CS 575 Project 3: Toward an Ecologically Valid Simulation

J. Wesley Borden

April 1, 2024

Abstract

This project continues a third iteration of building a network based simulation of neural firing in biological neuron networks. Advances include a simplified state machine that can be used to represent any part of a neuron (axon hillock, nodes of Ranvier, or axon terminal). Only one node, the axon hillock, is used to represent the neuron in a visualization, but the simulation uses intermediate notes to carry respective signals through the modeled neurons. This involved significant transfer of complexity from the neuron classes to the manager class. Additionally, graphs are built to be directed based on node metrics and weights are implemented as an input to each complex contagion calculation. These advances demonstrate the complexity of a future ecologically-valid simulation, and begin to point us toward details to be considered in building that model.

Introduction

Our most recent iteration of this project included an overview of the historical and biological significance of building a simulation to model neuronal firing activity in a mammalian nervous system. This iteration added several detailed advances to the simulation to identify node characteristics and model signal transmission through the network given varied inputs with respect to these measures. For example, signal transmission through the network when each edge was pointed toward the more central versus the less central node was considered. Dependent variables included generic graph metrics, as well as initially active nodes, edge direction based on centrality or degree of incident nodes. The simulation was built to handle edge weight and length, but these potential dependent variables were not tested during the experiments. Outcomes included continuous propagation of the signal through the network, steady state firing levels, and whether a steady state was reached or synchronous activity began. This last variable was not initially an outcome, as prior iterations indicated it had a low likelihood of being effected, but two experiments did show synchronous behavior, returning my interest to the outcome.

Hypotheses

This iteration of the simulation involved some iteration itself. Initially, I made most of the upgrades to the modules and tested the simulation with similar inputs to the project 2 inputs. In that case I hypothesized:

- "All networks will lead to self-sustaining activity"

This was a very conservative hypothesis because it was practically hard-coded: the new complex contagion calculation would include randomly firing with a 1% at each time step for an inactive neuron. This modification was consistent with biological examples as well as the finding from project 2 that a key aspect of continuously firing networks is occasional random (unstimulated) firing.
After confirming functionality of the simulation, I continued to build the experiments to handle more-flexibly crafted inputs. In the final set of 33 experiments, I hypothesized:

- "All networks continuously propagate" (restating from above)
- "Directed graphs have lower steady-state (Active)" In other words, there would be relatively less active nodes at the steady state for directed vs. bi-directional graphs.
- "Denser graphs have higher steady state (Active)" In other words, there would be relatively more active nodes at the steady state for more dense graphs.

After the experiments, I consider these hypotheses to be very conservative. I first intended to add these additional inputs (edge directionality, weight, length, intermediate nodes, and a new complex contagion algorithm based on edge weights) and demonstrate their use, and then I found that these inputs could be determined based on graph and node metrics.

Discussion

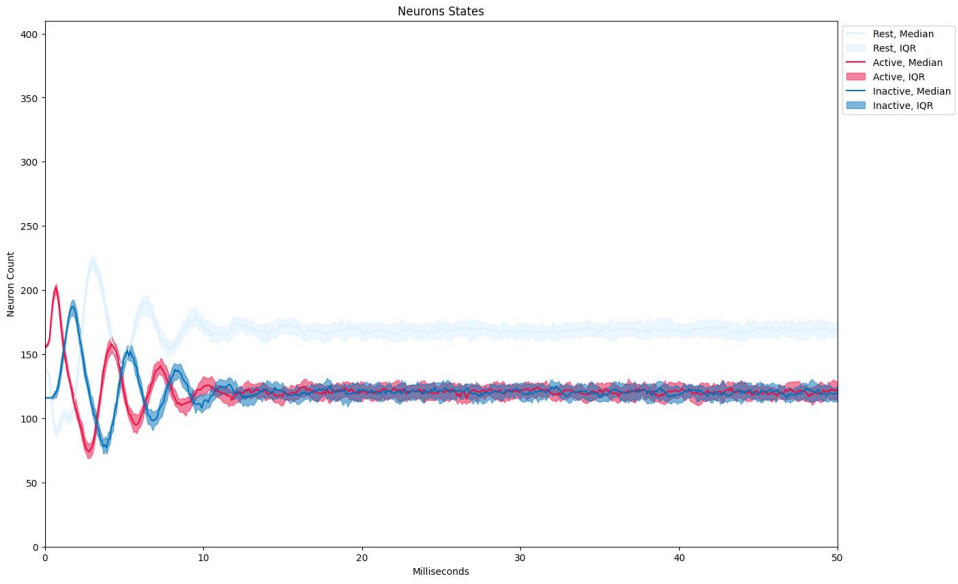
I performed many graph-specific tests of networks described in the project 2 specifications, as well as biological networks from the Network Data Repository (e.g., macaque-rhesus-brain-2). From my review of the available Network Data Repository documentation, I understand that these networks were obtained from electron microscopy imaging of tissue samples (method), and contributed to the Network Data Repository by Johns Hopkins Neurodata Laboratory, led by Dr. Joshua Vogelstein. Initially, all these graphs were simple, undirected, unweighted graphs. Before each simulation, graphs were programmatically updated to *nx.MultiDiGraph* objects, with node and edge attributes specifying initial node state, edge direction, edge weight, edge length, and inter-node distance. This function allowed specifying these attributes based on random chance, node centrality, and node degree.

Outputs were similar to outputs for my project 2. In addition, for each of the 33 experiments, a video of an example simulation was produced and is available upon request. Results from these outputs for the 33 most successful and most relevant experiments are summarized below:

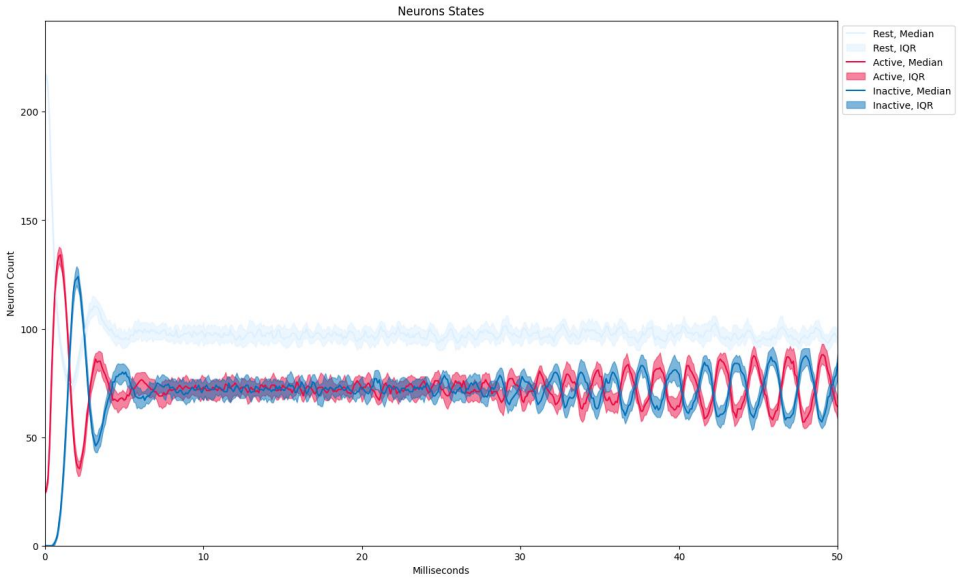
run_name	graph	initial_state_alg	direct	weight_alg	maximum_degree	average_degree	diameter	radius	density	louvain_modularity	Degree Distribution	Synchronous	Rest Median	Active Median	Rest Over Active
Circulant 20-12 1-10 Random	g_circ_20_12	given	random	random	4	4.00	NA	NA	0.11	0.82	Uniform	0	15	2	7.50
Circulant 20-12 1-2 All	g_circ_20_12	given	all	random	8	8.00	5	5	0.21	0.55	Uniform	0	10	4	2.50
Circulant 20-1234 1-2 All	g_circ_20_1234	given	all	random	16	16.00	3	3	0.42	0.34	Uniform	0	3	8	0.38
Circulant 20-1234 1-2-3-4 Random	g_circ_20_1234	given	random	random	8	8.00	5	3	0.21	0.34	Uniform	0	12	4	3.00
Dublin Highest-Degree FromC	g_infect_dublin	highest_degree	from_central	centrality	50	13.49	NA	NA	0.02	0.73	Right skew	0	100	150	0.67
Dublin Highest-Degree ToC	g_infect_dublin	highest_degree	to_central	centrality	50	13.49	NA	NA	0.02	0.73	Right skew	0	170	110	1.55
Dublin Random-Nodes Random	g_infect_dublin	random	random	random	50	13.49	NA	NA	0.02	0.71	Right skew	0	160	110	1.45
Figure 19-4 1-2 All	g_fig_19_4	given	all	random	10	6.82	7	4	0.21	0.54	Right skew	0	8	4	2.00
Figure 19-4 6-7-12 Random	g_fig_19_4	given	random	random	5	3.41	NA	NA	0.11	0.63	Right skew	0	13	2	6.50
Figure 19-4 7-8 Centrality	g_fig_19_4	given	to_central	centrality	5	3.41	NA	NA	0.11	0.46	Right skew	0	11	3	3.67
Karate 1-2 All	g_karate	given	all	random	34	9.18	5	3	0.14	0.50	Right skew	0	19	7	2.71
Karate 33-34 Random	g_karate	given	random	random	17	4.59	NA	NA	0.07	0.52	Right skew	0	25	4	6.25
Karate HighC FromC	g_karate	highest_centrality	from_central	centrality	17	4.59	NA	NA	0.07	0.50	Right skew	0	23	5	4.60
Karate HighC ToC	g_karate	highest_centrality	to_central	centrality	17	4.59	NA	NA	0.07	0.50	Right skew	0	23	6	3.83
Karate LowC FromC	g_karate	lowest_centrality	from_central	centrality	17	4.59	NA	NA	0.07	0.50	Right skew	0	23	6	3.83
Karate LowC ToC	g_karate	lowest_centrality	to_central	centrality	17	4.59	NA	NA	0.07	0.50	Right skew	0	23	6	3.83
Karate Random	g_karate	random	random	random	17	4.59	NA	NA	0.07	0.48	Right skew	0	25	3	8.33
Macaque HighC ToC	g_bio_macaque	highest_centrality	to_central	centrality	111	25.24	NA	NA	0.05	0.36	Right skew	1	100	60	1.67
Macaque LowC ToC	g_bio_macaque	lowest_centrality	to_central	centrality	111	25.24	NA	NA	0.05	0.36	Right skew	1	100	60	1.67
Macaque Random	g_bio_macaque	random	random	random	111	25.24	NA	NA	0.05	0.37	Right skew	0	100	60	1.67
MouseV1 HighC All	g_bio_mouse_v1	highest_centrality	all	random	62	4.44	8	4	0.01	0.81	Right skew	0	150	25	6.00
MouseV1 HighC FromC	g_bio_mouse_v1	highest_centrality	from_central	centrality	31	2.22	NA	NA	0.01	0.69	Right skew	0	150	25	6.00
MouseV1 HighC ToC	g_bio_mouse_v1	highest_centrality	to_central	centrality	31	2.22	NA	NA	0.01	0.69	Right skew	0	150	25	6.00
MouseV1 LowC All	g_bio_mouse_v1	lowest_centrality	all	random	62	4.44	8	4	0.01	0.79	Right skew	0	150	25	6.00
MouseV1 LowC FromC	g_bio_mouse_v1	lowest_centrality	from_central	centrality	31	2.22	NA	NA	0.01	0.69	Right skew	0	150	25	6.00
MouseV1 LowC ToC	g_bio_mouse_v1	lowest_centrality	to_central	centrality	31	2.22	NA	NA	0.01	0.69	Right skew	0	150	25	6.00
MouseV1 Random All	g_bio_mouse_v1	random	all	random	62	4.44	8	4	0.01	0.80	Right skew	0	150	25	6.00
Scale Free 100 Highest-Degree	g_scale_free_100	highest_degree	to_central	random	38	3.92	NA	NA	0.02	0.70	Right skew	0	70	20	3.50
Scale Free 100 Lowest-Degree	g_scale_free_100	lowest_degree	to_central	random	38	3.92	NA	NA	0.02	0.63	Right skew	0	70	20	3.50
Scale Free 100 Random-Nodes	g_scale_free_100	random	random	random	38	3.92	NA	NA	0.02	0.65	Right skew	0	80	10	8.00
Small World 100 Highest-Degree	g_small_world_100	highest_degree	to_central	random	6	4.00	NA	NA	0.02	0.76	Right skew	0	70	20	3.50
Small World 100 Lowest-Degree	g_small_world_100	lowest_degree	to_central	random	6	4.00	NA	NA	0.02	0.77	Right skew	0	70	20	3.50
Small World 100 Random-Nodes	g_small_world_100	random	random	random	6	4.00	NA	NA	0.02	0.77	Right skew	0	70	20	3.50

As anticipated, all these networks continuously propagated (circulation of the signal never ended) because of the increased sensitivity and excitability of each neuron. Directed (vs. bi-directional) networks had lower rates of active neurons in most cases, which followed my intuition that greater connectivity would lead to more activity. This finding also supports the hypothesis that more dense graphs would have a higher steady state–it's the same phenomenon. However, I realized after the experiments that there were too many structural differences between graphs (e.g., karate vs. small-world) to compare the outcomes across graphs.

In my first iteration of this simulation (Project 1), I attempted to model synchronous neural firing (similar to the synchronous firing in epileptic seizures, or sleep), but found that networks tended toward a steady state in the simulation. Here is an example of simulated states of nodes in the graph during simulation, showing how the network settles to a steady state:



This made sense in retrospect, as there was no structure to force the synchronous activity. Given this, I was surprised when one of the biologically-derived networks *did* show this synchronous activity:



This synchronous activity emerged after reaching a steady state, supporting the belief that the synchronicity was a result of graph structure rather than early adopters. In fact, the synchronicity was observed whether initially-active nodes were highest or lowest in centrality. It was also only observed in directed networks, where edges pointed from less central to more central nodes. The synchronicity was only observed in one of the biological networks, suggesting there may be a unique structure of this specific graph, such as the distribution of its k-cores that gives it this tendency toward synchronicity. Future work could test more similar networks, perhaps from varied parts of the brain, or with more consideration of edge directionality (e.g., away from the central nodes), to begin to better characterize what lets some networks tend toward synchronicity vs. a steady state.

Future Work

After a third iteration of this simulation, I have a greater appreciation of the complexity of agent-based modeling of networks, specifically with respect to cellular networks. There are many sub-cellular processes (e.g., genetic expression that leads to changes in behavior over time) that I can now see as a challenge to the ecological validity of a large-scale simulation. However, I believe that if combined with real-world data to determine input graph structure and node settings, network simulations might be improved. For example, could we build a simulation that parallels activity of a set of neurons, reading their current connectivity and activity, and then projecting an accurate anticipated firing pattern? While large-scale predictions may have challenges for feasibility, this could be useful for small circuits in key brain structures. In the future, I hope to integrate real-world electrophysiology data to build directed graphs with detailed node attributes to improve the ecological validity of the simulation.

In the nearer term, I also am considering how graph neural networks (GNNs) could better inform inputs to these simulations. As our class moves toward studying GNNs, I have a mind toward considering how the GNNs could be used to optimize the inputs to the simulation, or understand the attributes of nodes in the network.

Addendum

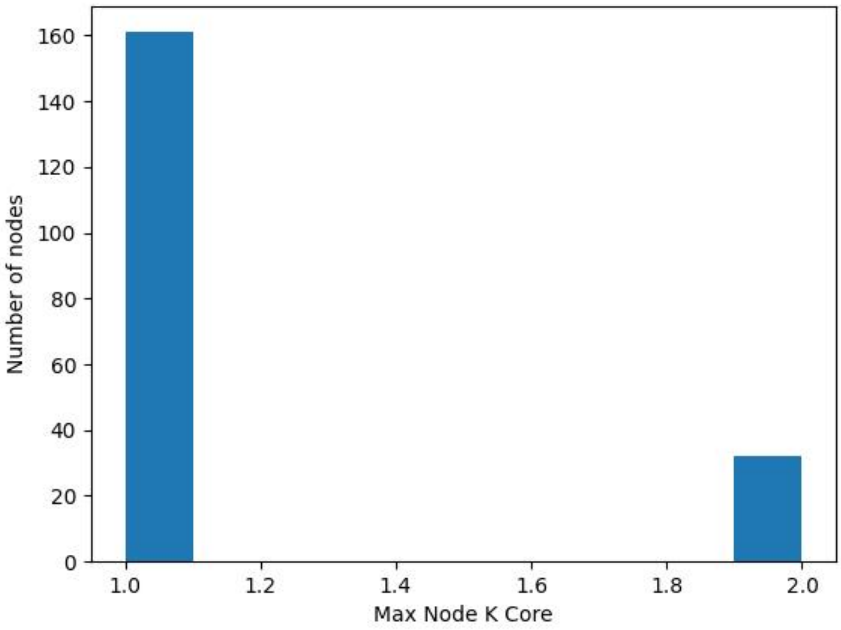
In reviewing these results and conclusions with Dr. Goodrich, we wanted to better characterize why it is that synchronous activity was observed in the macaque tissue model, but not in other graphs, which tended toward equilibrium. An intensive examination of all possible factors would be beyond the scope of this project, but we hypothesized:

- "The difference in synchronous activity is associated with the distribution of centrality, and of k-cores"

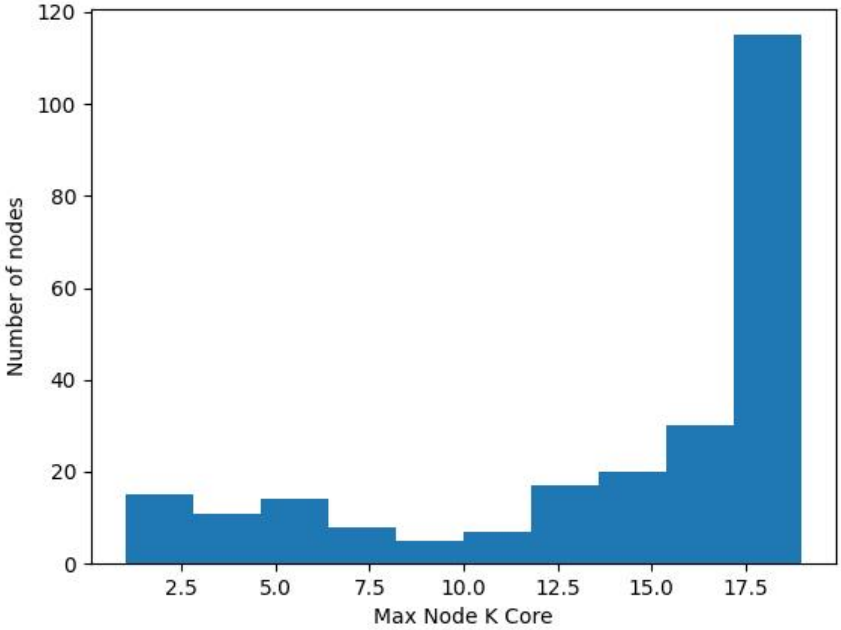
In other words, I anticipated that the synchronous activity was being caused by network shapes that include a central group with a distribution of peripheral agents. I had not originally measured or tracked a distribution of centrality or of core-periphery metrics, but made these measurements for graphs constructed with equivalent methods.

Core-Periphery Structures

This additional analysis revealed that nearly all graphs used in this and prior projects did not have remarkable core-periphery structure. For example, even a network measured with electron microscopy of mouse visual cortex tissue did not have a very strong core-periphery organization:



Unlike other input graphs, the macaque network where synchronous activity was observed did have a very strong core-periphery structure:

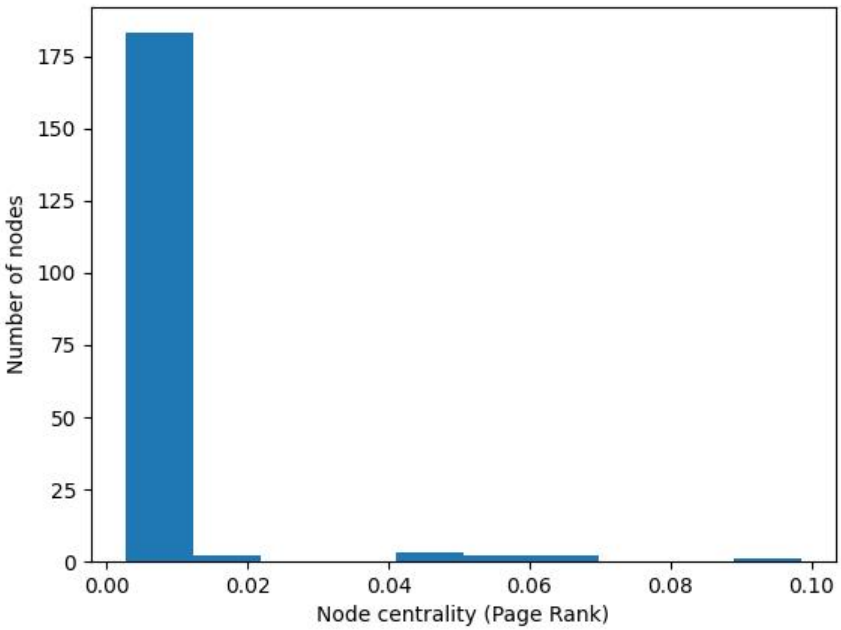


This evidence strongly supports the claim that core-periphery structure is key for synchronous signaling in biological neuronal networks.

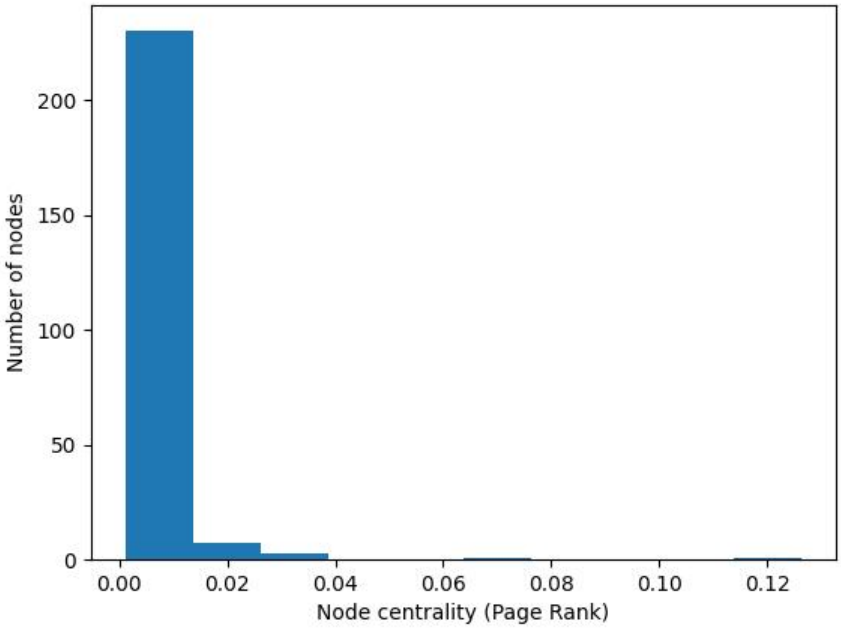
Centrality

Besides core-periphery structure, node centrality (page rank) was considered. Contrary to the hypothesis, neither networks demonstrating synchronous activity nor networks settling to a steady state had widely distributed centrality distributions. This suggests that while a core-periphery structure is key for synchronous activity, characteristics of centrality are less important.

The distribution of node centrality for a network based on mouse visual cortex tissue, which did not produce synchronous activity:



The distribution of node centrality for a network based on macaque brain tissue, which did produce synchronous activity:



LLM Acknowledgement

This report was generated using LaTeX, adapted from prior LaTeXfiles which had been written with the help of LLM's such as ChatGPT for formatting, etc. ChatGPT provided LaTeXformatting advice to questions such as "could you help me fix the format of this snippet? ...[LaTeXsnippet] ... This csv has this content: ...[csv file contents] ... And I'm getting the error ... [TeXcompiler error]". LLM's (ChatGPT, Google Search AI Overview) may have provided information for the development of code during the project (e.g., to answer questions about function call syntax), but did not directly write the code. Finally, ChatGPT may have been used to refine the report's tone, directly updating the text of the report as specified in the project description. My work reflects my best efforts at learning the content and building a promising simulation while leveraging the LLM's for things like productivity, speed, and technical writing tone.