Sanak Intertidal Food Web Dataset & Analysis

By Aasif, Sehbaan & Maksym

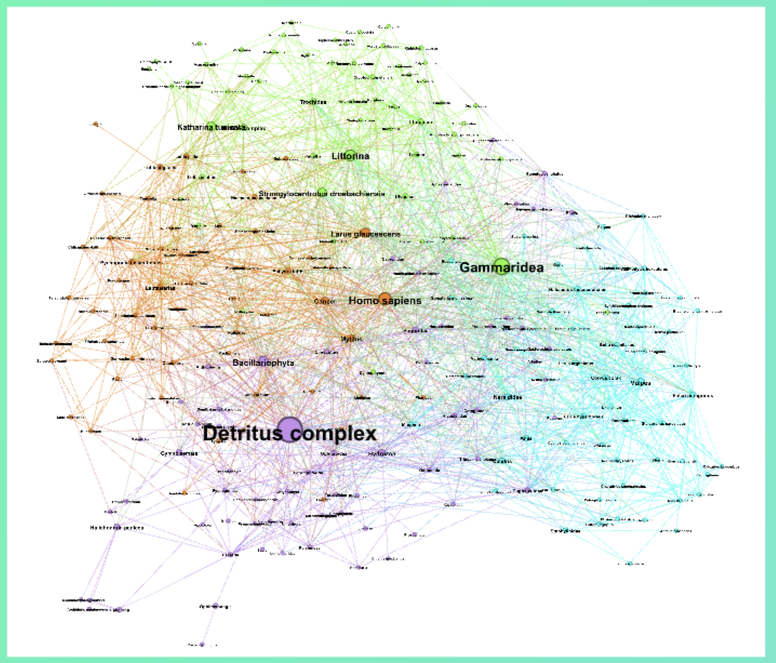
Project Brief & Proposal

The **Sanak Intertidal Food Web** illustrates a number of species and their **trophic links**.

- Trophic links are represented as an **edge** in the network, which indicates whether a species is part of a **trophic chain**.
- An **edge** signifies if the species consumes or is consumed by another.
- **Species** are represented as **nodes**.

The problem we wish to investigate is the **importance** of certain **species** within **communities** and the whole **ecosystem**.

The network is divided into communities using the **modularity class**. The network shows **4 communities** – the nodes within each community are more **densely connected** to each other than with nodes outside of their communities.



The Investigation

We wish to investigate the different **communities** and what kind of organisms are within each community (as divided by the clusters).

Additionally, we would like to use a second algorithm to judge the **centrality** of the network, to identify '**important**' nodes. From this, we wish to investigate what effect it would have on the network should these important nodes be **removed**.

Network Statistics

Number of Nodes: 235
- Number of species in the network.

Number of Edges: 1729
- Number of trophic links in the network.

Network Diameter: 5
- This is the **maximum distance** between all pairs of nodes in the network. A diameter of 5 indicates a trophic chain of no more than 5.

Average Path Length: 2.3589016184760867
- The **average number of steps** along shortest paths across all pairs of nodes with an edge between them.

Average Degree: 14.715
- A number indicating how many **neighbours** each node has, taken as an average across all nodes.

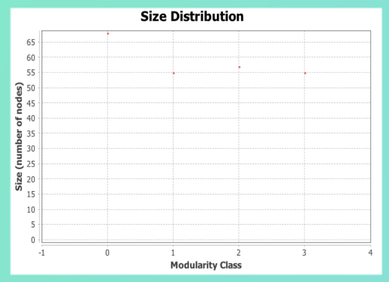
Average Clustering Coefficient: 0.22
- A measure of the **degree** to which nodes are **clustered** together.

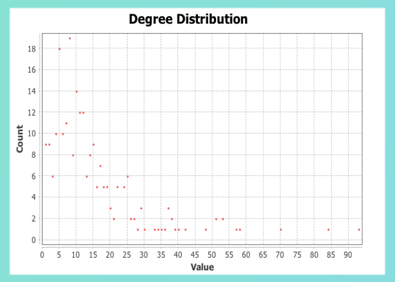
Modularity: 0.336
- A measure of which nodes are more **densely connected** together than to the rest of the network.

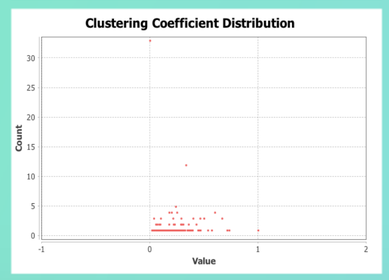
Network Information

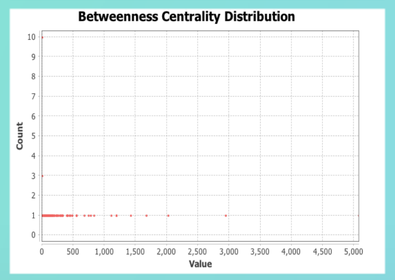
The **Homo sapiens** (or Humans) node size indicates that it has a **high Betweenness Centrality**, and the network also shows that it has **70 neighbours** – this demonstrates that Homo sapiens feed on 70 (**29.7%**) other taxa in the network.

They also have the 3rd shortest **mean path length**, which is the number of links that connects a species to another. This tells us that **94.4%** of the taxa were within **2 links** of humans.







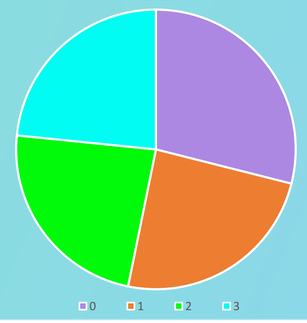


Algorithms

Algorithm 1: Girvan-Newman Community Detection
- This algorithm will be used to detect **communities** whose nodes are more densely connected, by removing edges.
- The goal is to illustrate the betweenness by focusing on edges that are most likely **between** communities.
- This will help to identify the **different communities** in our network for analysis and evaluation of any noticeable patterns.

Algorithm 2: Eigenvector Centrality
- This algorithm measures the **influence** or **importance** of a **node** in a network. A node is given a **higher score** depending on the scores of its **neighbours**. The score is based on **how many** neighbours a node has.
- This will determine which nodes in the network play a central role, and so a question such as, 'what would happen to the network **if these important nodes were removed**' can be investigated, and the resulting effects **analysed**.

Graph showing the Community Make-up in the Sanak Intertidal Food Web



Graph showing the distribution of the communities. Community constitutions – **0**: 28.94%, **1**: 23.4%, **2**: 24.26%, **3**: 23.4%

Evaluation

Important statistics, such as **clustering coefficient**, will be recalculated after a central node has been **removed**, to investigate how they have changed.
Analysis of the network can also help to predict what may happen if **new species** were **introduced** into the ecosystem.

Shortest path lengths (through **Eigenvector**) for nodes that are more **central** to the network will be **recalculated** to investigate the **impact** of the increased number of steps between species, resulting in species in **higher trophic levels** getting **less energy**. **Influential nodes** can be removed to investigate if there is a **breakdown** in the network. This could help to find out what would happen to the ecosystem if a species were to become **extinct**.