Structure Comparison of Binary and Weighted Niche-Overlap Graphs

Nayla Sokhn[1,2], Richard Baltensperger[2], Louis-Felix Bersier[1], Ulrich Ultes-Nitsche[1] and Jean Hennebert[1,2]

[1] University of Fribourg, CH 1700 Fribourg, Switzerland
[2] University of Applied Sciences of Western Switzerland, CH 1700 Fribourg, Switzerland

Abstract. In ecological networks, niche-overlap graphs are considered as complex systems. They represent the competition between two predators that share common resources. The purpose of this paper is to investigate the structural properties of these graphs considered as weighted networks and compare their measures with the ones calculated for the binary networks. To conduct this study, we select four classical network measures : the degree of nodes, the clustering coefficient, the assortativity, and the betweenness centrality. These measures were used to analyse different type of networks such as social networks, biological networks, world wide web, etc. Interestingly, we identify significant differences between the structure of the binary and the weighted niche-overlap graphs. This study indicates that weight information reveals different features that may provide other implications on the dynamics of these networks.

Keywords: Network Measures, Weighted Networks, Food-webs, Niche-Overlap Graphs.

1 Introduction

Complex systems have recently gained much interest. Many analyses have been conducted to understand the structure of these systems and to uncover their unique patterns [1, 2, 3]. Networks have emerged across many fields including biology, ecology, social networks [4, 5, 6] and many others. All these different networks were found to have a special architecture and a particular behavior. It was shown that social networks belong to the small word property [7], known as the «six degrees of separation» phenomena. Food-webs and niche-overlap graphs turned out to follow a single scale exponential distribution [8, 9] while other networks such as the biology cells and the World Wide Web were found to follow a scale-free power law distribution [10]. In these real systems, two entities are connected if there is a relationship between them. For instance, in a social network, the relationship would be «being a friend with», in a food-web «feeding on a species», in a niche-overlap graph «competition between species». However, in order to have a better understanding of these networks, it is important to quantify the relationship between nodes. This is done by giving a weight to the links of the network. For example, in the scientific collaborator

network, the weight is equal to the number of coauthored papers between two authors. For the world wide web network, the weight is defined by the load of data transferred between two hosts [11]. For niche-overlap graph, the weight is characterized by the number of common prey between two predators. In order to analyse weighted networks, researchers generalized some network measures by considering the weight of the links [12, 13]. Here, our aim is to first investigate the structure of weighted niche-overlap graph using four classical metrics: node degree, clustering coefficient, assortativity and betweenness centrality. We then compare the results with the ones obtained by analysing the binary niche-overlap graphs. To our knowledge, this is the first study that considers niche-overlap graphs as weighted networks and conducts an analysis to reveal their structure. The rest of the paper is organized as follows. Section 2 describes the food-webs and the niche-overlap graphs. Section 3 presents the structural properties used to inspect the binary and weighted networks. Section 4 illustrates and discusses the results. Finally, Section 5 concludes.

2 Ecological Networks: Food-webs and niche-overlap graphs

Food-webs are examples of ecological networks. They describe the interactions between consumers and resources. These complex systems are illustrated by a directed network. Nodes characterize species and directed links map the feeding connections between them. Other networks, namely niche-overlap graphs, are also examples of ecological networks. These graphs depict the competition between consumers. Two predators (consumer) are linked if they share at least one prey (resource). Niche-overlap graphs are drawn considering the information (who eat whom) retrieved from the food-webs. There are two different ways of using this information: (1) searching only for the common prey for each predator or (2) taking in consideration the number of common prey for each predator. In the second case, the weight $\omega_{i,j}$ assigned to each edge will be defined using the Jaccard index [14]:

$$\omega_{i,j} = \frac{|\mathrm{prey}_i \bigcap \mathrm{prey}_j|}{|\mathrm{prey}_i \bigcup \mathrm{prey}_j|}, \tag{1}$$

where prey_i and prey_j are the prey of predator i and j respectively.
These weights provide important information on the competition between predators. Two nodes might have the same number of links. However the strength of their links might be different.

3 Datasets and Network Measures

3.1 Datasets

We selected a collection of 15 real food-webs and built their corresponding niche-overlap graphs (Table 1). Weighted niche-overlap graphs were also generated to assess the comparison with the binary ones.

Graph	Food-web Order	Size	Niche-overlap Order	Size	Graph	Food-web Order	Size	Niche-overlap Order	Size
Chesapeake	33	71	27	95	Mangrove	90	1151	84	2148
Cypdry	68	468	53	855	LRL North Spring 2	144	2095	111	2520
Cypress	64	437	50	827	LRL North Summer	165	2706	121	3064
Cypwet	68	459	53	854	LRL North Winter	109	1257	86	1501
Everglades	63	617	58	1214	LRL South Winter	102	1328	83	1418
Gramdry	66	664	60	1267	LRL South Spring 1	151	2399	112	2965
Saint Martin	44	218	38	312	LRL South Summer	173	2901	119	3652
Mangrovedry	94	1210	86	2315					

Table 1: Empirical food-webs and their associated niche-overlap graphs are presented by their name, order and size (number of links).

3.2 Network Measures

In order to assess a comparison between the architecture of these binary and weighted graphs, we selected four classical network measures that were used to analyse different networks such as social networks, biological networks, world wide web networks and others [15, 16, 17]. These measures are presented below:

Degree : The degree D_v is the number of links that a node v has.

Weighted Degree : The node strength D_v^W is the sum of the weights of the links that a node v has.

By taking into consideration the strength of each link, we obtain additional information on the importance of the competition that a predator has. D_v^W is certainly lower than D_v since the weights $\omega_{i,j}$ are in the interval $[0, 1]$.

Clustering Coefficient : The clustering coefficient C_v measures the tendency that the neighbors of a node v are linked to each other's. It is given by:

$$C_v = \frac{2E_v}{D_v(D_v - 1)} = \frac{\sum_{j,h} a_{vj}a_{vh}a_{jh}}{D_v(D_v - 1)}, \tag{2}$$

where a_{vj} is 1 if species v and j are connected (i.e. in competition) and 0 otherwise. The factor $\dfrac{D_v(D_v - 1)}{2}$ is the potential number of links among the neighbors. E_v is the number of links among the neighbors of v i.e. the actual number of triangles in which node v participates: $\dfrac{1}{2}\sum_{j,h} a_{vj}a_{jk}a_{kv}$. The clustering coefficient of the whole network, is the average clustering coefficient C over all the nodes.

Weighted Clustering Coefficient : Many definitions of the weighted clustering coefficient have been proposed in the literature [13, 18, 19, 20, 21].
In this paper, we restrict our analysis on the following two definitions : the one

proposed by Barrat et al. [13] which reflects how much of node strength is associated with adjacent triangle edges and the one proposed by Onnela et al. [21] which shows how large triangle weights are compared to network maximum.

Barrat et al. take into account only two links of the triangle:

$$C_v^{W(B)} = \frac{1}{s_v(D_v - 1)} \sum_{j,h} \frac{(w_{vj} + w_{vh})}{2} a_{vj} a_{vh} a_{jh}, \tag{3}$$

where s_v accounts for the strength of node v :

$$s_v = \sum_j a_{vj} w_{vj}.$$

The factor $s_v(D_v - 1)$ is the normalization factor to ensure that the weighted clustering is in the interval $[0, 1]$ and $\dfrac{w_{vj} + w_{vh}}{2}$ is the weights' average of the links between node v and its neighbors j and h.

If $C^{W(B)} > C$, this shows that the interconnected triples are more likely to be created by the links with larger weights. If $C^{W(B)} < C$, this indicates that these triples are formed by the links with lower weights [22].

Onnela et al. consider all the three link weights of a triangle:

$$C^{W(O)} = \frac{2}{D_v(D_v - 1)} \sum_{j,h} (\hat{w}_{vj} \hat{w}_{jh} \hat{w}_{hv})^{\frac{1}{3}} = C_v \bar{I}_v, \tag{4}$$

where $\hat{w}_{vj}$ is equal to $w_{vj}/\max(w)$. The actual number of triangles in which node v participates is replaced by the average intensity $\bar{I}_v$ of the triangle, which is the geometric mean of the links' weights $(\hat{w}_{vj} \hat{w}_{jh} \hat{w}_{hv})^{\frac{1}{3}}$ [11].

Betweenness Centrality : The betweenness centrality of a node v introduced by Freeman [23] identifies the number of shortest paths that passes through node v (denoted by $\sigma_{st}(v)$) among all the shortest paths (σ_{st}) in the network. This measure is given by:

$$BC_v = \sum_{s,t \neq v} \frac{\sigma_{st}(v)}{\sigma_{st}}. \tag{5}$$

If the betweenness centrality BC of a node v is equal to 0, it belongs to *only one complete subgraph* (a clique) of a graph G [24].

Weighted Betweenness Centrality : The weighted betweenness centrality of a node v is calculated by taking into consideration the weights of the links in the network when finding the shortest path that passes through v, $\sigma_{st}^w(v)$ and the ones among the network σ_{st}^w :

$$BC_v^W = \sum_{s,t \neq v} \frac{\sigma_{st}^w(v)}{\sigma_{st}^w}. \tag{6}$$

In the weighted version of the betweenness centrality, if a species has a BC^W equal to 0, this does not ensure that it belongs to one unique clique.

Assortativity Coefficient : the assortativity coefficient R of a graph measures the tendency of degree correlation. It is calculated using the correlation coefficient of Pearson applied to the degrees of each node in the network.

It is defined as:

$$R = \frac{\frac{1}{M}\sum_{\Phi}(\prod_{v \in F(\Phi)} D_v) - (\frac{1}{2M}\sum_{\Phi}(\sum_{v \in F(\Phi)} D_v))^2}{\frac{1}{2M}\sum_{\Phi}(\sum_{v \in F(\Phi)} D_v^2) - (\frac{1}{2M}\sum_{\Phi}(\sum_{v \in F(\Phi)} D_v))^2}, \tag{7}$$

where M is the total number of links in the network, $F(\Phi)$ denotes the set of two nodes linked by the Φth link [15].

Weighted Assortativity Coefficient : the weighted assortativity coefficient R^W suggested by Leung et al. [25] is given by :

$$R^W = \frac{\frac{1}{H}\sum_{\Phi}(w_\Phi \prod_{v \in F(\Phi)} D_v) - (\frac{1}{2H}\sum_{\Phi}(w_\Phi \sum_{v \in F(\Phi)} D_v))^2}{\frac{1}{2H}\sum_{\Phi}(w_\Phi \sum_{v \in F(\Phi)} D_v^2) - (\frac{1}{2H}\sum_{\Phi}(w_\Phi \sum_{v \in F(\Phi)} D_v))^2}, \tag{8}$$

where H is the total weight of all links in the network and w_Φ denotes the weight of the Φth link.

If $R^W > R$, this implies that the links with a larger weights are pointing to the neighbors with larger degree. If $R^W < R$, this shows that the links with a larger weights are pointing to the neighbors with smaller degree [26].

4 Results and Discussion

The average degree of species of the 15 weighted niche-overlap graphs was significantly lower compared to the binary ones (Fig. 1 (a)). This indicates that even though species compete with many other species, they actually share few resources between them, thus providing weak links.

The distribution of the weighted clustering coefficient proposed by Barrat et al. $C^{W(B)}$ is slightly higher than the one for the binary clustering coefficient C whereas the one suggested by Onnela et al. $C^{W(O)}$ is considerably lower among the others (Fig. 1 (b)). The differences between both definitions comes from the fact that Onnela et al. take into account the weights between neighbors of node v and the weights of the edges between neighbors. On the other hand, Barrat et al. consider only the weights of the triangle forming the edges linked to node v but not the edges connecting the neighbors of v. Both weighted clustering

coefficient ($C^{W(B)}$ and $C^{W(O)}$) provide us with complementary information. $C^{W(B)}$ being close the C yields to two conclusions : (1) the absence of correlation (randomized network), (2) the network is divided in two sets, one where triples are constituted by larger weights and others by smaller weights. $C^{W(O)}$ being significantly lower is due to the weight normalization by the global $\max(w)$ and to a broad distribution of weights in networks [27].
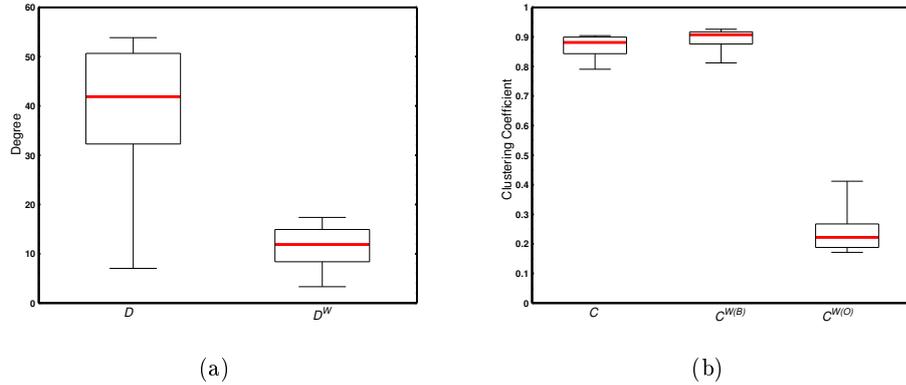

(a)   (b)

Fig. 1: Box plots (minimum, quartiles and maximum) illustrating the distribution of degree and clustering coefficient respectively of the 15 niche-overlap graphs. Medians are indicated by red lines. D and D^W correspond to the binary and weighted degree respectively. C denotes the binary clustering coefficient. $C^{W(B)}$ and $C^{W(O)}$ the one proposed by Barrat et al. and Onnela et al. respectively.

The percentage of species with a betweenness centrality equal to 0 differed between the binary and the weighted niche-overlap graphs (Fig. 2 (a)). A higher number of species with a $BC^W = 0$ was detected in the weighted version. This points out that some species have a stronger competition (a high number of shared prey) among the others.

Interestingly, the assortativity coefficient for the weighted networks was positive whereas for the binary ones it was close to 0 and slightly negative (Fig. 2 (b)). This points out that by considering the strength of links, niche-overlap graphs reveal a fairly tendency to be assortative. This expresses that predators with a high number of common prey tend to be connected with predators who also have a high number of common prey. Nevertheless, by considering simply the presence or absence of links (ignoring the weights), highlights a different assemblage of predators, indeed for binary niche-overlap graphs, predators tend to be linked randomly.

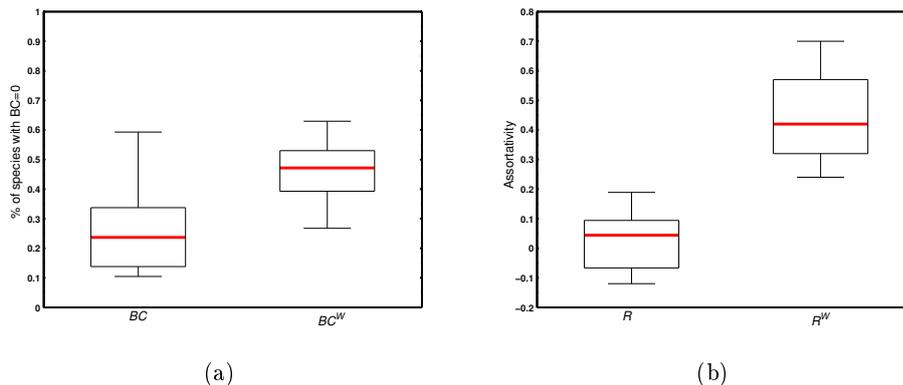(a)   (b)

Fig. 2: Box plots (minimum, quartiles and maximum) illustrating the distribution of betweenness centrality and assortativity respectively of the 15 niche-overlap graphs. Medians are indicated by red lines. BC and BC^W correspond to the binary and weighted betweenness centrality respectively. R and R^W correspond to the binary and weighted assortativity respectively.

5 Conclusion

In this work, a set of 15 real networks was considered to conduct a comparison between the structure of the binary and weighted niche-overlap graphs. Our analysis showed significant differences between both structures indicating the influence of the weights on the architecture and on the assemblage of species. We believe that our study provides new insights and additional topological information on the structure of niche-overlap graphs studied in the context of foodwebs.

References

1. R. Albert and A.L. Barabási. Statistical mechanics of complex networks. *Reviews of modern physics*, 74(1):47, 2002.
2. M. EJ. Newman. The structure and function of complex networks. *SIAM review*, 45(2):167–256, 2003.
3. M. EJ. Newman. Fast algorithm for detecting community structure in networks. *Physical review E*, 69(6):066133, 2004.
4. J. Little, D. Shepley, and D. Wert. Robustness of a gene regulatory circuit. *The EMBO Journal*, 18(15):4299–4307, 1999.
5. J. A. Dunne, R. J. Williams, and N. D. Martinez. Network structure and biodiversity loss in food webs: robustness increases with connectance. *Ecology Letters*, 5(4):558–567, 2002.
6. M. Girvan and M. EJ. Newman. Community structure in social and biological networks. *Proceedings of the National Academy of Sciences*, 99(12):7821–7826, 2002.

7. D. J. Watts. *Small worlds: the dynamics of networks between order and randomness*. Princeton university press, 1999.
8. J. A. Dunne, R. J. Williams, and N. D. Martinez. Small networks but not small worlds: unique aspects of food web structure. *Proc. Nat. Acad. Sci*, 2002.
9. N. Sokhn, R. Baltensperger, J. Hennebert, U. Ultes-Nitsche, and L.F. Bersier. Structure analysis of niche-overlap graphs. *NetSci2013*, 2013.
10. A.L. Barabási, R. Albert, and H. Jeong. Scale-free characteristics of random networks: the topology of the world-wide web. *Physica A: Statistical Mechanics and its Applications*, 281(1):69–77, 2000.
11. A. Ioannis and T. Eleni. Statistical analysis of weighted networks. *arXiv preprint arXiv:0704.0686*, 2007.
12. M. EJ. Newman. Analysis of weighted networks. *Physical Review E*, 70(5):056131, 2004.
13. A. Barrat, M. Barthelemy, R. Pastor-Satorras, and A. Vespignani. The architecture of complex weighted networks. *Proceedings of the National Academy of Sciences of the United States of America*, 101(11):3747–3752, 2004.
14. P. Jaccard. Distribution de la flore alpine dans la bassin de dranses et dans quelques regions voisines. *Bulletin de la Societe Vaudoise des Sciences Naturelles*, 37:241–272, 1901.
15. M. EJ. Newman. Assortative mixing in networks. *Physical review letters*, 89(20):208701, 2002.
16. L. A. Adamic. The small world web. In *Research and Advanced Technology for Digital Libraries*, pages 443–452. Springer, 1999.
17. L. C. Freeman. Centrality in social networks conceptual clarification. *Social networks*, 1(3):215–239, 1979.
18. B. Zhang, S. Horvath, et al. A general framework for weighted gene co-expression network analysis. *Statistical applications in genetics and molecular biology*, 4(1):1128, 2005.
19. G. Kalna and D. J. Higham. Clustering coefficients for weighted networks. In *Symposium on Network Analysis in Natural Sciences and Engineering*, page 45, 2006.
20. L. Lopez-Fernandez, G. Robles, and J. M. Gonzalez-B. Applying social network analysis to the information in cvs repositories. 2004.
21. J.P. Onnela, J. Saramäki, J. Kertész, and K. Kaski. Intensity and coherence of motifs in weighted complex networks. *Physical Review E*, 71(6):065103, 2005.
22. S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D.U. Hwang. Complex networks: Structure and dynamics. *Physics reports*, 424(4):175–308, 2006.
23. L. C. Freeman. A set of measures of centrality based on betweenness. *Sociometry*, pages 35–41, 1977.
24. R. Grassi, R. Scapellato, S. Stefani, and A. Torriero. Betweenness centrality: extremal values and structural properties. In *Networks, Topology and Dynamics*, pages 161–175. Springer, 2009.
25. C. C. Leung and H. F. Chau. Weighted assortative and disassortative networks model. *Physica A: Statistical Mechanics and its Applications*, 378(2):591–602, 2007.
26. M. Barthélemy, A. Barrat, R. Pastor-Satorras, and A. Vespignani. Characterization and modeling of weighted networks. *Physica A: Statistical mechanics and its applications*, 346(1):34–43, 2005.
27. J. Saramäki, M. Kivelä, J.P. Onnela, K. Kaski, and J. Kertesz. Generalizations of the clustering coefficient to weighted complex networks. *Physical Review E*, 75(2):027105, 2007.