ECOLOGICAL NETWORK ANALYSIS
COMPLEX NETWORKS PROJECT

SULAY SHAH (201601406) & ANERI PATEL (201601445)

November 18, 2019

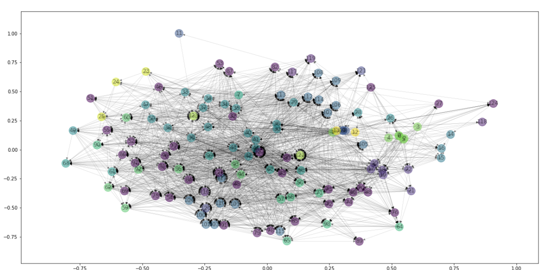
1 Introduction

Networks provide one of the best representations for ecological communities, composed of many species with sometimes complex connections between them. In addition to how many species are there, and which species are in a local area, knowledge of their interactions is indeed an additional layer of information that network measures exploit to quantify biodiversity.
In this report, we analyze a food web network. Food web networks describe who eats whom among an assemblage of species. In food webs, species are the nodes and feeding interactions are the edges. The edges are directed flows of biomass and energy that occur when one species feeds on another.

1.1 Data

The food web graph used for this analytical study is a simple, unweighted, unipartite and directed. The network contains the carbon exchanges in the cypress wetlands of South Florida. In this network, nodes are compartments (organisms and species or sources of energy in the food web) and edges represent energy flow (A directed edge from `i` to `j` means that carbon is transferred from `i` to `j` i.e. organism j eats organism i).

In addition, vertices are partitioned in different types as follows: 1 - Living/producing compartment 2 - Other compartment 3 - Input 4 - Output 5 - Respiration. Included with this network data are group classification labels (derived from ATLSS). Examples groups include "Zooplankton Microfauna" and "Pelagic Fishes".



The general structure of the data is as follows:

1. Order of the network is 128 - this is a measure of the species richness of the network.

2. Size of the network is 2106 - this measure quantifies the total interactions in the network.

3. Connectance is 0.13 - this measure represents the proportion of established interactions relative to the possible number of interactions.

2 Problem Formulation

1. Understand the general structure of the food web network.

2. Observe patterns by clustering similar nodes - detecting communities and motifs in the food web

3. Highlight important species using various centralities

3 Observations and Conclusions

3.1 General structure of the food web

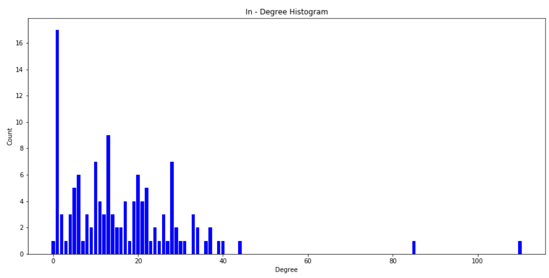


Figure 1: In - degree distribution

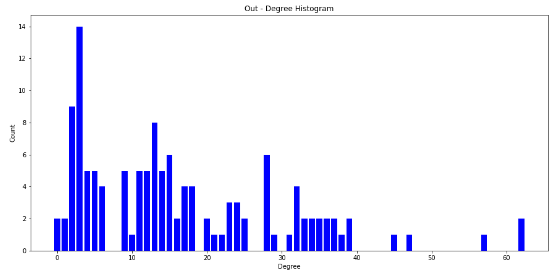


Figure 2: Out - degree distribution

In & out degree (k_{in} & k_{out}) - The degree of a node can give an idea of the species' vertical position in a food web— i.e., its trophic level. Species that do not consume any other species in the web (i.e., those nodes with an in-degree of zero) are primary producers. At the other extreme, species with no predators (i.e., those with an out-degree of zero) are top predators. Those with both predators and prey (i.e., non-zero in- and out-degrees) are intermediate consumers.
The out-degree and in-degree of each species represent correspondingly their *vulnerability* (number of predators) and *generality* (number of resources).
Both the distributions don't follow the power law.

3.2 Community Detection

Structures formed by groups of multiple nodes can be analyzed by looking at what types of relationships the nodes (representing a species) are typically embedded in (e.g.competition,intraguild predation - intraguild predation is the killing and sometimes eating of potential competitors. This interaction represents a combination of predation and competition, because both species rely on the same prey resources and also benefit from preying upon one another), by determining if there are nodes found in dense clusters or non-overlapping compartments, forming modular communities or using motif distribution.

Louvain algorithm which optimizes the modularity to form communities is used on this network to form communities. 4 communities are formed with varying size.

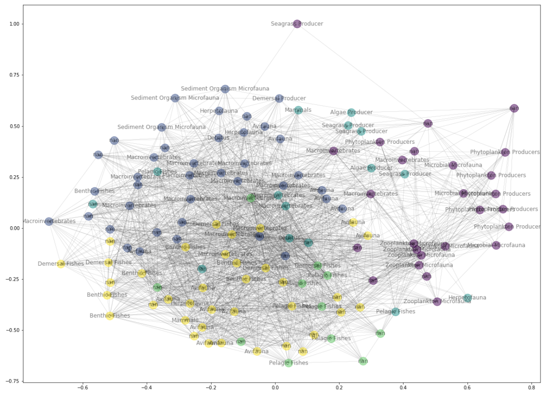


Figure 3: Louvain Algorithm

Interestingly, we don't get a clear distinction between each group through community detection. Many groups are evenly distributed in all the communities. This phenomenon presumably indicates that the roles of these species in the carbon exchange cannot be derived from the traditional divisions in a trivial manner. For example, though both are mammals, the manatee and the dolphin have very diverse diets. The manatee feeds on submergent aquatic vegetation, and the dolphin feeds on small fishes and shrimps. Consequently, one would expect that the manatee and the dolphin play different roles in the carbon exchange.

Community detection via Clique Conductance is one such algorithm which works well in this network and splits it into communities based on groups.

3.3 Motif Analysis

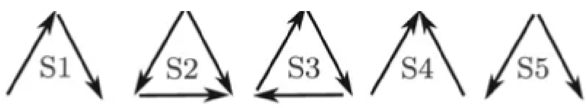


Figure 4: 3-Node Motif types

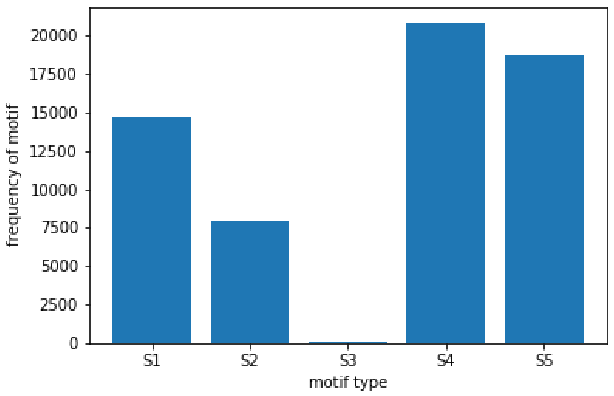


Figure 5: motif distribution

This shows that the network doesn't have many of the 3-node cycle motif type. The network will generally have cycles with path length more than 3. Also the 2^{nd} maximum motif type is of the competition kind.

3.4 Important species

Centralities are used to identify possible keystone species in ecological networks.

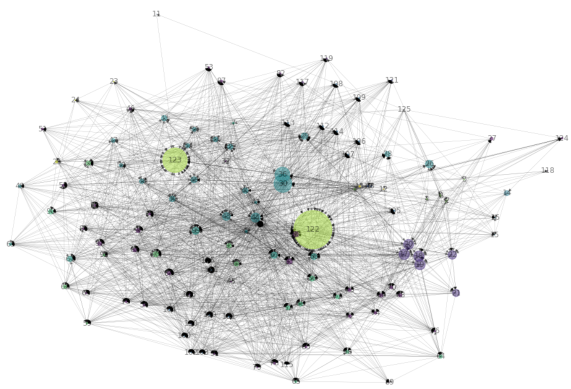


Figure 6: Betweenness Centrality

'Water POC' (node 122) has the highest betweenness centrality out of all the nodes. The second highest betweenness centrality is that of 'Benthic POC' (node 123). Both of these nodes belongs to the group 'Detritus'. Detrivores consume detritus and almost every organism gets converted into detritus. Therefore, due to high out-degree of these nodes their betweenness centrality is high.

This measure is ideal to study the influence of species loss on fragmentation processes (Population fragmentation occurs when groups of animals living in the wild become separated from other groups of the same species.It is often caused by habitat fragmentation, which as the name implies describes the emergence of discontinuous habitat (fragmentation) in the environment.)

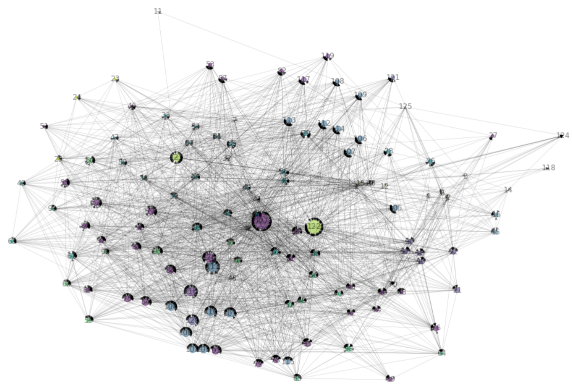


Figure 7: Eigenvector Centrality

As 'Respiration' is connected to all the living organisms, the out-degree is high and will be connected to important nodes which leads to the highest eigenvector centrality.
The second highest eigenvector centrality is that of 'Water POC' (node 122). This measure shows a similarity with the results of the betweenness centrality in this case.
The 3^{rd} highest C_E value is that of the raptor species, followed by Avifauna as the 4^{th}.

A species' eigenvector centrality is its position on the first axis of variation in the structure of the network. Species with extreme values of eigenvector centrality can therefore be viewed as strong contributors to the stability or instability of a food web.

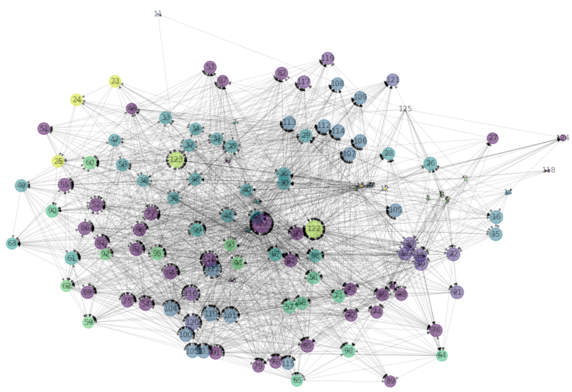


Figure 8: Closeness Centrality

Similarly, 'Respiration' has the highest closeness centrality of all the nodes. Therefore, respiration is the node

which will influence the whole network the most rapidly. Almost every living organisms gets carbon through respiration so it has the highest closeness centrality.
The second highest closeness centrality is that of 'Water POC' (node 122). This measure shows a similarity with the results of the betweenness and eigenvector centralities in this case.
The 3^{rd} highest C_C value is that of the Benthic node while the 4^{th} highest closeness centrality is that of a living species and not a source of energy - the dolphin.

4 References

- Robert E. Ulanowicz, Cristina Bondavalli, and Michael S. Egnotovich, 1997, Network analysis of trophic dynamics in South Florida ecosystems–The Florida Bay Ecosystem: Annual Report to the U.S. Geological Survey, Biological Resources Division.

- Austin R. Benson, David F. Gleich, and Jure Leskovec. "Higher-order Organization of Complex Networks." (2016). Science, 353.6295 (2016): 163–166.