GA Final

Tuesday, May 7, 2019

8:58 PM

This is Sarah Kirby

For my final project, I did an analysis of human connectomes.

* 1. The Human Connectome Project is a project funded by NIH that aims to map all the structural and functional neuronal connections in the human brain.
     1. A connectome refers to this graph of neuronal connections.

* 1. The human brain has billions of neurons, so rather than map invididual neurons, the project is starting by mapping areas called "regions of interest". They have identified 1015 of these "grey matter" areas, and mapped the connections by analyzing MRI images.

BrainGraph.org is a separate project that converted these data sets to graphml. The graphml datasets are published at braingraph.org.

These are some images from the Human Connectome website.

* 1. So I mentioned the connectome graphs were created from MRIs. Each vertex corresponds to a Region of Interest (ROI), which are 1-1.5cm2 areas of grey matter.

The edges were created by analyzing the MRIs to detect fibers running between the ROIs.

The braingraph.org project assigned direction to some of the edges based on the hypothesis that are common to most people's connectomes were formed earlier in the brain's development, and less common edges were formed later. They proposed that the newer connections grew towards the older ones.

* 1. This is an example of one individual's connectome. You can clearly see the left and right hemisphere, with several connections between them.

* 1. One of the questions I wanted to answer was are these connectomes scale free? Biological networks are often cited as examples of scale-free networks. So I wanted to see if this data set supported the idea that neural structures are scale free.

The properties of scale free graphs are: a degree distribution that follows the power-law. So few nodes have a very high degree, and most of the nodes have a low degree.

The clustering coefficient distribution also follows a power-law distribution. Finally, the clustering coefficient decreases as node size increases.

These graphs below are examples of a power-law distribution. The left images are histograms, and the right is a log-log plot of the distribution. In a power-law distribution, a log-log plot should show a linear relationship. This example clearly shows a straight line.

* 1. First I looked at the degree distributions, starting with total degree. The histogram on the left shows the average total degree distribution. The data set I downloaded had 424 individual connectomes, so I averaged the degree for each node and plotted that distribution.
     1. The histogram has a long tail to the right, which is indicative of a power-law distribution. However, the log-log plot is included on the left and you can see it's not quite linear.

* 1. I was also interested in what these distributions looked like for just the left or right hemisphere, so I plotted those as well. The left hemisphere is on the left, and the right is on the right. You can see that these are both almost identical, and they match the distribution of the total network. They also make it clearer that the log-log relationship is not linear.

* 1. Next I looked at just the in-degree distribution. Again, this is averaged over all the connectomes.
     1. The in-degree does appear to be closer to a power-law distribution. It has an even longer tail in the histogram, and the log-log plot is closer to linear.
  2. Looking at each hemisphere, you can see that again the individual hemispheres have the same distribution as the total network.

* 1. Finally, I looked at the out degree distribution. This one is clearly not a power-law, with a small tail in the histogram and a log-log plot that is not linear.
  2. Again, each hemisphere has the same distribution as the total graph.
     1. One thing to keep in mind about the in-degree and out-degree is that the direction labeling was based on a hypothesis about neuron growth, so it isn't necessarily accurate. Not all edges were labelled, too, so we're looking at an incomplete data set.

* 1. So the degree distribution was close to power-law but not quite. Next I looked at the clustering coefficient distribution.
     1. You can see from the histogram and log-log plot that this does not follow a power-law.
  2. Looking at each hemisphere, you can see that their individual clustering coefficients also do not follow a power law.

* 1. The last property of scale free graphs was the relationship between degree and clustering coefficient.
     1. You can see from the graph on the left that clustering coefficient decreases as node degree increases, which is in line with what you would expect from a scale-free graph.
     2. Interestingly, if you plot this as log-log, you can see that there is a linear relationship, so the relationshio between clustering coefficient and degree follows a power law, even though they don't on their own.
        1. This is specifically a property of hierarchical scale-free graphs, although it is not a property of all scale-free graphs.

* 1. So to answer the question, are connectomes scale free?
     1. The degree distribution did not follow a power-law, although the in-degree was close.
     2. And the clustering coefficient didn't follow a power law.
     3. The clustering coefficient did decrease as node degree increased, and there was a power law relationship between clustering coefficient and degree, However, overall I don't think we can say with any confidence that the connectomes are scale free.

* 1. The next question I had was are the connectomes small-world? Some have claimed that neural networks are small-world, so just like with the scale-free claim, I wanted to see if these graphs back that up.
     1. The properties of a small world network are:
        1. Short average path lengths
        2. High clustering
     2. There are small world metrics that evaluate this by comparing a graph to a random graph and a lattice graph of similar size and degree.
        1. In order to perform this comparison, I generated a random graph and lattice graph using cytoscape's Randomizer plug-in.
        2. The table below shows the results of analysis of those graphs. I picked one connectome and generated an equivalent random graph and lattice graph. The lattice graph was created using the Wattzstrogatz model, with no reconnecting of nodes so you just get the lattice.

* 1. This is an image of the random graph that was generated. It has the same number of nodes and average degree as the original connectome.
  2. This is the lattice graph. It also has the same number of nodes as the original, and close to the same degree.

* 1. The two metrics I calculated are the small-world metric and the small-world index
     1. …explain…
     2. Based on these results, I would not say that this connectome is small-world.

* 1. To summarize…