Analysis of Mouse Brain Neural Network

Aditya Jasuja
aditya.jasuja@wsu.edu

EECS
Washington State University
Pullman, WA-99164

Tejas Ghanwat
tejas.ghanwat@wsu.edu

EECS
Washington State University
Pullman, WA-99164

Abstract

Studying the human brain neural network has been a topic of critical importance in the study of the neural networks or network science in general. In order to study brain neural connectivity, the C.elegans worm neural network and mouse brain neural networks are studied as they are less complex and have fewer nodes in the network. Various measures are used, specifically from igraph package for studying the networks and analysis is performed. Few measures used include coreness, degree distribution, k-core decomposition, betweenness and various other centrality measures. Using the results from those measures, we also generated visualizations to observe patterns and to get a deep insight of the structural properties and connectivity. The result indicate many similarities and points us to the specific neuron which drives the information around the network and which of the neurons hold high importance and functionality. We conclude the paper by giving out comparisons between both the neural networks and stating the possibility of future work that can be performed using our analysis.

1 Introduction

The study of neuroscience could be dated back to the Edwin Smith Surgical Papyrus written in the 17th century BC, which consists of the earliest recorded reference to the brain. In the previous centuries, many biological technologies have been emerged through which the researchers are finally able to generate the real neural networks of various animals, which greatly benefit the study of neuroscience. Since the human brain is too complex to understand therefore scientists have been studying the behavior of creatures with simple neural networks. C. elegans is the only creature that we have complete knowledge of its nervous system. Many analyses have been performed on the C. elegans nervous system since connectome data was made available.

The architecture of nervous systems of most of the animals shows bilateral symmetry. C. elegans worm brain also shows this kind of symmetry. In our work instead of studying C. elegans behavior and biological properties, we focused on exploiting network analysis and techniques to get a deeper understanding of worm neural network structure and compare it to the network properties of mouse brain neural network structure. Our work consists of two parts. First, the properties of the worm neural network were analyzed where we computed basic statistics. Second, we tried to analyze the same properties on mouse brain neural network and compare it with the findings of the worm neural.

The complexity of the nervous system continues to protract efforts to understand its development. We try to attempt computational and statistical representation of the neural development of C. elegans based on availability of the biological data, enabling a spatiotemporal analysis of the developing neuronal network. we carry out a structural analysis of neuronal networks. Thus, we focused mainly on this topic, and performed experiments and thorough analysis on the C. elegans worm neural network and mouse brain neural network from the aspect of a graph and compare the differences on both the network by studying the practical observations which lead to these changes.

The following report discusses the dataset and resources throughout the project. After defining the datasets used in the project, we will discuss the different approaches of network analysis to gain information about the connectivity of the nodes and use those data to compare and conclude the observed results. We have used the following properties to analyze the network structure: clustering coefficient, page rank, k- core decomposition, degree distribution, etc. and gained from these properties, applied to our understanding, Following this discussion, we generate a comparison of the two datasets and conclude our project giving out the possible future work that can be implemented on these datasets. Understanding brain connectivity can shed light on the brain cognitive functioning that occurs via the connections and interactions between neurons. To build effective interfaces for brain connectivity analysis, we identify common visual analysis tasks that neuroscientists carry out in brain connectivity analysis based on an in-depth review of the domain literature.

1.1 Motivation and Problem Statement

The main motivation behind this project is to get a deeper insight into the working of neural networks and various parameters that affect its performance. Our goal is to analyze such brain neural networks that will help us to get a better picture of the working of the neural network of a brain. As mentioned earlier, it is difficult to analyze the human brain because the human brains have synaptic neuron connections which add to the complexity of the neural network structure. So, researchers choose to study lesser complex brain such as the C. elegans worm neural network or the mouse brain neural network that consists of less number of neurons, approximate 300 nodes. On the same lines, we have chosen the C.elegans neural network and the mouse brain neural network to perform our study and we have found some interesting observations in the process that have helped us get a better picture of neuron behavior and identify the critical neurons in the brain.

1.2 Data

We have used two datasets for this project, first is the C. elegans neural network which consists of 297 nodes and 2345 edges. In this dataset the first column represents the id of the original neuron, second represents the weight of the edges and the third column represents the destination neuron. This is a weighted directed network [1].

Second dataset that we have used is the mouse brain neural network which consists of 213 nodes and 21.7 K edges. In this dataset, the first column represents the id of the original neuron and the second column represents the destination neuron. The mouse brain network is a unweighted directed graph [3].

2 Analysis Measures

Various measures used to perform the analysis are mostly part of the 'igraph' library, however, we have also used other measures that are not part of igraph but are equally effective for our study. The first measure we used in the 'coreness measure'. This measure helps to identify the tightly interlinked groups within a network. This way we get to visualize an overall structure of the network and the interlinked cores in it. Degree of a node is simply the number of adjacent edges connected to that node. Based on that the next measure we use is the 'degree distribution'. This measure helps to identify the possibility of a randomly chosen node in the graph having degree 'k'. Next measure is the clustering coefficient. As we know a cluster is a set of nodes that are tightly connected, the clustering coefficient measure helps us to identify what portion of the nodes' neighbors are actually connected. K- core decomposition is actually a process in which we repeatedly remove nodes from the graph with degree less than 'k', so finally, all the nodes left the graph to have degree greater than or equal to 'k'. Fano-factor is defined as the ratio of the variance of the spike rate to the mean of the spike rate. It indicates the variability in neurons' underlying firing rate. 'Rewire' is a unique technique of reconfiguring all the edge connections in the network. In simple terms, it maintains the degree distribution of the nodes and the edges are randomly connected to the nodes. We iteratively repeat the following process,

first randomly select two directed edges e1 = (A,B) e2 = (C,D).

Randomly select one end point of e1 and 1 end point of e2, connect them and also connect other end point of e1 with other end point of e2.

Continue, Randomly select one weight from w1 and w2, assign it one newly created edge, and assign the other weight to the other edge.

Make sure there is no self edge or multi edge.

3 Network Analysis

3.1 Network properties:

1. **Degree distribution:** For a graph, degree centrality assigns an importance score based purely on the number of links held by each node. These are used to find the individuals who are likely to hold most information. It is given by a mathematical expression, P(k) = Nk/N, where

Nk represents the number of nodes with degree k and P(k) is the probability.

2. **Clustering Coefficient:** Clustering coefficient is measured as cn= tn/(dn(dn-1)) where tn is the number of triangle around node n and dn is the degree of node n. We take the clustering coefficient for each individual and take the average as a measure of the network.

3. **K- core decomposition:** It is a process where we iterate the following process: a) Remove all nodes with degree less than k. b) Repeat until convergence. It decomposes all the nodes with degree less than k so in the end, all we are left with are the nodes with degree equal to or greater than k.

4. **Strength:** It is the summation of the edge weights of the adjacent edges for each vertex. This provides us the nodes with the highest strength value depicting us the major importance of those set of nodes.

5. **Maximum flow:** This measure calculates a matrix of maximum pairwise flows within a (possibly valued) input network. It computes the maximum flow from each source vertex to each sink vertex, assuming infinite vertex capacities and limited edge capacities.

6. **Closeness centrality:** This measure scores each node based on their 'closeness' to all other nodes within the network. Centrality measures the average distance from a given starting node to all other nodes in the network. It is defined as:

$$Cij = \frac{1}{\sum j \in Ner(i)Wij} \tag{1}$$

   where j is the neighbor of node i and Wij is the weight of the edge between node j and node i.

7. **Page Rank:** PageRank is a variant of EigenCentrality, also assigning nodes a score based on their connections, and their connections' connections. The difference is that PageRank also takes link direction and weight into account – so links can only pass influence in one direction, and pass different amounts of influence. It estimates how important one node is in the network in term of information flow. Page rank is given by

$$Rn = \frac{1-d}{N} + d\sum \frac{RiAn, i}{\sum jAn, j} \tag{2}$$

   where d is the damping factor and A is the weighted adjacency matrix.

8. **Betweenness centrality:** It measures the number of times a node lies on the shortest path between other nodes. This measure shows which nodes act as a bridge between nodes in a network. It does this by identifying all the shortest paths and then counting how many times each node falls on one. It can be calculated using the formula:

$$ck = \sum \frac{\sigma i, j(k)}{\sigma i, j} \tag{3}$$

   i ,j is the number of shortest path among node i and j and i ,j(k) is the number of shortest path among node i and node j through node k.

9. **Authority and hub score:** The authority scores of the vertices are defined as the principal eigenvector of t(A)*A. The hub scores of the vertices are defined as the principal of A*t(A). Authorities a and hubs h are defined as

$$a = \alpha Ah; h = \beta ATa \tag{4}$$

   where A is the adjacency matrix. These scores are calculated for each node and we use the largest values as two measures of the network.

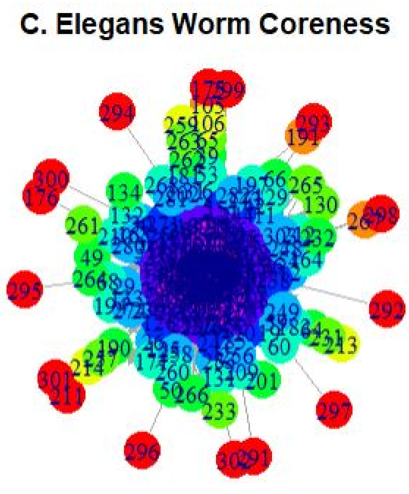
10. **Modularity:** Modularity is one measure of communities in the network, defined as:

$$Q = \sum [e_{uu} - (e_{vv})^2] \tag{5}$$

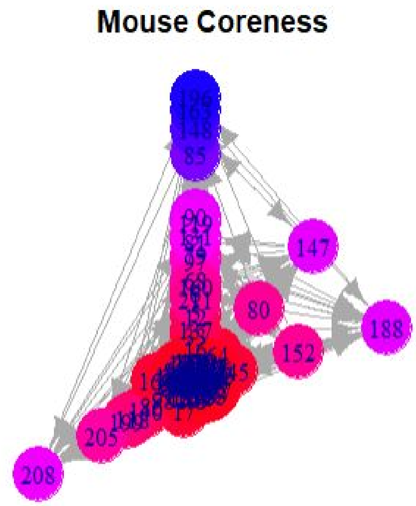
    where the network is fully subdivided into a set of non-overlapping modules M, and euv is the portion of all edges that connect nodes in module u with nodes in module v. This function calculates how modular is a given division of a graph into subgraphs.

3.2 Visualization

Both the datasets are relatively small making it possible and reasonable to visualize. Using color visualization could bring us valuable information and help us to compare both the network structure. The first figure, Fig. 1 (a) and (b) gives us the layout of both the network. Looking at the figures, we can see that both the network differ a lot from one another. The second figure, Fig. 2 shows us the degree distribution of the network. The spikes generated in the graph gives us the idea of what the nodes represent

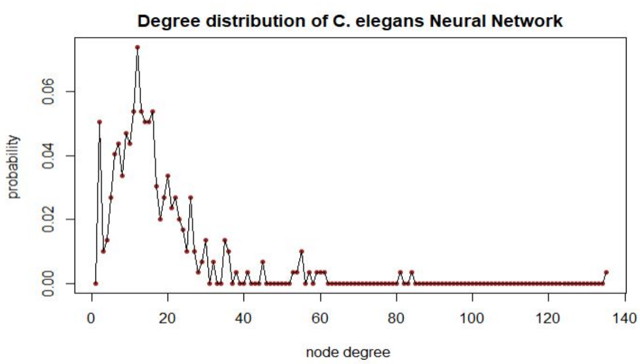


(a) Worm Coreness

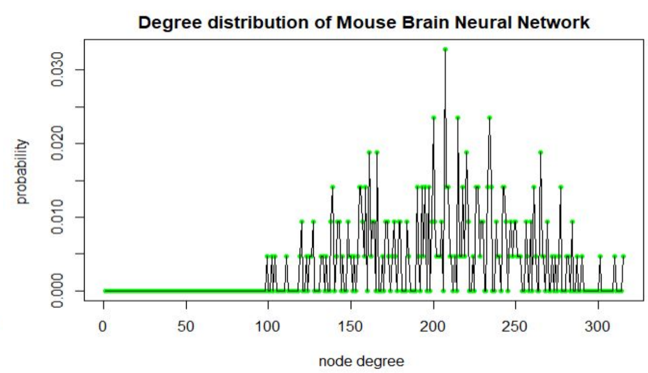


(b) Mouse Coreness

Figure 1: Coreness Measure for both Networks

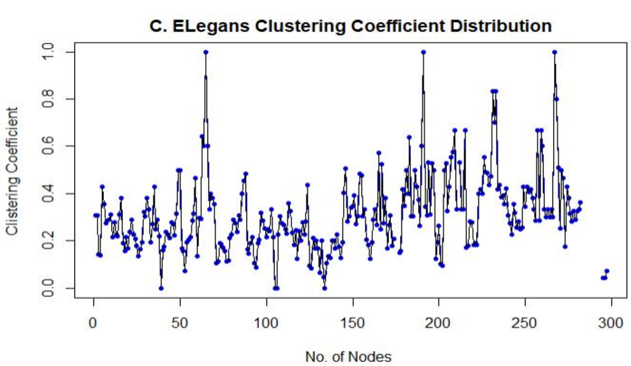


(a) Worm Degree Distribution

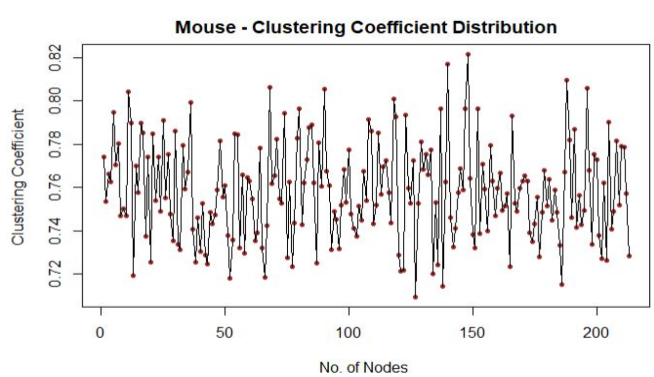


(b) Mouse Degree Distribution

Figure 2: Degree Distribution

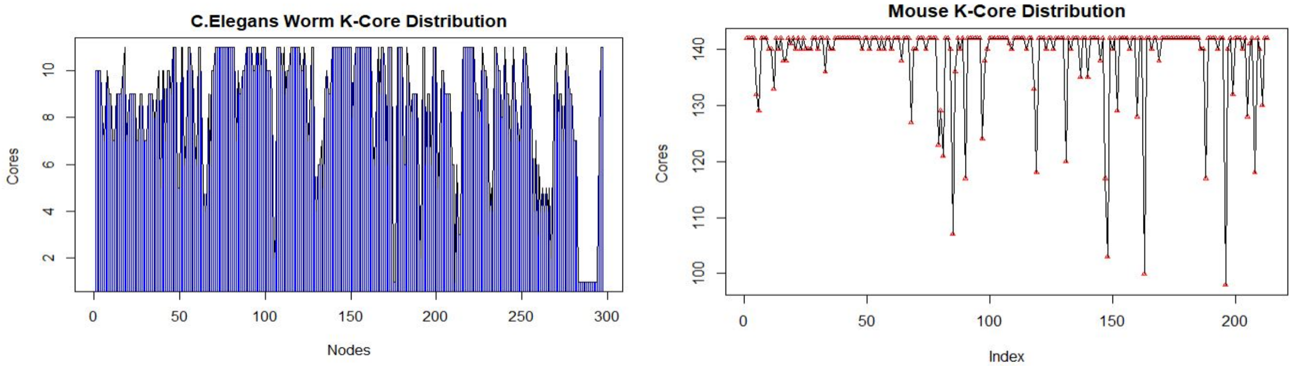


(a) Worm Clustering Coefficient



(b) Mouse Clustering Coefficient

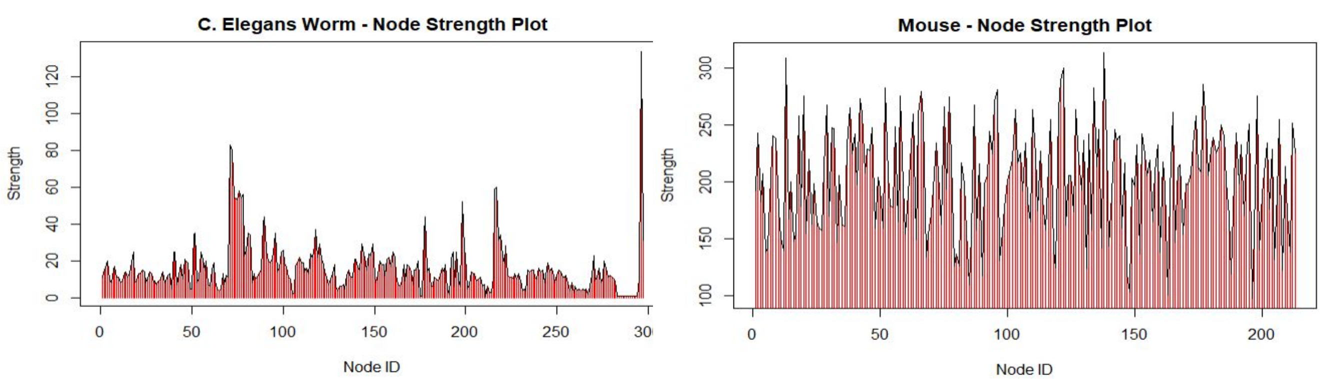
Figure 3: Clustering Coefficient



(a) Worm K-Coreness Distribution

(b) Mouse K-Core Distribution

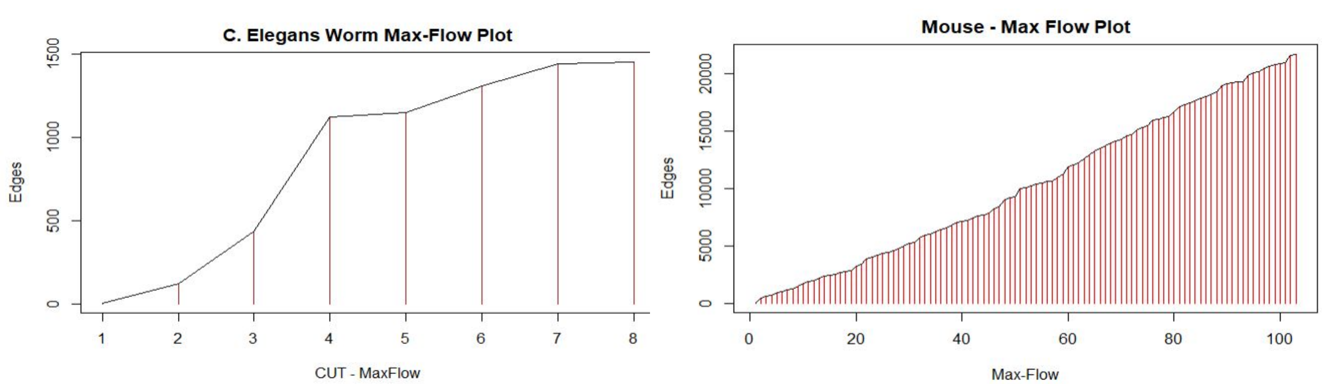
Figure 4: K-Core Distribution



(a) Worm Neuron Strength

(b) Mouse Neuron Strength

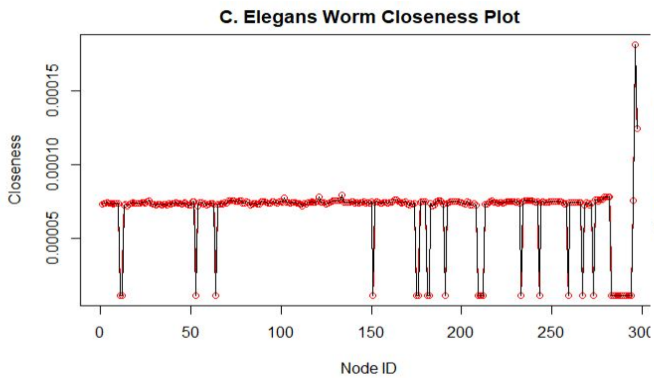
Figure 5: Strength of Neurons



(a) Worm Node Max-Flow

(b) Mouse Node Max-Flow

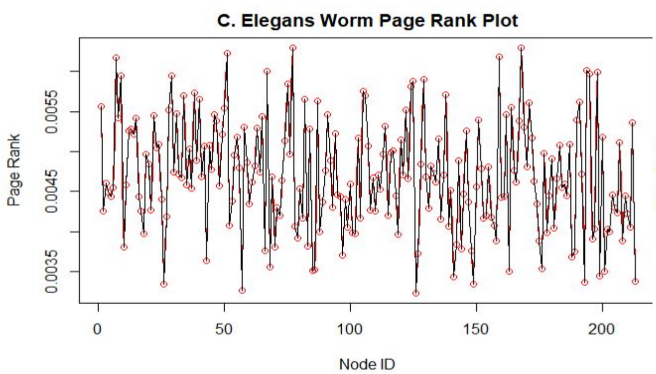
Figure 6: Max-Flow of Nodes



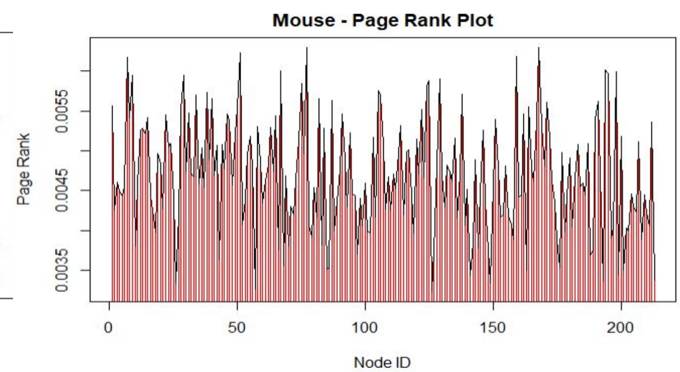
(a) Worm Closeness Centrality

(b) Mouse Closeness Centrality

Figure 7: Closeness Centrality Measure

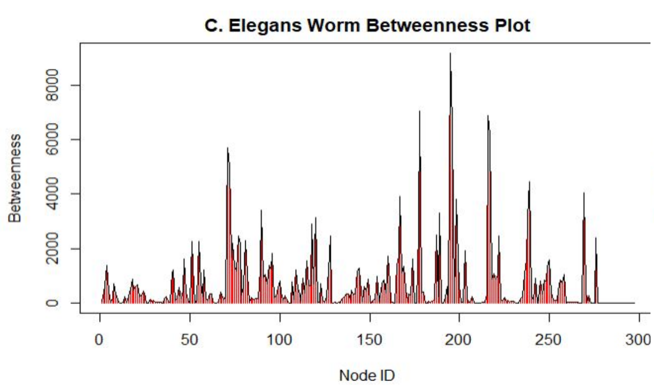


(a) Worm Page-Rank

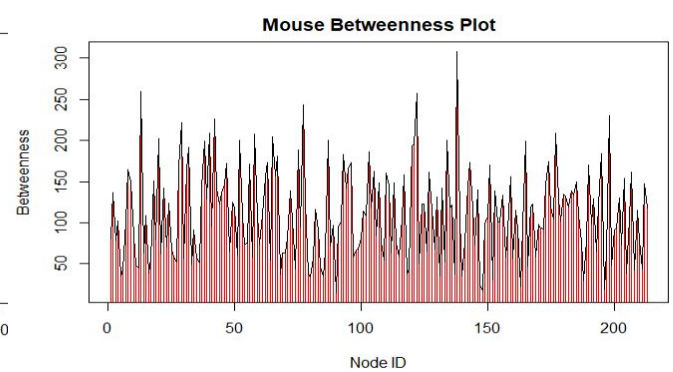


(b) Mouse Page-Rank

Figure 8: Page Rank

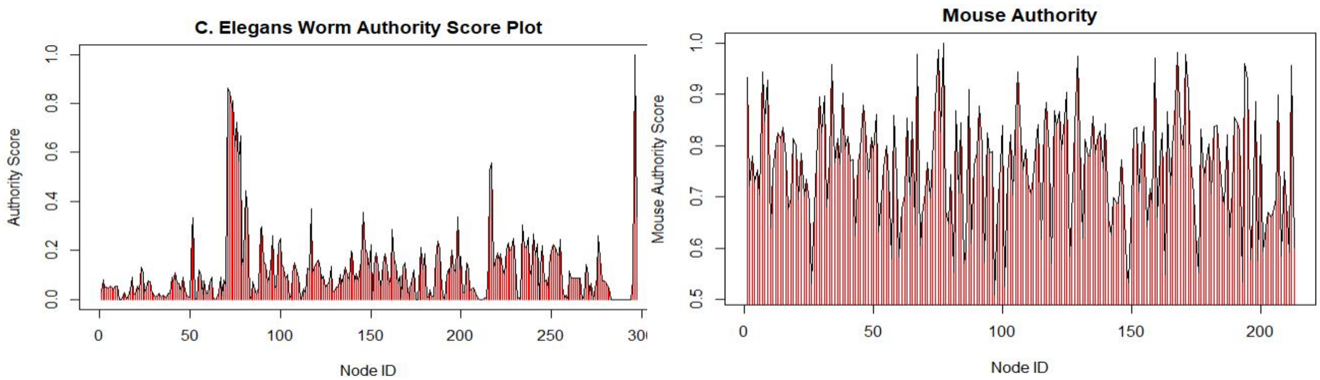


(a) Worm Betweenness



(b) Mouse Betweenness

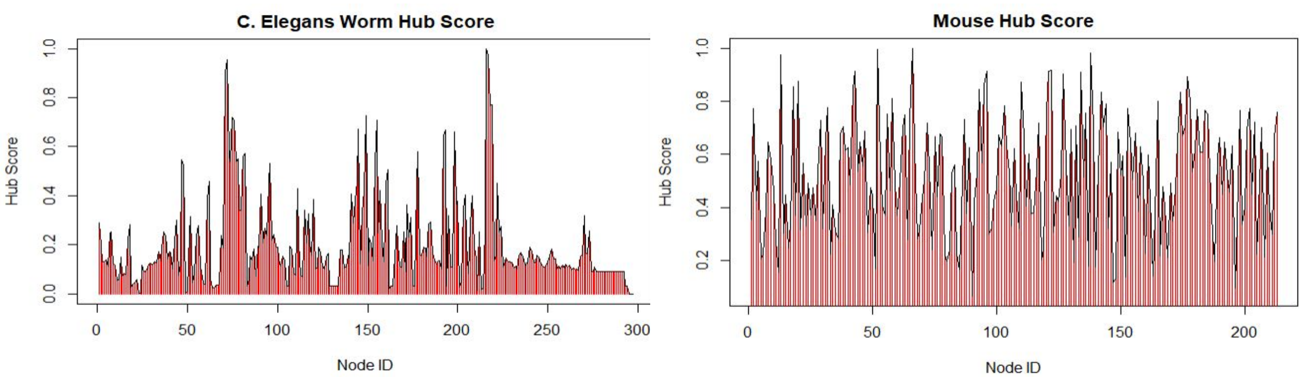
Figure 9: Betweenness



(a) Worm Authority Plot

(b) Mouse Authority Plot

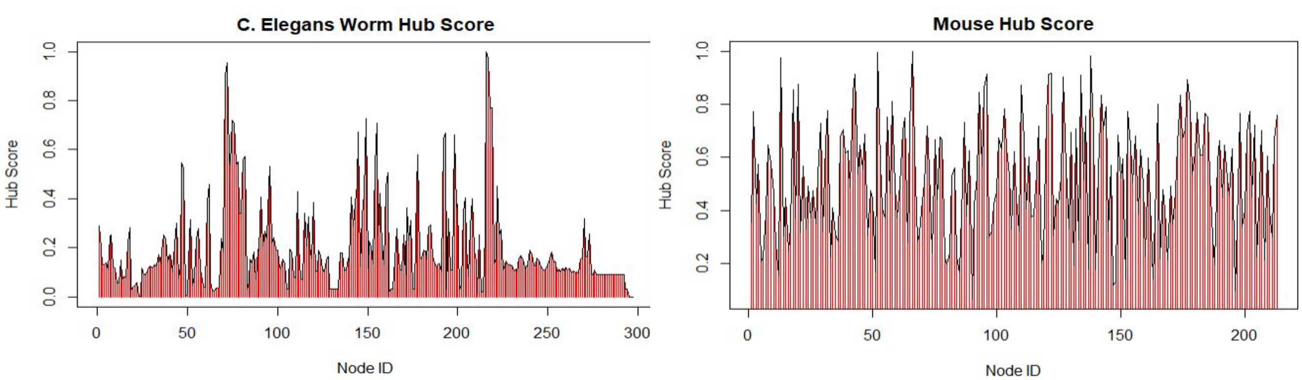
Figure 10: Authority Score



(a) Worm Hub Score Plot

(b) Mouse Hub Score Plot

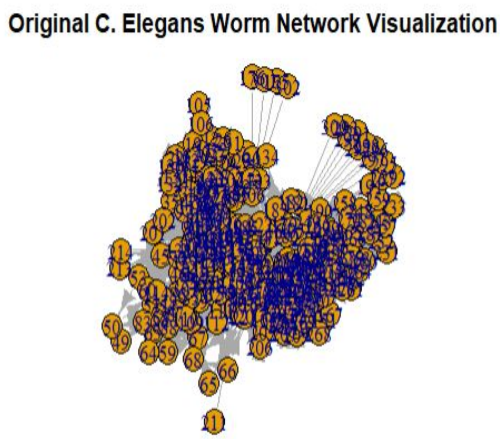
Figure 11: Hub Score



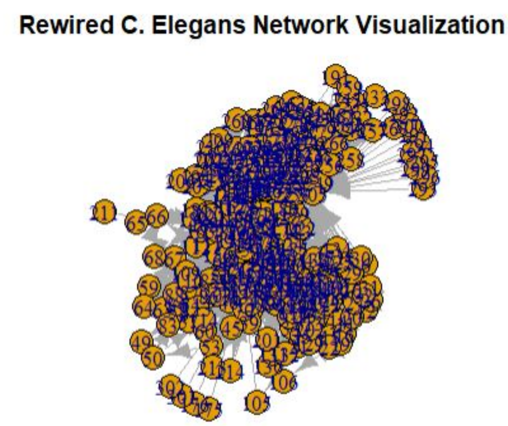
(a) Original Worm Network

(b) Rewired Worm Network

Figure 12: Rewired Worm Network

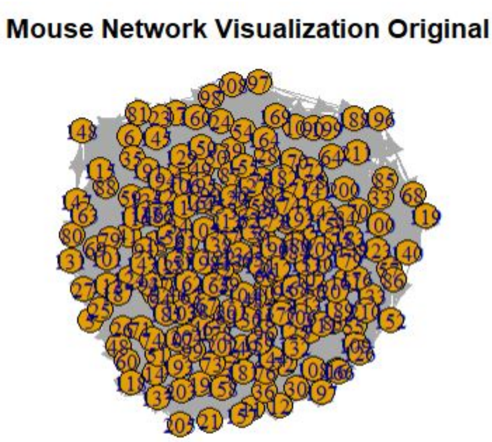


(a) Original Worm Network

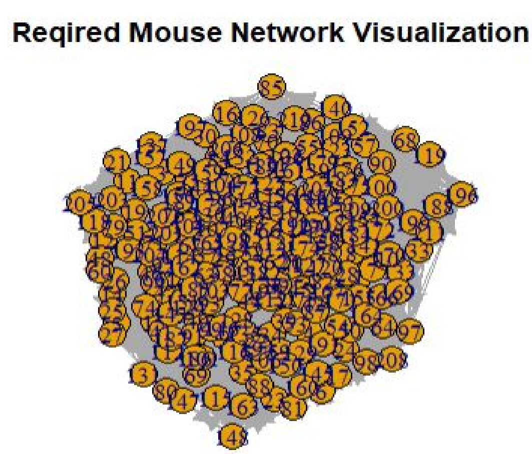


(b) Rewired Worm Network

Figure 13: Rewired Worm Network



(a) Original Mouse Network



(b) Rewired Mouse Network

Figure 14: Rewired Mouse Network

and what action do they perform in the network. The third figure, Fig. 3 is the clustering coefficient which gives us the idea on the group that performs the major role in a network. The fourth figure, Fig. 4 is the K-core decomposition of the network which displays how the information is passed on in the network from one node to another and which nodes are responsible for the flow of information. Fig. 5 depicts the strength of the network i.e. the high- ranking node among all the nodes and controls the incoming and outgoing signals in a network. Figure 6 gives us the maximum flow in a network with minimum loss in the information. Figure 7 is calculating the closeness centrality of the nodes which tells us the minimum distance between the nodes and how close a cluster is in a network. The next figure, Fig.8 is of the page rank, which is another measure of the importance of a node in the network regarding the information flow. Fig. 9 gives us the visualization of the betweenness centrality and the junctions which lie in a path during the transfer of the data. Fig. 10 and Fig. 11 gives us the authority and the hub score as a principal eigen vector in a graph. Figure 12 gives us the modularity or the measure of the communities. Visualizations are shown on next pages.

4 Results and Discussions

The component neurons of the nervous system of C. elegans have simple, unbranched morphologies. Few neurons have more than two processes, and many are monopolar with only a single process. Processes of neurons run in parallel bundles except in the immediate vicinity of their cell bodies, where they join the bundle. Branching typically occurs when a neuron has a process that leaves the main bundle to run out as a commissure or at a discontinuity, where one bundle joins another.

Neurons with a branched structure generally have very similar patterns of branching in different animals; however, there are a few interesting cases where differences occur between animals, or between sides of the same animal. The nerve ring has a high degree of bilateral symmetry and the process runs in a similar position relative to the neighboring processes whether it runs on the left or the right.

The processes of many classes of neuron terminate at the point of contact with a process from a neighboring member of the same class. There is usually a gap junction at this, although there is one case where processes touch and terminate with no gap junction. Chemical synapses in C. elegans occur en passant between neighboring parallel processes. The presynaptic process has vesicle-filled varicosity and a specialized, darkly staining region in the membrane adjacent to the point of contact with the postsynaptic elements. Chemical synapses in C. elegans usually have no visible specializations on postsynaptic elements and consequently, there is often some ambiguity as to the identities of these elements. The graph from the clustering coefficient depicts us with this information. It gives us the idea of the high synaptic regions in the membrane which means that those junctions are of high importance and functionality. As compared to a worm, the mouse brain network has more junctions or synaptic regions.

Many classes of a neuron are found to have regions of the process that are devoid of presynaptic specializations. This could be because the particular class of neuron does not have many synapses in total or that these regions corresponded to regions where there are no suitable postsynaptic partners.

Neuromuscular junctions or the cortical network are responsible for the transfer of the information and are highly dense. Neurons in the cortical network are of two types, excitatory neurons, and inhibitory neurons. The excitatory neurons are responsible for triggering a positive change in the membrane while the inhibitory neurons are responsible for triggering a negative change in the membrane[4]. In the worm neural network, based on the K-core decomposition, we can see that the core greater than 9 lie in the excitatory Neurol regions while the rest lie in the inhibitory neural regions. Similarly, in mouse brain network, the core which is greater than 136 lies in the excitatory region while the rest of them lie in the inhibitory regions.

Degree distribution graph gives us the idea of the motor-sensory neurons. The high spike in the graph tells us that these neurons are responsible for the actual information transfer and data retrieval in the network.

The modularity of the mouse neural network comes out to be 53% and that of the C. elegans worm network is 37% which tells us the degree to which it is modular and is recombined with the benefit of flexibility and variety in use.

Based on the betweenness of the neural networks, it indicates the high holding authority of the nodes and tells us which nodes controls the structure and is in the center of every data flow. Among 3012 nodes in the worm network, node number 195, 178, 216, 196, 217 are few of the nodes which indicate a high score of this measure and are in the path of the information travel.

The average closeness of the worm neural network is 25% and that of the mouse neural network is 31% which gives us the idea of the proximity of nodes in the network. Page rank of worm neural network is 33% and of the mouse neural network is 46% which shows the importance of these nodes in the network and how important is their influence over the other nodes. Based on the strength measure, node 305 is the most critical node in the worm neural network and node 138 has the maximum strength of scaler value 314 which tells the critical node of the network.

The mean authority score of the worm neural network is 11% and that of the mouse neural network is 75% which gives us the approximate idea of those which are connected to many other vertices which are, in turn, connected to many others.

After performing the rewire functionality on both the graphs, we see that the value of the measure drops significantly indicating the randomness of the graph.

5 Related Work

Jingying Yue from Stanford University conducted an experiment on the Analysis of elegans of worm neural network which inspired us to further extend his work and based on his study on the neural network structure, implement the measures on the mouse brain neural network and compare both the network structure with adding few other measures on it [4]. C. Qian, P. Yuan, and X. Jiang also did analysis on the structural and symmetrical properties of C. elegans Neural Network which helped us to understand the technical and biological implementation of our findings[2]. R. Badhwar gave us the idea of the distance constrained synaptic plasticity model of C[5]. elegans neural network. We used his findings and tried to understand how the network is wired for optimal long-range synaptic connections. Our work in this project is using their findings and adding our measures on another different dataset to conclude the similarity and differences in those and generalizing the idea of the analysis on the neural network structure.

6 Conclusion

In this project, we analyzed the structural properties of the C. elegans worm neural network and compared it with the mouse brain neural network. Based on our results, we can say that the mouse brain network is far better structured and well organized. We think that our results imply that the mouse brain neural network is far complex and is still a mystery to work on. This is an open end and we still need answers on the arrangement of nodes and their interactivity with the membrane. The analysis and the model successfully capture specific driver neurons and the high synaptic nodes which indicate high functionality with greater accuracy. In the future, we can try to use more measures and analyze the neural network more dynamically.

References

[1] C.Elegans Worm Neural Network *LINK - http://opsahl.co.uk/tnet/datasets/celegans*$_n$306*.txt*

[2] Chen Qian, Peng Yuan, Xinyi Jiang, *Analyzing Structural and Symmetrical Properties of C. Elegans N,eural Network*

[3] Mouse Brain Neural Network *http://networkrepository.com/bn.php*

[4] Jingying Yue, *Analysis of Elegans Worm Neural Network*

[5] Rahul Badhwara, Ganesh Bagler *A distance constrained synapticplasticity model of C.elegans neuronal network,* Physica A: Statistical Mechanics and its Applications.

Appendix

`https://drive.google.com/drive/folders/1u_59F9We7-5wPLsZSlCvz2kT2oEM8dVL?usp=sharing`