

Using the kcorebip package

Javier Garcia-Algarra

May 17, 2016

Introduction

This package performs the *k-core decomposition* analysis of a bipartite graph and provides two new kinds of plot: *polar* and *ziggurat*.

Decomposition and definition of k-magnitudes

The *k-decomposition* is an iterative algorithm that prunes links of nodes with degree equal or less than k . The process starts pruning nodes with $k = 1$ until all the resting nodes have two or more links. Then it continues with $k = 2$, and so on. After performing the *k-decomposition*, each species belongs to one of the *k-shells*. The highest value of k , i.e ks_{max} , corresponds to the innermost *core* $ks_{max} \equiv C^{A,B}$. For each *k-shell* there are two subsets, one per guild (A and B), that we call K_j^A, K_j^B where j is the *k-shell* index.

In order to quantify the distance from a node to the innermost shell of the partner guild, we calculate the average of the shortest paths to all the nodes in that set. We define the k_{radius} of node m of class A as the average distance to the species of C^B :

$$k_{radius}^A(m) = \frac{1}{|C^B|} \sum_{j \in C^B} dist_{mj} \quad m \in A \quad (1)$$

where $dist_{mj}$ is the shortest path from species m to each of the j species that belong to the set C^B . The same definition is valid for species of the guild B computing the average distance to C^A . The minimum possible radius value is 1 for one node of the maximum shell directly linked to each one of the maximum shell set of the opposite guild.

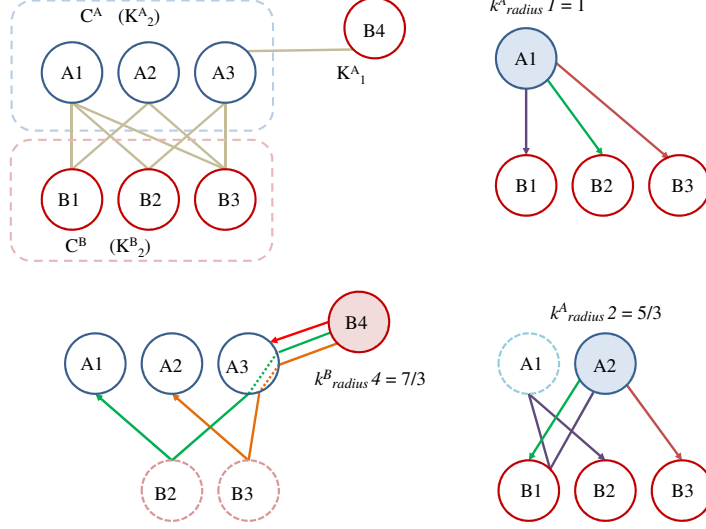


Figure 1: Examples of k_{radius} in a fictional network.

Figure 1 is a very simple fictional network, with only seven nodes. The upper left graph shows network structure. In the upper right figure, species $A1$ belongs to C^A . The distance to each of the nodes of C^B is 1, so its $k^A_{radius1}(1)$ is 1. In the bottom right figure, node $A2$ also belongs to C^A but there is not any direct link with $B2$, so the distance between them is 3 and $k_{radius}(2)$ is $\frac{5}{3}$. In the bottom left figure, node $B4$ is not part of C^B , and as may be expected, the value of its k_{radius} is higher.

A global value can be defined averaging this magnitude across network:

$$\bar{k}_{radius} = \frac{1}{|A \cup B|} \sum_{l \in A \cup B} k_{radius} l \quad (2)$$

In our example network of Figure 1, the value is $11/7$.

k_{radius} is a useful magnitude *to measure nestedness* but it not a good measure of centrality. For instance, its value for an isolated specialist linked to the maximum core is low. *To attend this necessity*, we define a second k -magnitude, the k_{degree} :

$$k_{degree}^A m = \sum_j \frac{a_{mj}}{k_{radius} j} \quad m \in A, \forall j \in B \quad (3)$$

where a_{mj} is the element of the interaction matrix that represents the link. So the $k_{degree} m$ is the sum of the inverse of k_{radius} for each node linked to m . A node of the innermost shell will have a high degree, whereas specialists have only one or two links and so a low k_{degree} . In the example of Figure 1, this magnitude is $1 + 3/5 + 3/5 = 11/5$ for node $B3$, while only $3/7$ for node $B4$.

Visualizations

The most common visualization of a mutualistic network is the bipartite graph. It underlines the existence of two communities, but links between species are difficult to distinguish even with a reduced number of nodes. It is an exercise of graphical intuition to identify the innermost shells and almost impossible to guess the centrality of species. As a result, it is uncommon finding graphs of mid-size networks in the research literature, although typical mutualistic networks are in the range of hundred species.

We have designed two new kind of visualizations for mutualistic networks, using k_{radius} and k_{degree} . We call the first one *Polar Diagram*. Its purpose is showing species centrality and having a quick overview of network distribution. It was inspired by the *fingerprint-like* graph, developed by Alvarez-Hamelin *et al.* [AH+05] to plot very large *k-decomposed* networks.

Species are depicted with their centers located at k_{radius} . Angles are assigned at random by the visualization algorithm to reduce overlapping, with the condition that each guild lies inside one of the half planes. Shape underlines the guild separation. The area of each node is proportional to its k_{degree} and the color represents the k_{shell} . This visualization does not include links and does not perform any kind of information reduction. The user may choose adding the histograms of *k-magnitudes*, a handy option because they convey a wealth of structural information.

Polar graph is specially useful when comparing different networks side by side, even if they are of very different sizes. In Figure 2, we have picked three networks located in both edges and in the middle of Figure ?? (right). The pollinator network *PL_010* has a low value of *Modularity*, 0.25, and high \bar{k}_{degree} , 4.57. Nodes appear unlabeled, except the three ones with smaller

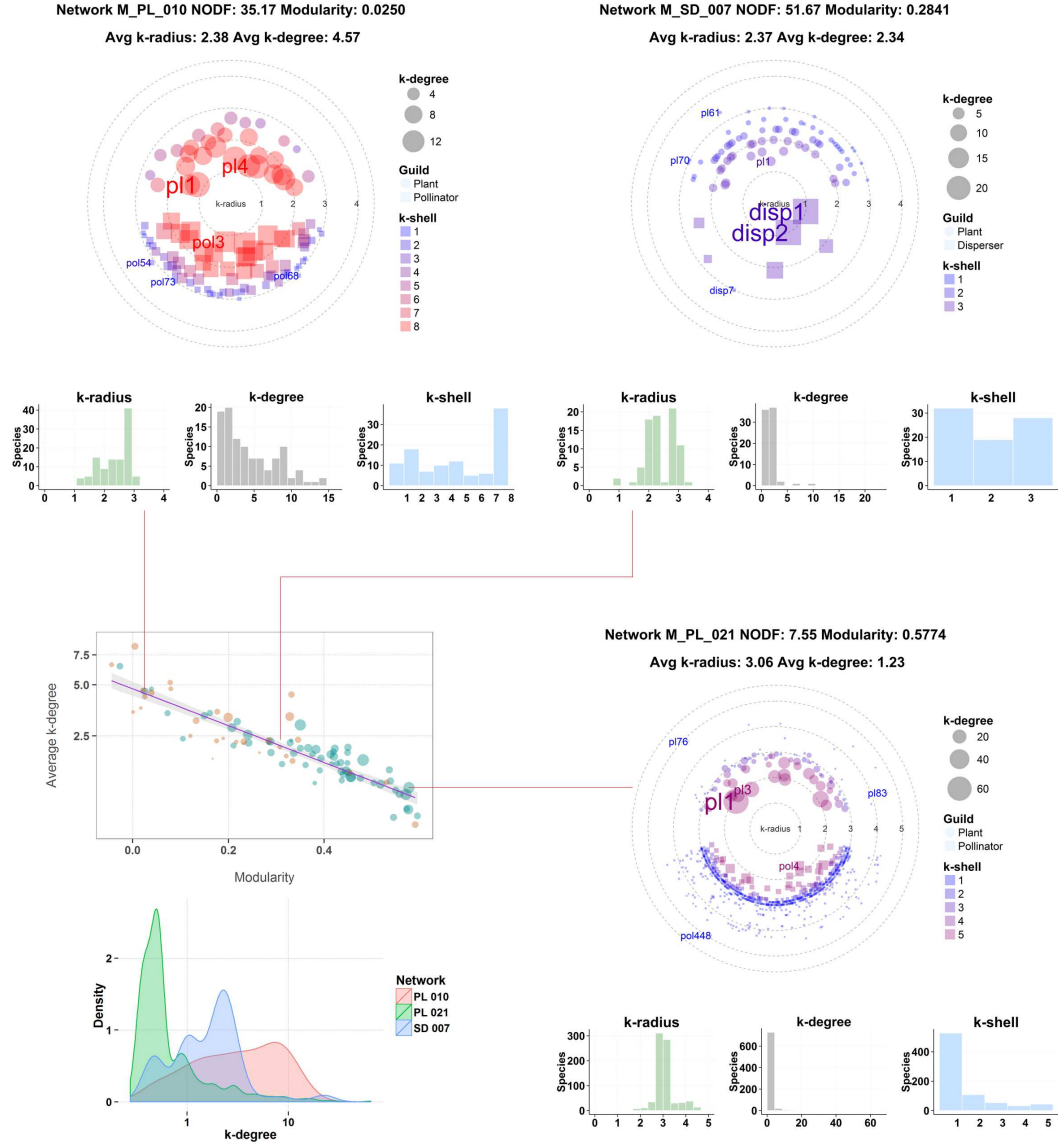


Figure 2: Polar diagrams of three mutualistic networks. The arrows show their positions in the plot $\bar{k}_{degree}/Modularity$. On the lower left corner, \bar{k}_{degree} density plots with logarithmic scale on the x axis.

k_{radius} and the three ones more distant from the innermost core. The polar

graph shows the onion like organization and asymmetry of mutualism.

Modularity (0.28) and \bar{k}_{degree} (2.34) lie in the middle range for the disperser network *SD_007* [Cro75]. The \bar{k}_{degree} distribution is sharper with two very dominant species. There is a sense of being more hierarchical than *PL_010*.

The network *PL_021* [Kat+90] is big and strongly modular, with *Modularity* 0.57 and \bar{k}_{degree} 1.23. The \bar{k}_{degree} distribution is even sharper. We have included the density distributions in a separate plot that shows how they become flatter as *Modularity* decreases and this behavior explains the negative correlation.

Ziggurat diagram is the second type of visualization, created from scratch. The idea behind it is splitting species in sets by their *k-shells*. Each of these groups are represented as small ziggurats. The maximum *k-shell* is located on the left side, the other ones are arranged following an almond distribution. Species of this shell are ordered by their k_{degree} with the higher one leftmost. Areas do not convey meaning, their height decreases just by convenience of visualization.

Nodes of 1-*shell* are scattered around the ziggurats. When multiple species of this shell are connected to the same node of any of the ziggurats they are clustered to reduce the number of lines. For instance, in Figure 3 plants species 22, 30 and 33 are specialists linked to the generalist disperser 3. With this organization, we get a clear view of structure and interconnections. The almond shape leaves a wide space in the center of the graph to depict the links and they do not overcross the boxes of the different species.

The network of Figure 3 has 95 links. It is still possible to see details in the bipartite graph, as nodes highly connected. With a higher number of interactions it becomes a *spaghetti dish* mess. The ziggurat gives a clear glimpse of the internal organization with four shells and a tiny external layer. Patterns may be visually discovered, for instance the low number of links between species of lower shells, or the relative low centrality of disperser species number 2 that appears leftmost in the bipartite graph.

References

- [AH+05] José Ignacio Alvarez-Hamelin et al. “k-core decomposition: A tool for the visualization of large scale networks”. In: *arXiv preprint cs/0504107* (2005).

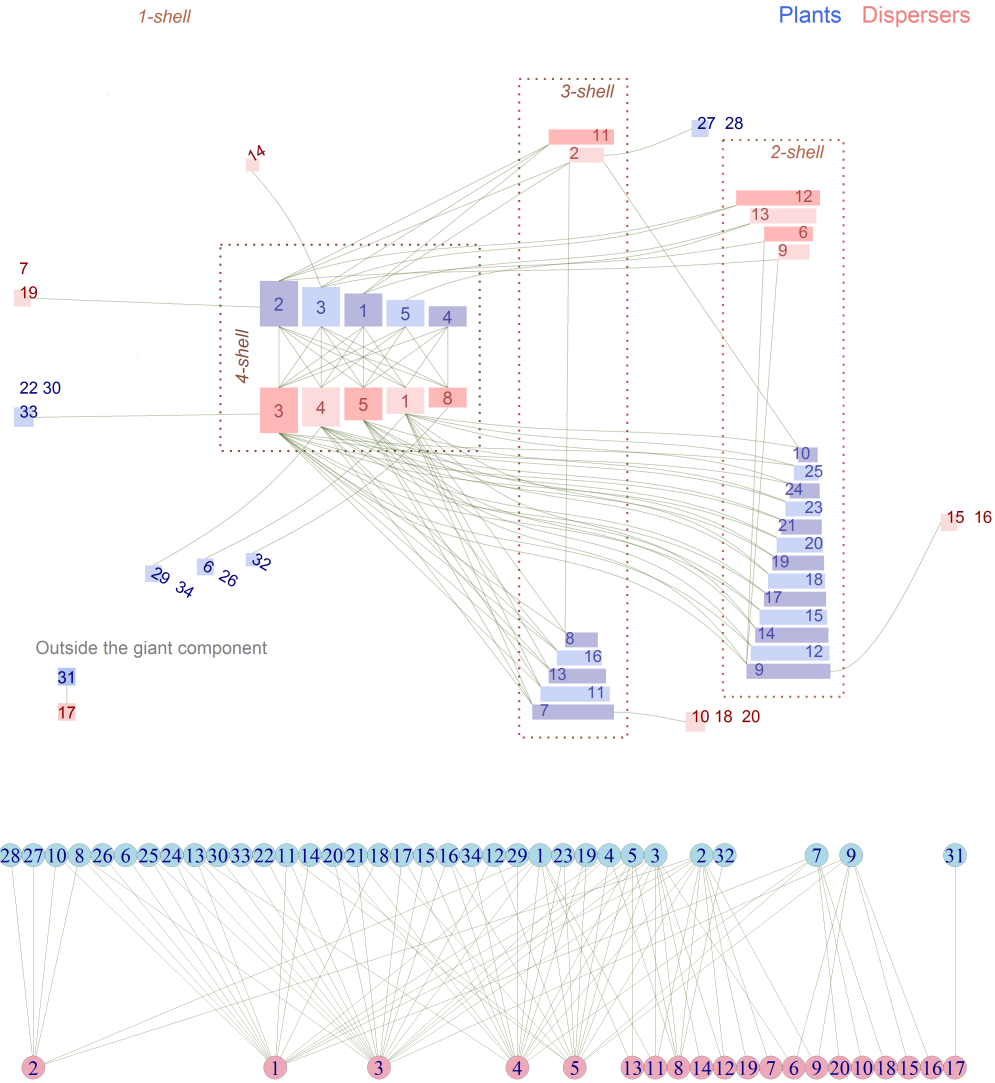


Figure 3: Ziggurat graph of an avian frugivore community in Puerto Rico with 54 species and 95 links[CCG03]. Below, the classical bipartite graph of the same network.

[Bas+09] Ugo Bastolla et al. “The architecture of mutualistic networks minimizes competition and increases biodiversity”. In: *Nature* 458.7241 (2009), pp. 1018–1020.

- [CCG03] Tomás A Carlo, Jaime A Collazo, and Martha J Groom. “Avian fruit preferences across a Puerto Rican forested landscape: pattern consistency and implications for seed removal”. In: *Oecologia* 134.1 (2003), pp. 119–131.
- [Cro75] FHJ Crome. “The ecology of fruit pigeons in tropical Northern Queensland”. In: *Wildlife Research* 2.2 (1975), pp. 155–185.
- [Kat+90] Makoto Kato et al. “Insect-flower relationship in the primary beech forest of Ashu, Kyoto: an overview of the flowering phenology and the seasonal pattern of insect visits”. In: (1990).
- [LM05] Vito Latora and Massimo Marchiori. “Vulnerability and protection of infrastructure networks”. In: *Physical Review E* 71.1 (2005), p. 015103.