Little Rock Food Web

Petar Petrov h11729352

The Little Rock food network is represented as a directed graph, consisting of 183 vertices(animals/plants) and 2494 edges (food chain connection between them). Each consumer (predator) is represented as a start vector and each source (prey) as end vector. The first thing, that I did, was to compute the graph properties and plot some of them in order to get a sense of the distribution and other interesting attributes of the graph. Fig.1 shows the total degree distribution. By looking at it and comparing it to the graph models studied in the previous lectures, I can assume, that this is a scale free network, but I will check on this further in the following lines. We can see that there is a high amount of low-degree vertices and low amount of high-degree vertices (hubs). The minimum degree is 1, the maximum – 108. Next the amount of connected components is checked and it equals one, which means that the entire network is connected. Furthermore, the vertex and edge connectivities were examined and the result for both is 0. If we use the function as.undirected() ,we get vertex and edge connectivity of 1, but as our graph is directed and connected, we can assume, that it is only weakly connected. This means, that we are not able reach every vector, starting from every other by following the direction of the edges. As a next step, the diameter is checked. It equals to 6, which means, that the longest shortest path between 2 distinct vertices is 6 steps. The average path length is equal to 1.898, which means, that on average animal A can reach animal B in under 2 steps. The transitivity is then computed, it is equal to 0.332, which is a rather high value, when it comes for a scale free graph, but it rather resembles more the small world graph structure. This means, that there is a good number of clusters, with vertices connected to each other and then to some hub probably. However, based on the previous assumption and the other properties, like APL, degree distribution and diameter, I would still suggest, that this is a scale-free network, but with high transitivity.

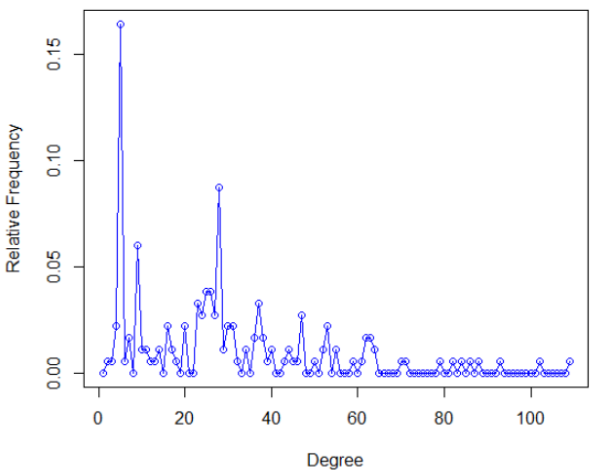

Fig. 1. Degree distribution

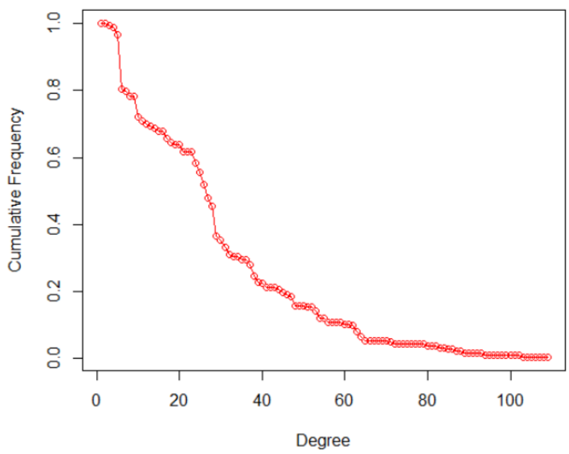

Fig. 2. – Cumulative Degree Distribution

Now, that I’ve completed the initial graph description and analysis, I will start to remove vectors, based on their centrality score, using different centrality measures. I will remove the vector with highest centrality score from the initial graph, then recompute it for the remaining graph, remove it again and repeat it one more time, so at the end, the three vectors with highest centrality will be removed from the graph.

I am going to start with the degree centrality. Degree centrality measures how many neighbors a vertex has. The one with the highest centrality score, is actually the vector with maximum edges. As the graph is directed, the in-degree and the out-degree are also relevant, aside from the total degree, but as they show similar results, I will provide numbers and explanations about the total degree centrality. As expected, the first thing we notice is that the max degree of the graph is now 107(down from 108), which means that there were two main hubs with 108 and 107 edges each. On the next steps, the max degree vector is reduced more, and after the third removal is equal to 86. The other properties of the graph don’t show significant differences – the diameter, the transitivity and the APL remain almost the same, which means, that this removal method doesn’t have a great effect on them, only on the max degree of the graph.

Next comes the betweenness centrality removal. Betweenness centrality measures the extent to which a vertex lies on the shortest paths between other vertices. A vertex with high centrality score has big influence on the information (food) flow within a network. If a vector with high betweenness centrality is removed the shortest paths will become longer and it is even possible to disconnect the network. Here actually, the opposite thing happens. After the third removal, we can observe, that the APL and the diameter are reduced, down to 1,53 and 4 respectively. The transitivity, min and max degree remain the same.

I continue with the closeness centrality, which refers to how close a vertex is to all the others. The mean distance to all the other vertices is calculated and the vertex with the minimum mean distance, has the largest closeness centrality. Usually those types of vertices in a network should be used as a source, for example in social networks, they can quickly spread rumors and news across the network. Here we can observe the first really significant change in the graph structure – it has split into 2 parts(clusters) after the second removal (i.e. became disconnected). The max degree, diameter, apl and transitivity don’t change much. But we can see a change in the min degree of the network – it has dropped from 1 to 0. This means, that when I removed the vector with highest closeness centrality, it was directly connected to some border vertex with a degree of 1 and became disconnected.

Eigenvector centrality is described as: A node is important, if it is linked to other important nodes. This centrality measure doesn’t care about the degree of the node measured, but rather for the degree of its neighbors. Even if a node has low degree, but is connected to few high degree nodes, it may have a large eigenvector centrality score and be able to have impact on the network. It is used mainly with undirected graphs or with strongly connected digraphs, but here R gives a solution for this one, so it means, that it is not weakly connected, as I assumed above. We can observe, that with each subsequent removal, the max degree vertex decreases by one, which means, that the vector, that has been removed, has been connected to the vector with maximum edges. The other values remain the same.

Next, an improvement on Eigenvector centrality measure is the Katz centrality. It addresses the problem, that eigenvector centrality has with weakly connected digraphs and the spider traps, that may exist in them. It introduces an attenuation factor, which is between 0 and 1 and contributes with the required weight, based on the distance between the vertices. Longer paths will have higher score, which will dampen the effect of the spider trap, that may exist in a network. From the results, I can see, that the diameter has been reduced to 5, the mean distance and the transitivity stay the same. The max degree has been reduced slightly, which means, that nodes connected to the max degree node were removed. This is expected, because, as I mentioned earlier, the higher the degree of the neighbors, the higher the eigenvector centrality score, and it is highly likely, that the vector with the highest ev centrality score is connected to the hub.

The last centrality measure observed, is the page rank centrality. It is an upgrade to the Katz centrality, because it considers also the out degree of the neighbors of a specific vertex. Another dampening factor is introduced, such as, when a vertex is connected to a high centrality vertex and is one among many, it should be less important for the page rank centrality of the observed vertex. After removing the vertices, the results for my network are pretty similar with the Katz centrality ones: Slightly reduced max degree, diameter down to 5 and the diameter, apl and transitivity stay pretty much unchanged.

As a conclusion, I will make brief summary: The removal of vectors with the highest closeness centrality and degree centrality has changed the structure of the graph most significantly by disconnecting it and removing the hubs. The other methods(Eigenvector, Katz and Page Rank) don’t have such a high influence, rather a small one on the diameter. This is probably due to the small number of vectors, that have been removed in comparison to the overall number of vectors. The transitivity and average path length scores don’t differ too much between all the models.