

Deciphering the Organisational Principles of the *C. elegans* Connectome

TP Sanjan (BT11B036)
Shantanu Gupta (CS11B047)

December 1, 2014

1 Introduction

Over the course of evolution, biological systems tend to become more and more efficient and robust. Several examples can be cited in this regard. For example, polypeptides are known to achieve 3D conformation corresponding to the lowest energy 1 ; metabolic pathways such as glycolysis have evolved to meet cellular objectives (such as higher ATP and maximal negative Gibbs free energy of the overall pathway) better 2. Likewise, it has been hypothesised that neuronal connections are optimised to minimise the total length of fibre used. This is commonly referred to as the “save wire” principle 3.

C. elegans belongs to the phylum *Nematoda* and is characterised by a transparent, unsegmented body. It has about 302 neurons and the data pertaining to its connectome i.e. the connection of neurons is available. In this study, we have mathematically modelled the *C. elegans* neuronal network and attempted to justify the spatial distribution of neurons in the context of the “save wire” principle.

2 Objectives

- Topological analysis of the *C. elegans* connectome.
 - Degree distribution.
 - Motif detection.
- Justifying the spatial distribution of neurons based on optimal wire length.

3 Methodology

The connectivity information was obtained from the Dynamic Connectome Lab website [4-6]. The coordinates of the positions of neurons, sensory organs, body muscles were acquired from the Wormatlas database 7. For the ease of computation, the connectivity information was represented in the form of adjacency matrix A , such that

$$A_{ij} = \begin{cases} 1 & \text{if neurons } i \text{ and } j \text{ are connected via synapse} \\ 0 & \text{otherwise} \end{cases}$$

The dataset acquired from the website linked below was parsed using Python to generate adjacency matrices and position maps in the form of `csv` files.

1. Using this adjacency matrix, degree distribution plots were generated, and were further analysed for their fit to power law graph models. For a power law degree distribution,

$$P(k) \propto k^{-\gamma} \quad \text{for } k > k_{min}$$

where

- k : degree
- $P(k)$: fraction of nodes having degree k
- k_{min} : minimum value of k above which power law behaviour is exhibited

Typically, the degree exponent γ lies between -3 and -2 for a strongly power law network.

2. The *C. elegans* connectivity network was compared with 100 random graphs to detect over represented 3-node patterns commonly known as motifs. For this purpose, the **mfinder** tool 8 was used.
3. Finally, the spatial distribution of the *C. elegans* neuronal system was obtained based on the component placement optimisation problem formulated with the following objective function:

$$C = \sum_{\substack{i,j \\ \text{connected}}} ||\mathbf{r}_i - \mathbf{r}_j||^2 + \sum_{\substack{\text{neuron-} \\ \text{periphery} \\ \text{pairs}}} d_{\text{neuron-periphery}}$$

which sums over all inter-neuron, as well as neuro-muscular and neuro-sensory connection lengths.

As the typical dimensions of *C. elegans* is around 1 mm in length, with a width of the order of $10\text{ }\mu\text{m}$. The worm was modelled as a rectangular box having 20 partitions along the length and 2 along the width (in accordance with the bilateral symmetry of the organism). Neurons were assumed to be non-overlapping, and to have a diameter of $1\text{ }\mu\text{m}$. The optimisation problem was solved along with appropriate constraints (explained in results and discussion) using the simulated annealing method. The codes were written in MATLAB.

4 Results and Discussion

4.1 Topological analysis of *C. elegans* connectome

The in-degree as well as out-degree distribution (Fig 1) of the *C. elegans* connectome showed a power law like distribution.

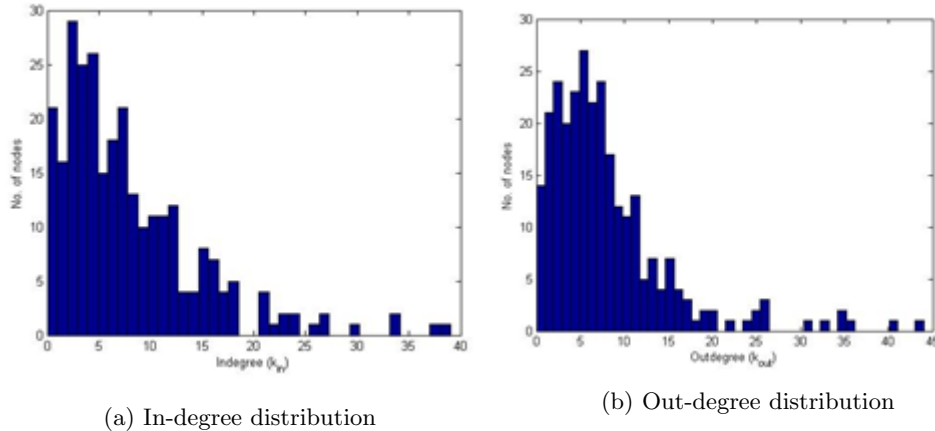


Figure 1: Degree distributions

The power law exponent was calculated using MATLAB's Curve-Fitting Toolbox : **cftool**. The value of the exponent γ was found to be -0.9015 (R^2 value 0.8264) and -1.683 (R^2 value 0.9215) with k_{min} of 2

and 5 for in-degree and out-degree distribution respectively. This suggests that the neuronal connectivity network has a high number of low degree nodes and a small number of high degree nodes which are known as hubs. This is consistent with observations across many real world networks such as social networks and biological networks.

4.2 Motif Detection

Among the 13 possible 3-node motifs for a directed graph, 6 subgraph patterns¹ were found to be overrepresented. Surprisingly, the most abundant motif was the feed forward network which is also commonly found in gene regulatory networks.

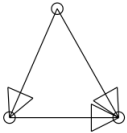
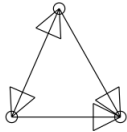
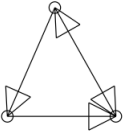
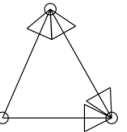
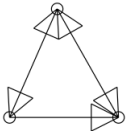
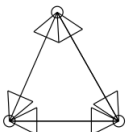
Motif detected	Frequency
	1521
	299
	170
	487
	139
	16

Table 1: Motifs and their frequencies in the *C. elegans* connectome

4.3 Spatial distribution of *C. elegans* neuronal system

In order to avoid the trivial case of all neurons clumping to one point the following constraints were imposed:

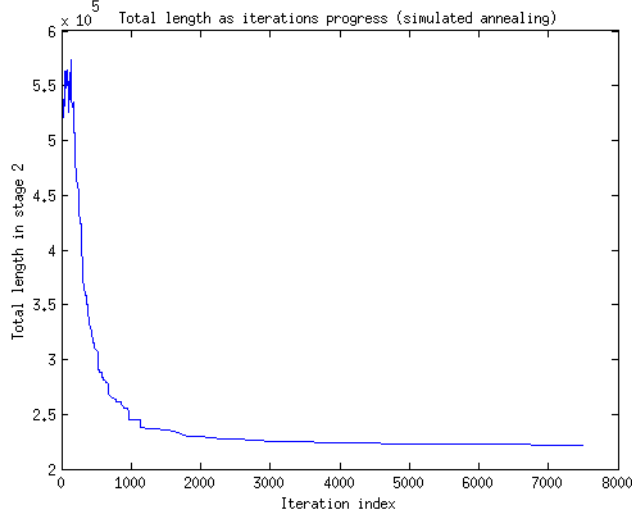


Figure 2: Total length as iterations increase

4.3.1 Minimum neurons per compartment

A minimum of one neuron in each compartment was imposed. Since the anterior region compartments would ideally have higher number of neurons due to the presence of sensory organs, a parameter h_{min} was used to find the optimal value. It was found that when h_{min} was in between 7 and 20, non-trivial (no clumping in one single box) steady state distribution was obtained. So, a minimum of 12 neurons was chosen for the 6 leftmost compartments.

In case the number of neurons in a compartment was above the minimum criteria, it was allowed to be relocated to any other box as dictated by the simulated annealing move generator. Otherwise, the neuron was forced to be relocated within the compartment by restricting the range of new positions coordinates generated for that particular neuron.

Simulated annealing parameters When an exponential cooling schedule was used instead of the linear schedule, the distribution was found to resemble the actual spatial distribution of neurons better. The cooling schedule used was:

$$T_k = \frac{T_0}{(1.1)^k}$$

where

T_0 : initial temperature

T_k : temperature at the k th iteration

One of the optimal distributions found in accordance to the above constraints and heuristics is shown below:

12	12	12	16	35	21	5	5	1	1	1	1	1	1	1	1	1	1	1
12	12	14	27	35	24	11	1	1	1	1	1	1	1	1	1	1	1	1

Table 2: Distribution with minimum-cell-occupancy constraints

4.3.2 With neuron-periphery connection data

Later, we used the coordinates of sensory organs and body muscles (referred to as sensors and effectors) in *C. elegans* obtained from Wormatlas database as prior edge constraints. The number of connections among the neurons is 2105 whereas the number of connections from neurons to sensors and neurons to effectors is 644. Moreover, only 199 neurons have connections to the fixed points sensors and effectors. So, to avoid the trivial distribution, we performed the neuron placement in two stages. The neurons having connections with fixed points were placed initially 3. Later, the remaining neurons which don't have any connections to fixed points were placed 4. The variation of the total fibre length as iterations progress can be seen in 2.

0	3	25	17	4	3	3	5	15	1	8	4	3	3	2	3	1	2	0	0
1	3	16	14	10	6	4	10	2	4	8	4	4	3	2	5	1	0	0	0

Table 3: Distribution after placing the peripherally-connected neurons alone

0	3	27	30	10	5	9	9	15	1	8	4	3	3	2	3	1	2	0	0
1	3	19	22	17	7	24	15	3	4	8	4	4	3	2	5	1	0	0	0

Table 4: Final distribution after placing the remaining neurons as well.

The final distribution obtained at the end of stage two shows majority of neurons being located at the anterior end which is consistent with the well-known fact of brain(which contains majority of the neurons) being located at the frontal part of an organism. Although *C. elegans* doesn't have a well-defined organ called brain, most of the neurons are found to be located just behind the pharynx.

For comparison, the actual distribution of the neurons from the original dataset is at 5:

0	8	30	29	9	0	2	0	0	1	0	1	1	1	0	1	13	0	0	0
0	13	37	23	16	7	8	7	4	7	7	7	7	8	6	4	22	0	0	0

Table 5: Actual neuron distribution

5 Conclusions

Topological analysis The degree distribution analysis of *C. elegans* connectivity network indicated power law like behaviour. The motif detection analysis revealed that 6 of the 13 possible 3-node motifs are present in the *C. elegans* connectome with feed forward loop being the most abundant.

Save wire This computational study validates the fact that component placement optimisation (CPO), also known as save wire principle is one of the primary driving force in organisation of neuronal system. The study carried out at individual neuron level shows that majority of the neurons are situated in the anterior end consistent with the brain's position. In future, with the availability of connectome data of other organisms, it would be good to see if the CPO wins over other competing factors such as information processing, functional roles of the ganglia, electrophysiological mechanisms etc.

6 References

1. Kuhlmann B. *et al* (2003) Design of a Novel Globular Protein Fold with Atomic-Level Accuracy. *Science* 302, 1364 doi:10.1126/science.1089427

2. Waddell TG. *et al* (1997) Optimization of Glycolysis : A New Look at the Efficiency of Energy Coupling. *Biochemical Education* 25 (4) : 204-5
3. Cherniak C. (1994) Component Placement Optimization in the Brain. *The Journal of Neuroscience* 14 (4) : 2418-2427
4. Kaiser M, Hilgetag CC (2006) Non-Optimal Component Placement, but Short Processing Paths, due to Long-Distance Projections in Neural Systems. *PLoS Computational Biology* 2:e95
5. Ktter R (2004) Online retrieval, processing, and visualization of primate connectivity data from the CoCoMac database. *Neuroinformatics* 2:127-144.
6. Choe Y, McCormick BH, Koh W (2004) Network connectivity analysis on the temporally augmented C. elegans web: A pilot study. *Society of Neuroscience Abstracts* 30:921.9.
7. Wormatlas Database : <http://www.wormatlas.org/neuronalwiring.html#NeuronalconnectivityII>
8. mFinder (tool): <http://wbs.weizmann.ac.il/mcb/UriAlon/download/network-motif-software>