

# Analysis of Protein-Protein Interaction Network in Yeast

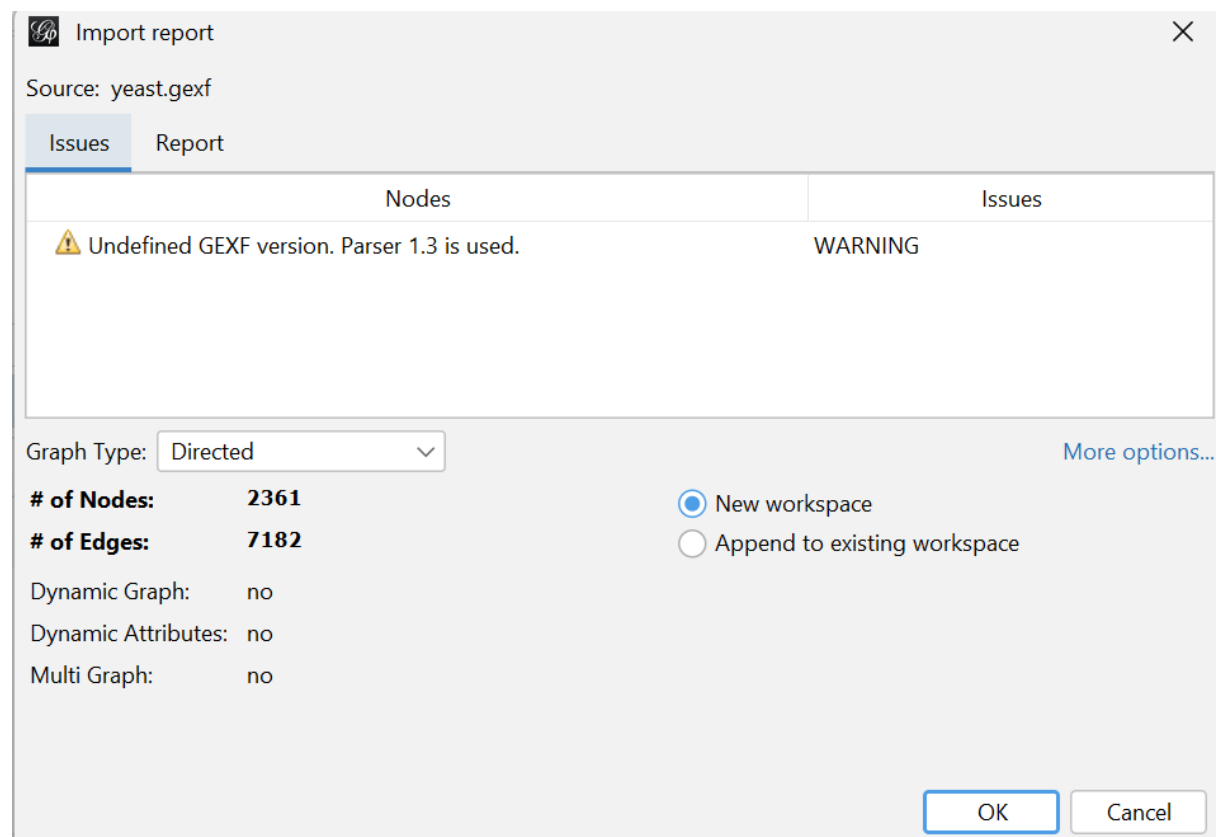
## [ Using GEPHI ]

The dataset represents a protein-protein interaction network in yeast, which is critical for understanding the complex biological processes that occur within cells. Protein-protein interactions (PPIs) are essential for almost all cellular functions, including immune responses and metabolic pathways. This network provides a visual and statistical representation of how proteins within yeast cells interact with one another, forming a complex web of interactions that are vital for the organism's survival and functionality.

In this analysis, the dataset is visualized and examined using GEPHI, a powerful network analysis tool. By exploring various network statistics such as degree distribution, graph density, and modularity, aim is to uncover key structural properties of the protein interaction network. These properties can shed light on the roles of individual proteins, identify critical hubs within the network, and highlight potential protein complexes that may be involved in specific biological functions.

This analysis will help in better understanding the topological structure of the yeast protein-protein interaction network, providing insights that could be valuable for further biological research and applications, such as drug discovery and understanding disease mechanisms.

### STEP 1: Load the dataset GEXF File



The image shows the 'Import report' dialog box in GEPHI. The 'Source' is 'yeast.gexf'. The 'Issues' tab is selected, showing a warning: 'Undefined GEXF version. Parser 1.3 is used.' The 'Report' tab is also visible. Below the issues, the 'Graph Type' is set to 'Directed'. The statistics show '# of Nodes: 2361' and '# of Edges: 7182'. The 'Dynamic Graph' and 'Dynamic Attributes' are both set to 'no'. The 'Multi Graph' is also set to 'no'. The 'New workspace' radio button is selected, and the 'Append to existing workspace' radio button is unselected. The 'More options...' link is visible. The 'OK' and 'Cancel' buttons are at the bottom right.

Nodes	Issues
⚠ Undefined GEXF version. Parser 1.3 is used.	WARNING

Graph Type: Directed

**# of Nodes:** 2361  
**# of Edges:** 7182

Dynamic Graph: no  
Dynamic Attributes: no  
Multi Graph: no

☒ New workspace  
☐ Append to existing workspace

[More options...](#)

OK Cancel

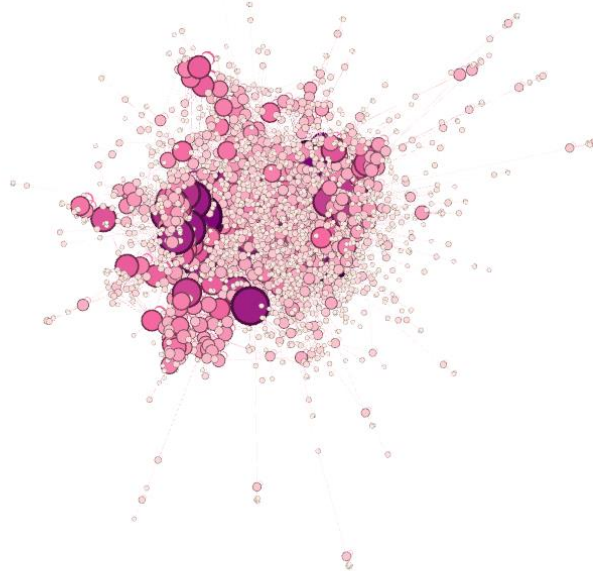
The number of nodes is 2361, and the edges i.e. connections are 7182.

**Nodes:** 2361

**Edges:** 7182

Directed Graph

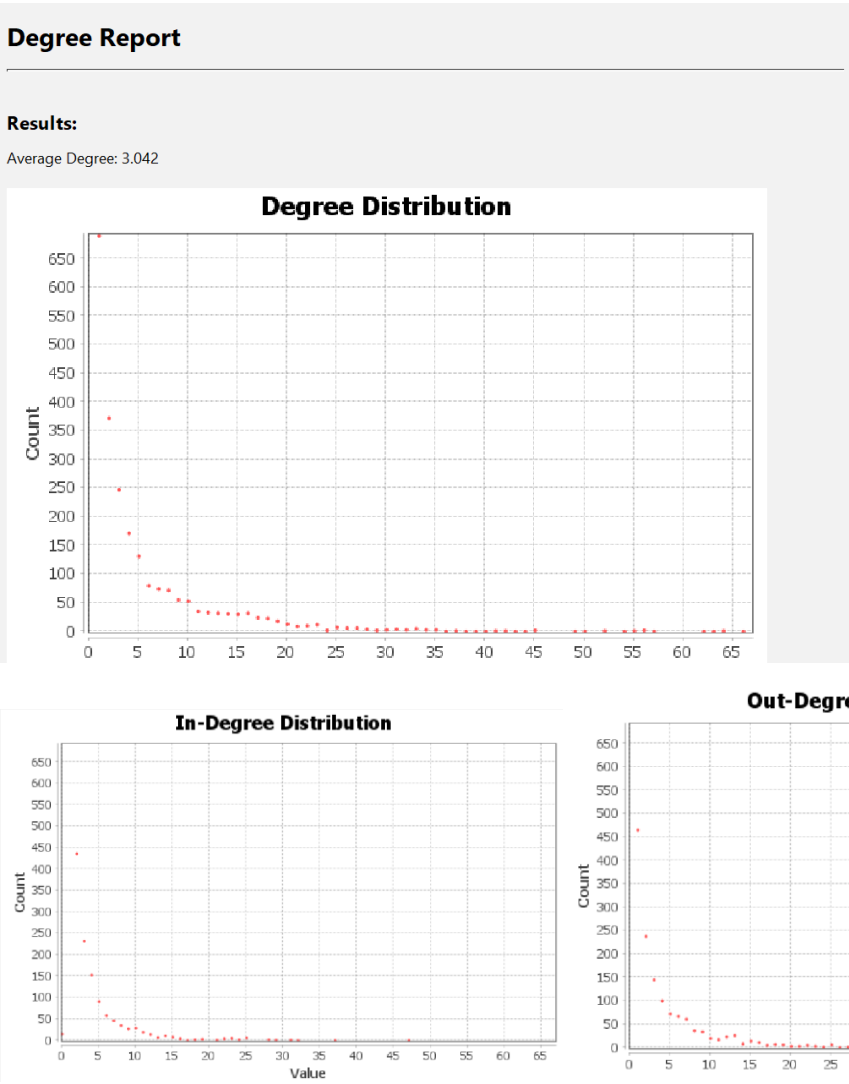
Layout- Force ATLAS 2



STEP 2: Conduct Statistical Analysis

1] Average Degree

The average degree of a network is a key metric that represents the average number of connections (edges) each node (protein) has within the network. In this analysis, the average degree of the yeast protein-protein interaction network is calculated to be 3.042.



2] Graph Density

**Graph Density Report**

**Parameters:**

Network Interpretation: directed

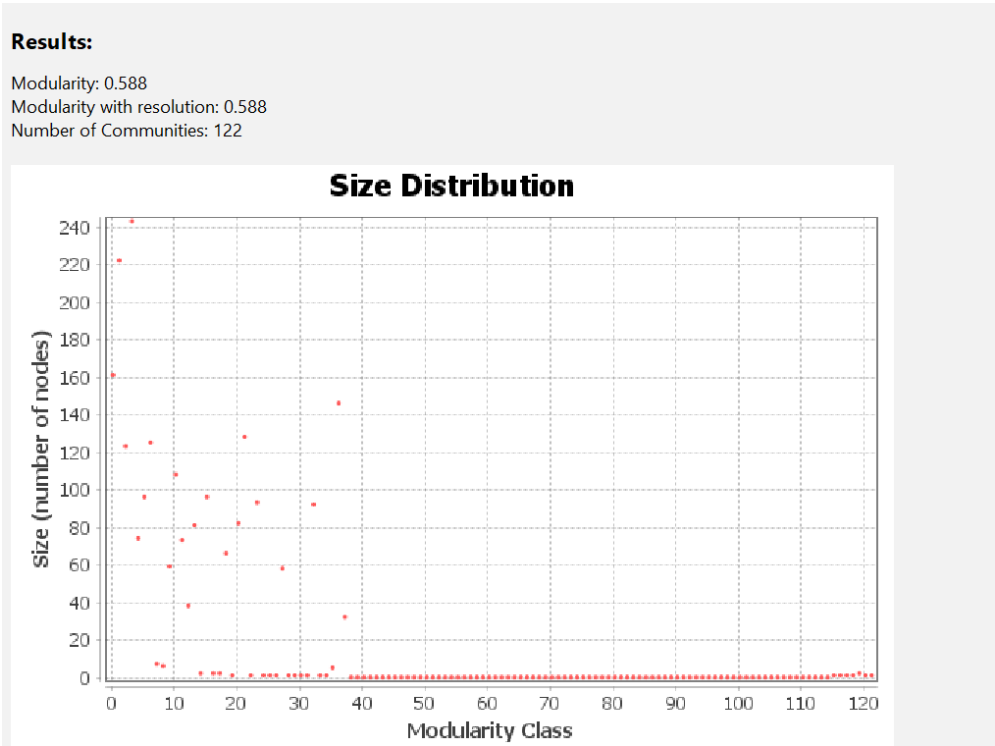
**Results:**

Density: 0.001

The graph density is a measure of how many edges are present in the network compared to the maximum possible number of edges. For the yeast protein-protein interaction network, the density is calculated to be 0.001.

The graph density of 0.001 for the yeast protein-protein interaction network indicates a sparse and highly specialized network structure. This low density reflects the biological reality that proteins tend to engage in specific, functionally relevant interactions rather than forming numerous non-specific connections. The sparse nature of the network highlights the importance of modularity and the potential existence of distinct functional communities within the network. Understanding this low density is crucial for interpreting the overall architecture of the yeast protein-protein interaction network and for identifying key protein clusters and interactions that are vital for cellular processes.

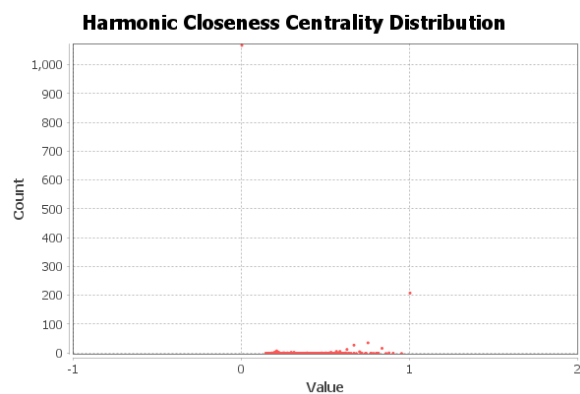
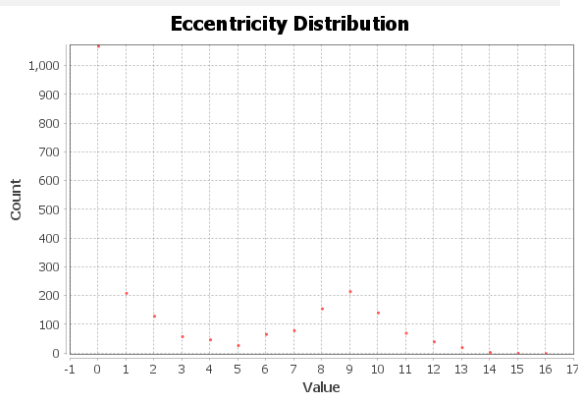
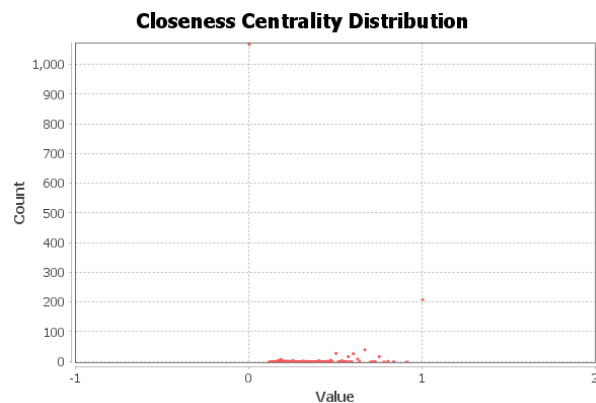
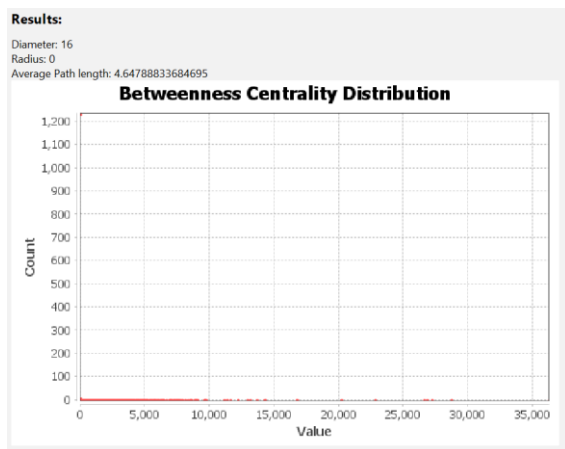
3] Modularity



In the yeast protein-protein interaction network, the modularity is calculated to be 0.588, and the network is divided into 122 communities. A modularity score of 0.588 suggests that the network has a well-defined community structure. This network organization suggests that proteins in yeast are grouped into many specialized clusters, each playing specific roles within the cell

#### 4] Average Path Length

The average path length of 4.648 in the yeast protein-protein interaction network indicates a moderate level of connectivity among the proteins. This metric suggests that, on average, it takes just under five interactions to connect any two proteins in the network, highlighting the efficiency of communication within the cellular environment. A lower average path length often correlates with a more robust and accessible network, which is crucial for the rapid transmission of signals and the coordination of cellular functions.



## **Conclusion**

Based on the statistical analysis results for the yeast protein-protein interaction network:

1. Average Degree (3.042): This indicates that, on average, each protein interacts with about three other proteins. This relatively low average degree suggests a network where individual proteins maintain specific interactions, potentially leading to functional specialization.

2. Graph Density (0.001): The low graph density value indicates that the network is sparse, with only a small fraction of potential interactions actually realized. This sparsity may facilitate focused interactions, allowing proteins to maintain distinct roles without overwhelming connectivity.

3. Modularity (0.588): A modularity score of 0.588 suggests that the network has a significant community structure, with distinct groups of proteins that interact more frequently with one another than with proteins in other groups. This modularity is essential for organizing biological functions and may indicate the presence of specialized protein complexes.

4. Number of Communities (122): The identification of 122 communities within the network reflects a high level of organization, suggesting that the proteins can be grouped into various functional modules, each possibly contributing to different biological processes.

5. Average Path Length (4.648): The average path length of 4.648 indicates efficient connectivity among the proteins, as any two proteins can be linked through approximately five interactions. This efficiency is crucial for rapid communication and coordination of cellular functions.

In conclusion, the analysis demonstrates a well-organized protein-protein interaction network in yeast, characterized by moderate connectivity, clear modular structures, and efficient communication pathways. These insights can aid in understanding the biological significance of protein interactions and may provide a foundation for further investigations into cellular processes and disease mechanisms.