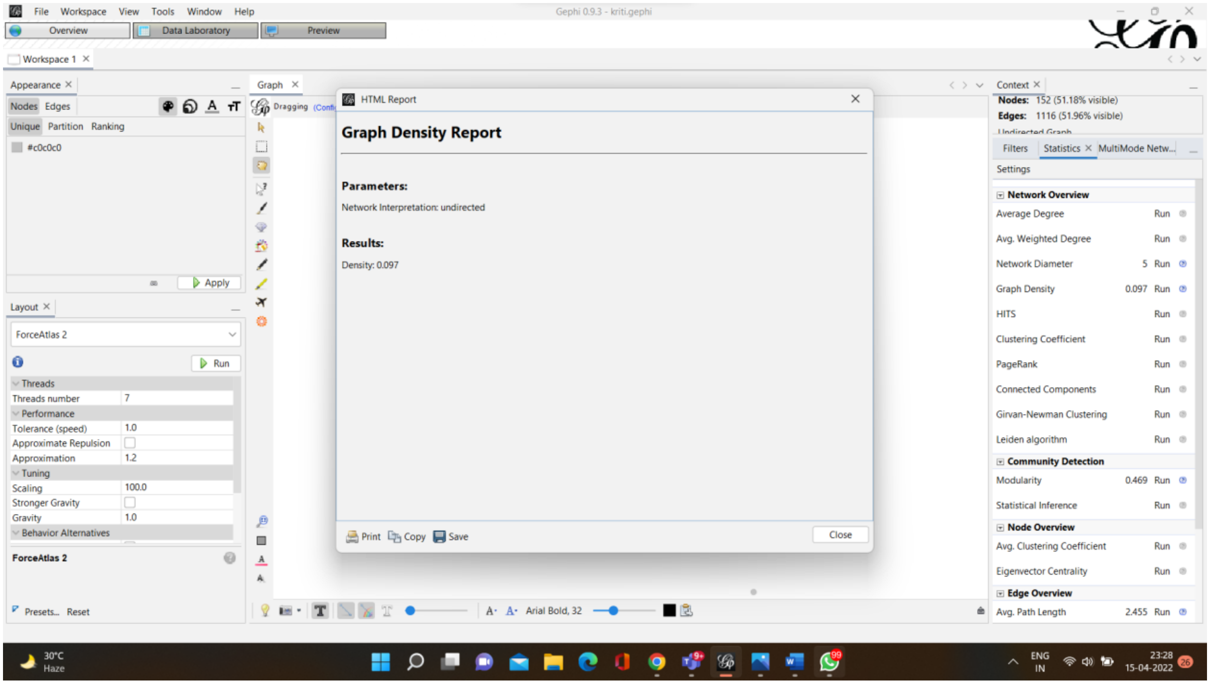
ECCENTRICITY DISTRIBUTION:



STATISTICS → AVERAGE DEGREE REPORT = 14.684



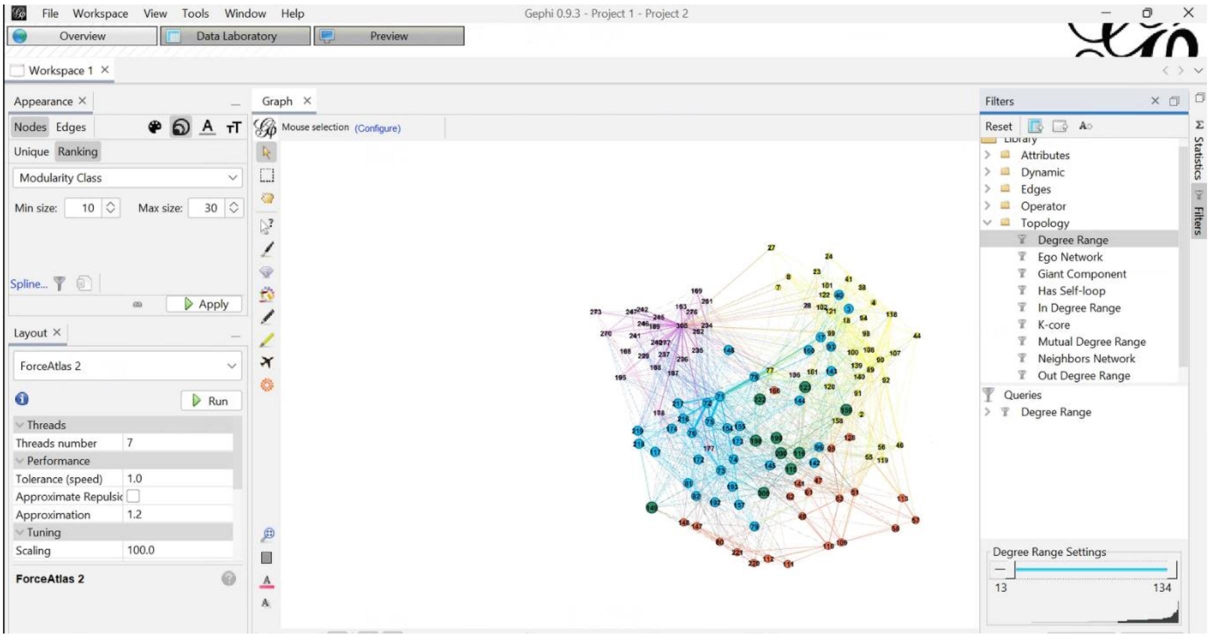
STATISTICS → GRAPH DENSITY REPORT = 0.097



FILTERS

Topology → Degree Range (with Filter)

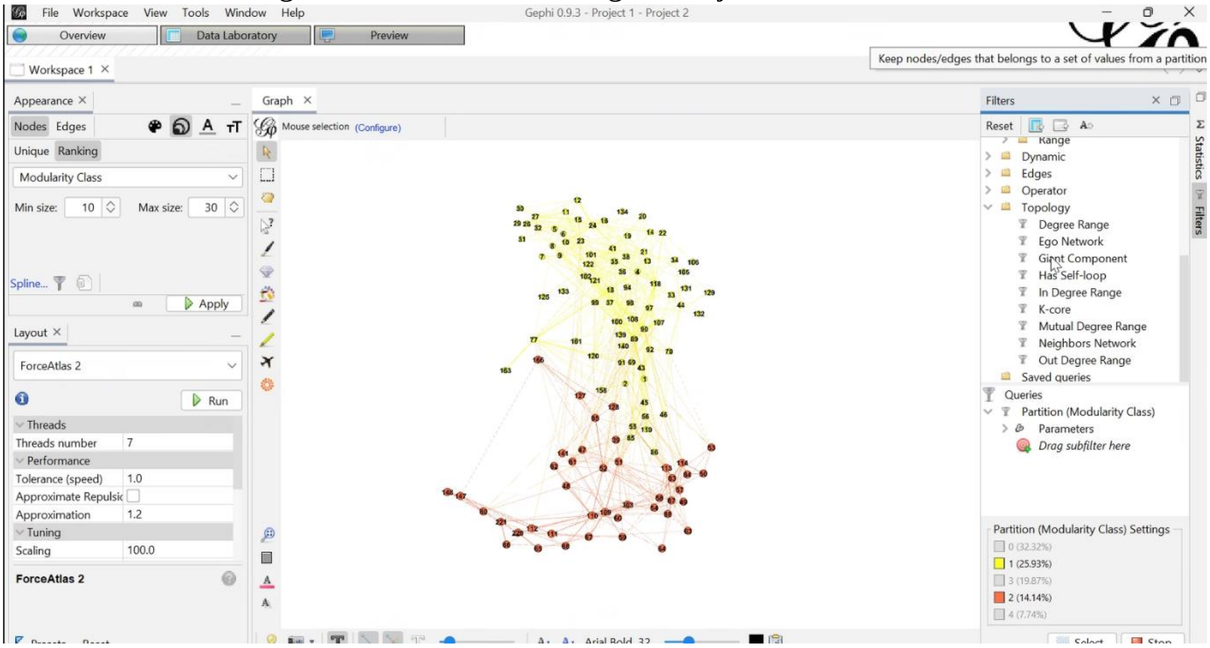
Here we selected the nodes whose degree was more than 13 and less than 134.

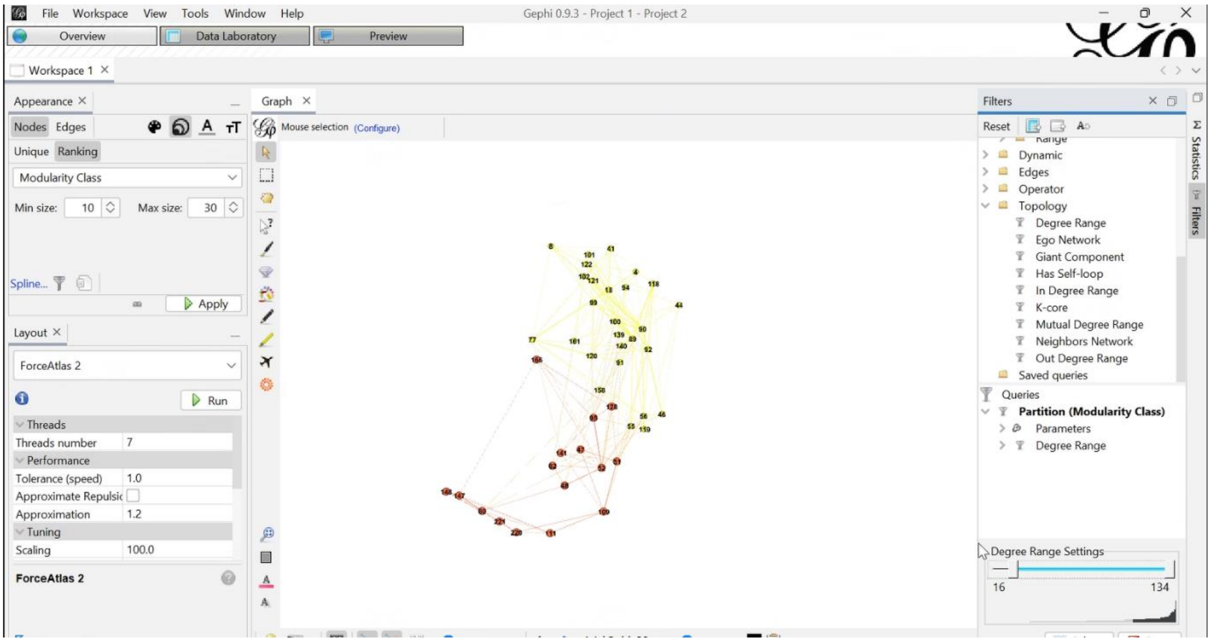


By changing the degree range of the network, we can see that there is a significant change in the graph as we have opted for the directed graph.

Attributes → Modularity Class

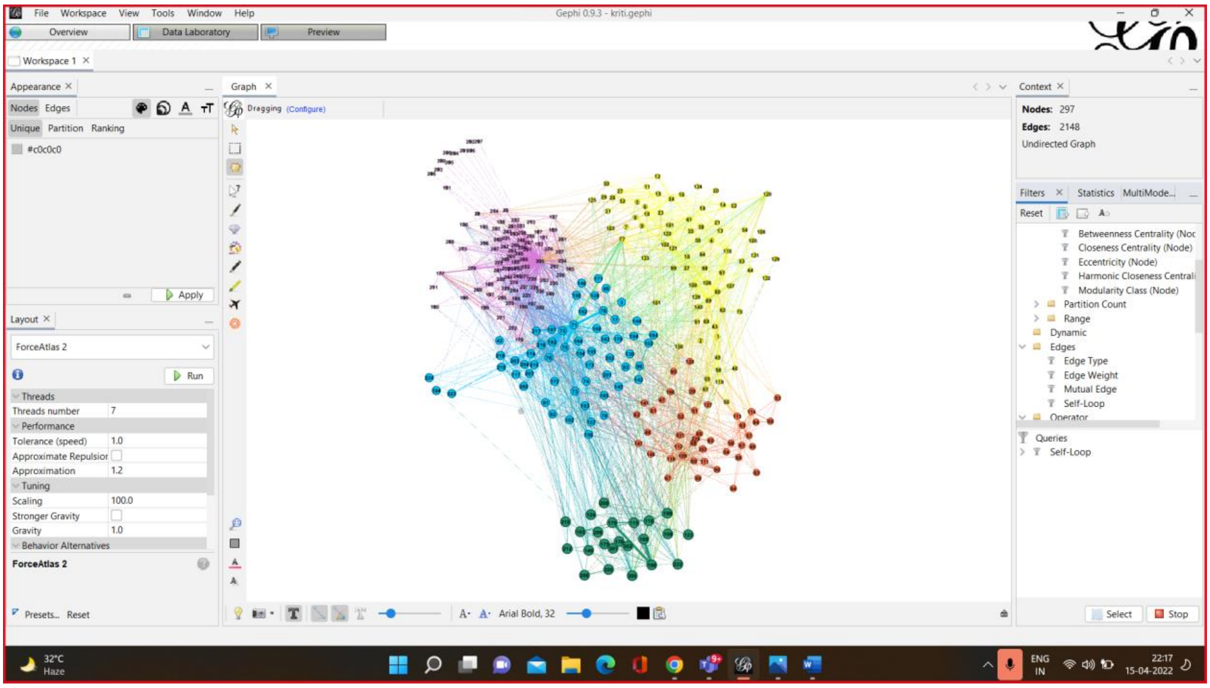
Partitioning was done based on Modularity, the top 2 communities were selected, i.e., the communities with Degree above 16 and forming the major communities were selected





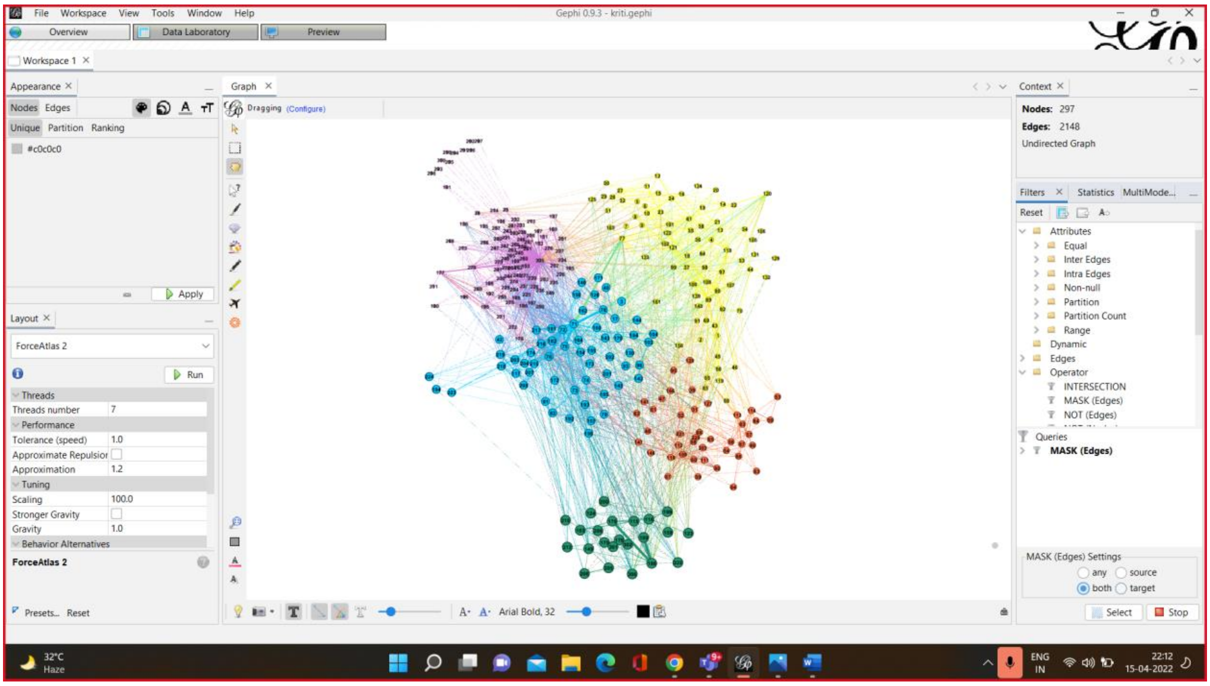
Edges → Self-Loop

This filter removes all the self-loops that exist in the network. After proper comparison, we can see the difference between networks

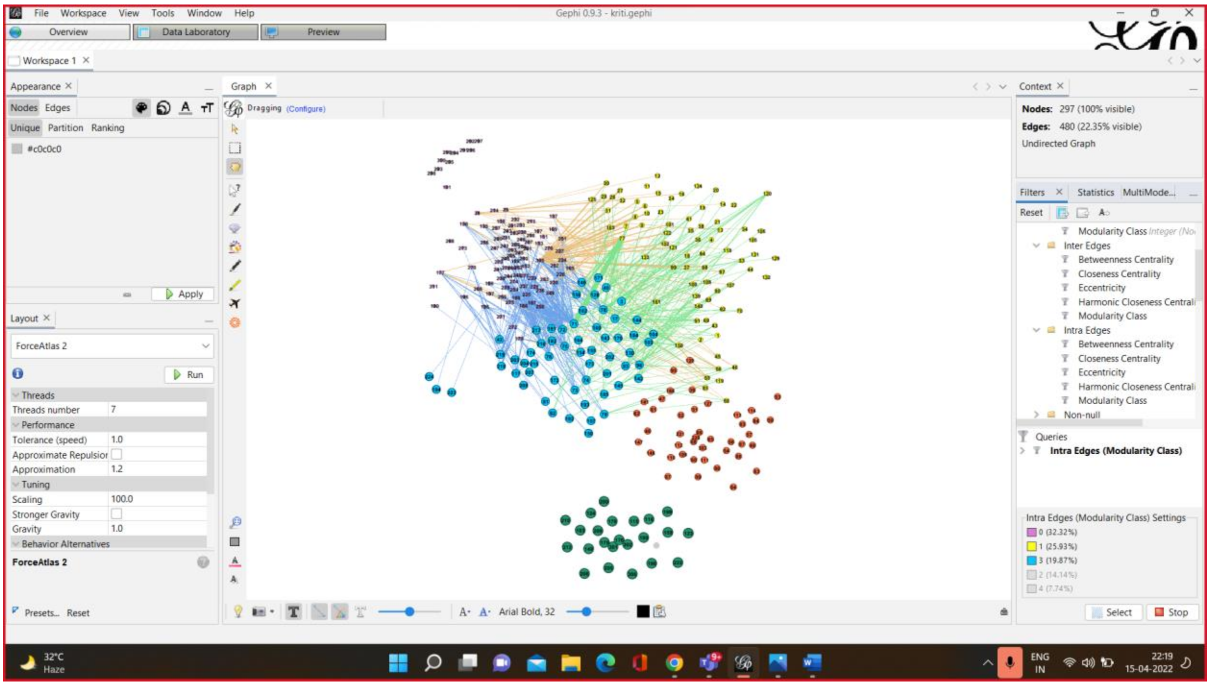


Operator → Edge (MASK)

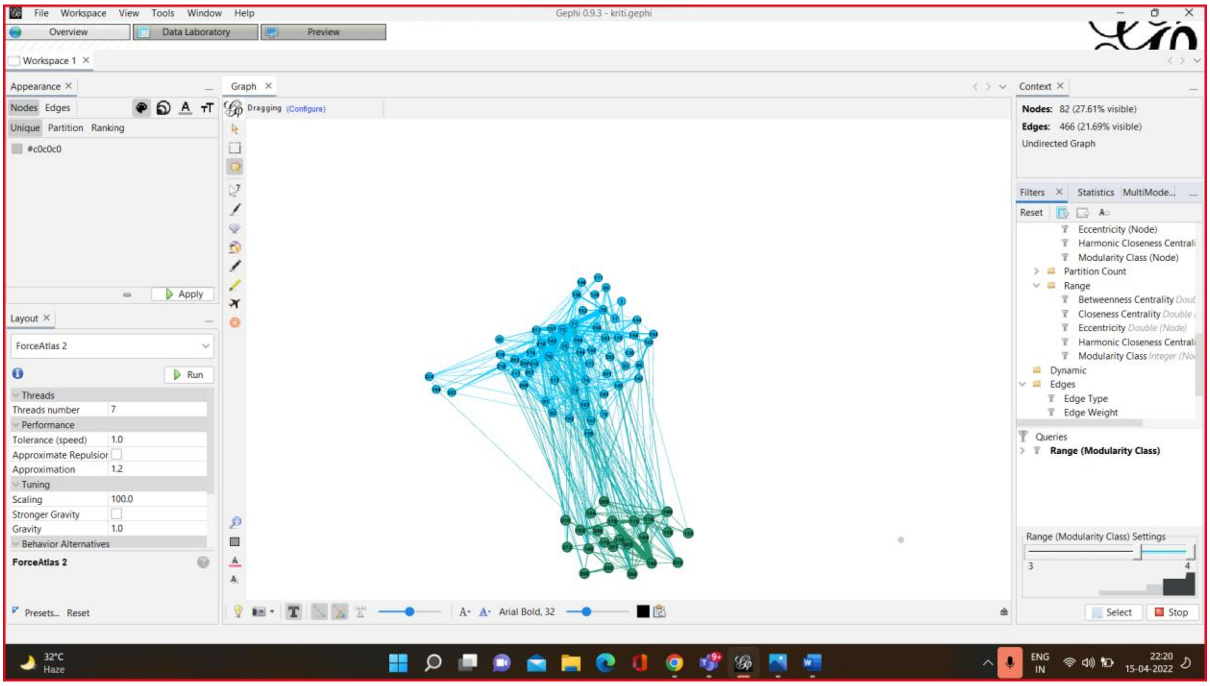
This filter returns the complete graph with only edges from the node filter sub-query.



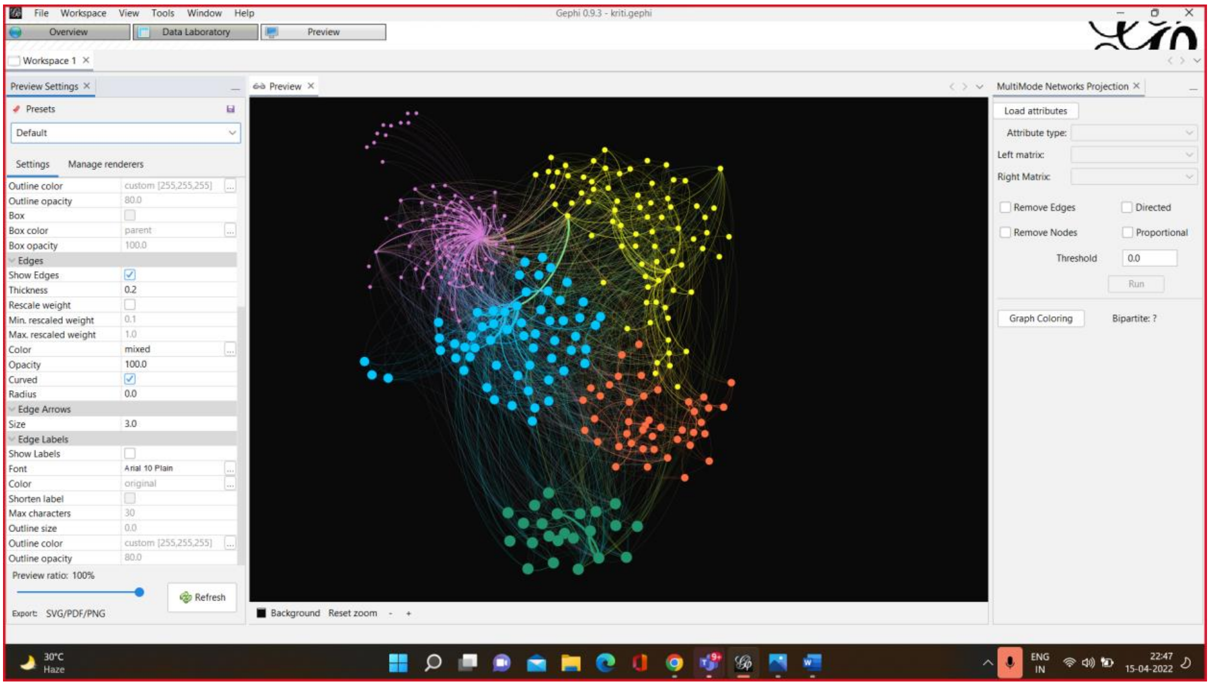
Intra Edges (Modularity Class) (3)



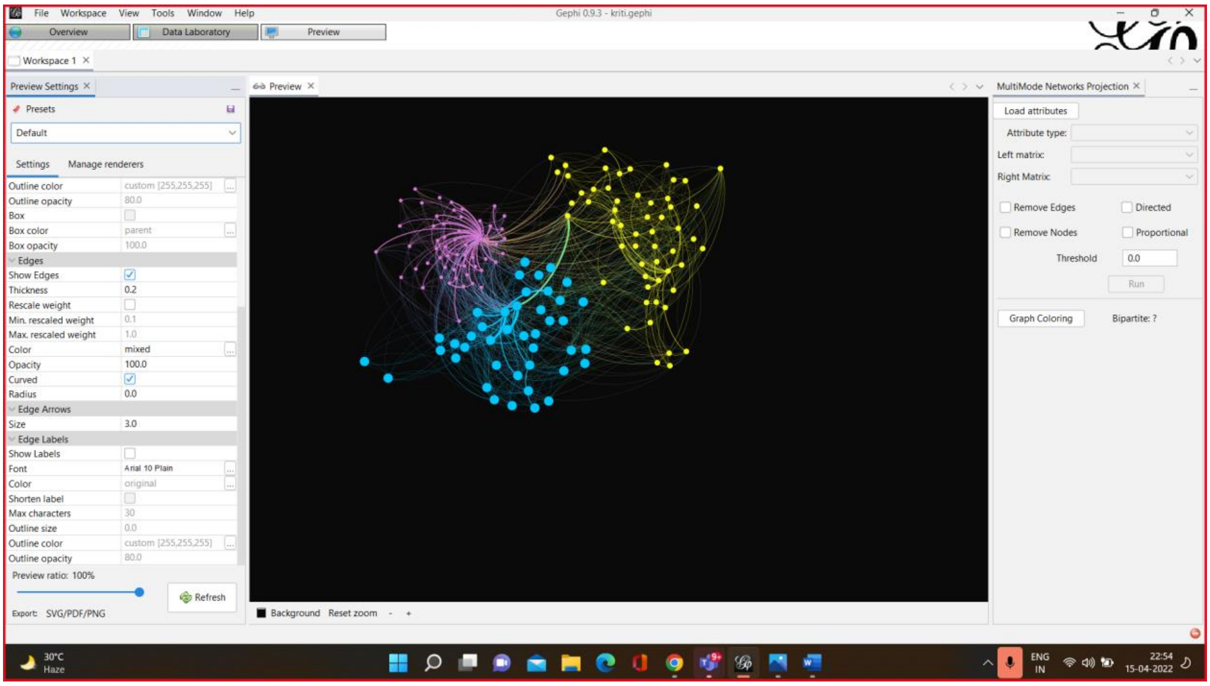
Range → Modularity Class (3)



Preview



Filtered Preview



Conclusion:

The last image suggests that the dataset has many items which are interrelated to each other and the neural networks of C. Elegan are having a strong relationship with each other. There are 5 communities and filtering them suggests that the top 3 communities are bonded strong enough.