Connectivity and Nestedness in Bipartite Networks from Community Ecology

Gilberto Corso¹, A. I. Levartoski de Araujo², Adriana M. de Almeida³

¹ Departamento de Biofísica e Farmacologia, Centro de Ciências, Universidade Federal do Rio Grande do Norte, UFRN - Campus Universitário, Lagoa Nova, CEP 59078 972, Natal, RN, Brazil
² Instituto Federal de Educação, Ciência e Tecnologia do Ceará Av. Treze de Maio, 2081 - Benfica CEP 60040-531 - Fortaleza, CE, Brazil
³ Departamento de Botânica, Ecologia e Zoologia, Centro de Ciências, Universidade Federal do Rio Grande do Norte, UFRN - Campus Universitário, Lagoa Nova, CEP 59078 972, Natal, RN, Brazil
E-mail: corso@cb.ufrn.br

Abstract. Bipartite networks and the nestedness concept appear in two different contexts in theoretical ecology: community ecology and islands biogeography. From a mathematical perspective nestedness is a pattern in a bipartite network. There are several nestedness indices in the market, we used the index $\nu$. The index $\nu$ is found using the relation $\nu = 1 - \tau$ where $\tau$ is the temperature of the adjacency matrix of the bipartite network. By its turn $\tau$ is defined with help of the Manhattan distance of the occupied elements of the adjacency matrix of the bipartite network. We prove that the nestedness index $\nu$ is a function of the connectivities of the bipartite network. In addition we find a concise way to find $\nu$ which avoid cumbersome algorithm manipulation of the adjacency matrix.

1. Introduction
The last decade have experienced an explosion of papers on network theory [1]. Despite the fact that some of most popular networks are bipartite the works on bipartite networks are a minority. A very popular bipartite network is the actors collaboration network [2] where actors and movies are two sets and elements of these two sets have connections every time an actor participate in a movie. However, this bipartite network is presented in a simplified form where movies do not take part. The usual actor network has actors as nodes and a link between two actors is established every time they collaborate in the same movie. Of course there is also the movie network, but it is not so popular. An other well known network that comes from a projection of a bipartite graph is the scientific colaboration network [3].

An important application of bipartite networks are the interaction networks that comes from community ecology. Plants and their pollinators or animals and their parasites are examples of interaction networks in community ecology. We call specialist a species that interact with one or few species and generalist a species that interact with many. Nestedness is investigated in these networks characterizing the generalist-specialist balance in communities [4]. In nested interaction matrices, generalist species interact with many species, while a specialist interacts
Figure 1. A sketch of a bipartite network of plants and animals, the flower-polinater network is the standard example. In the left we show the usual bipartite representation while in the right the lattice-like representation.

with few ones. In nested matrices, species are organized in such a way that specialist species only interact with more generalist ones. Ecologists suggest that nestedness patterns can distinguish between mutualist (e.g., pollinator and plants) and antagonist (animals and parasites) networks [4]. In addition, the nestedness degree is important in the study of species co-evolution [5].

Bipartite networks have also been used in island biogeography [6], a very traditional branch of community ecology. Given a set of islands (from an archipelago) and a set of species, every time a species is found on an island a link between elements of these two sets is established, originating a bipartite network. In general, larger islands will support more species. But one can pose the question: species present in small islands will always be present in larger ones? A rough answer to this question is yes. This problem was studied for many archipelagos and several taxa (birds, lizards, beetles, etc) [7]. This pattern of species-poor sites composing a subset of species-rich sites is called a nested structure. Since forest remnants can be considered islands in a sea of anthropogenic disturbed landscape, conservation studies have revived some aspects of island biogeography in last three decades [8].

In figure 1 we show two representations of an interaction network between a set of plants \(i\) and animals \(j\). We sketch a plant-polinater network just to fix the ideas, in fact the theoretical framework developed is valid for any bipartite network. In the lattice format the full spaces indicates the presence of interactions while the blank spaces the absences of interactions. The same information is also show in a graph representation. In this case the mathematical object that describes this phenomenon is a bipartite network which establishes links between two sets: animals and plants. The main quantity that characterizes the network is the connectivity of animals and plants, \(k_A(i)\) and \(k_P(j)\), which count the number of interactions of animals and plants.

The main objective of this manuscript is to show that a recent index of nestedness \(\nu\) introduced in the literature [9] depends only in the connectivity of the bipartite network. The work is organized as follows: in Section 2 we outline the meaning of nestedness in community
Figure 2. Two adjacency matrices with the same occupation, the matrix in the left is more nested than the matrix in the right.

ecology and introduce other indices used by the specialists in the area; in Section 3 we present the nestedness index $\nu$, its main mathematical properties and derive a relation between $\nu$ and the interactions of animals and plants; finally in Section 4 we conclude the work.

2. The nestedness concept

Nestedness is not a straightforward concept, and this fact leads to misunderstandings in the literature as well as a proliferation of estimators [10] (this situation is similar to the definition of diversity, another key concept in community ecology that also shows several conflicting definitions and estimators [11]). To visualize nestedness from the adjacency matrix it is common in this area to rank rows and columns of the matrix. In this way the matrix usually show in the lattice format of fill and empty squares, present a decreasing connectivity where the large number of connections is packed in the upper left corner.

The oldest and most popular nestedness estimator is the ”temperature” of Atmar and Patterson [12]. This index is constructed using a median line in the ranked adjacency matrix. This line equally separates holes and full cells in the matrix (zeros and ones). The ”temperature” index $\tau$ is a measure of the dispersion of holes in respect to the median. There are other indices [13, 14] that basically count the number of vacancies in the matrix. The difficult point of these indices is that they are defined by cumbersome algorithms that impede further analytical developments.

In figure 2 we show two adjacency matrices to help to clarify the nestedness idea. The two matrices have the same occupation $\rho = 7/16$, the occupation is defined by the number of ones divided by the total number elements of the matrix. In both cases the sum of the total marginals (connectivities $k$) is ranked as we have previously explained. The matrix in the left is more nested than the matrix in the right because it has more holes. In this matrix the occupied elements are more distant from the upper left corner, and we will use this property to define the nestedness index in the next section.

3. The nestedness coefficient and the connectivity

We investigate nestedness using a nestedness index $\nu = 1 - \tau$ developed in [9]. This estimator is based on distances on the adjacency matrix of the network. The intuitive idea behind this nestedness index comes from the dispersion of ones and zeros through the adjacency matrix. As exposed in the previous section we start packing the adjacency matrix of the observed interactions. A highly nested matrix is the one that, after packing, present a minimal mixing of
ones and zeros. We estimate nestedness using the average distance \( d \) of the matrix:

\[
D = \frac{1}{N} \sum_{k=1}^{N} d_{i,j}(k)
\]  

(1)

for \( N \) the number of ones in the matrix and \( d_{i,j} = i + j \) the Manhattan distance. To normalize the temperature index \( \tau \), two artificial matrices are used: the perfectly nested matrix and the equiprobable random matrix. These two matrices have the same sizes \( L_1, L_2 \) and occupancy \( \rho \) of the adjacency matrix, and they work as benchmarks to properly define the nestedness index. In other words, we parametrize \( \tau \) with help of the distance \( D \) of \( D_{nest} \) and \( D_{rand} \). The first is the average distance related to a completely nested matrix and the second to the random matrix. \( \tau \) is defined as follows:

\[
\tau = \frac{D - D_{nest}}{D_{rand} - D_{nest}}
\]  

(2)

A zero \( \tau \) corresponds to a state of minimal disorder (zero temperature), where all elements are perfectly nested. Conversely, \( \tau = 1 \) corresponds to a state of equiprobable randomly dispersed matrix elements (high temperature). The main advantage of \( \tau \) is that it is posed in an analytical way. Because of that we have already discussed the relation between nestedness and abundance [15] and now we intend to work the connection between nestedness and connectivity.

A puzzling question when studying the nestedness pattern is the following: is it possible that nestedness can be defined only by the connectivity of matrix elements? An answer to this question will strongly help to demystify the nestedness concept. Using our index expressed in (1) and (2) we answer positively to this question. We start with the definition of \( D \) as expressed above:

\[
D = \frac{1}{N} \sum_{i,j} d_{i,j}
\]  

(3)

where we emphasize that the sum is taken over the \( N \) occupied elements. As the Manhattan distance of each element is \( (i + j) \) and the matrix elements in this equation may be written in the following form:

\[
D = \frac{1}{N} \sum_{i,j} a_{i,j} (i + j) = \frac{1}{N} (\sum_{i,j} a_{i,j} i + \sum_{i,j} a_{i,j} j)
\]  

(4)

where the elements \( a_{i,j} \) assume the values 1 or 0 for occupied and absent elements. Using the definition of \( k_A(i) = \sum_j a_{i,j} \) and \( k_P(j) = \sum_i a_{i,j} \) we conclude:

\[
D = \frac{1}{N} (\sum_i i k_A(i) + \sum_j j k_P(j)).
\]  

(5)

This formula sets that the distance \( D \) depends only on the connectivity of animals and plants. As the nestedness \( \nu = 1 - \tau \) that depends on \( D \) we have proved that nestedness rely on the connectivities of animals and plants. In addition Equation (5) enable ones to find \( \nu \) without complex algorithmic manipulation of the bipartite network. In this way Equation (5) shows the nestedness in a compact form, this situation is quite distinct of other algorithms that calculate nestedness that are based on elaborated algorithmic manipulation of the bipartite netowork [12, 13, 14].
4. Final Remarks

We use the metric concept of nestedness previously developed, to prove that the nestedness index is completely determined by the connectivity of the bipartite network. Indeed, we prove a simple statement that can simplify the theoretical framework of community ecology. Rigorously speaking, the nestedness index is not a necessary concept, the connectivity is enough to characterize the nestedness of a bipartite network. Another important aspect of this work concerns the algorithmic work to find the nestedness index. The Patterson and Atmar index or the Brualdi index cited in Section 2 have a computational complexity much larger that our index. Indeed $\nu$ can be easily calculated by Equation (5).

It is worth to note that Equation (5) gives that the sum of the distances $D$ is a function of plants and animal connectivities, that means: $D = D(k_A(i), k_P(j))$. The usual way of considering connectivity in network theory [1, 16] is employing the distribution of connectivity $P(k)$ that is the integral of the connectivity $k(l)$. As $D$ is not a constant, but a function of the connectivities, we have not proved that the nestedness is a function of $P(k)$. Therefore the present work analytically advances the original paper [15] that defines a nestedness metric, but does not provides any hint of the relation of this metric with the connectivities of the bipartite network.

We believe that a well defined nestedness estimator can be useful to investigate order and structure in bipartite networks such as the metabolic network of a cell or social webs like actors bipartite network [16]. To conclude, nestedness is a subject of high current interest in biogeography, community ecology and conservation biology. The Web of Science shows more than 300 papers on the subject in the last 14 years. In this way, this work will be of interest to researchers in a great multitude of areas, as well as to physicists working with patterns in bipartite networks.

Acknowledgements

The authors gratefully acknowledge the financial support of CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico, Brazil). We also thanks to Thomas Lewinsohn and Paulo Inacio Prado for the encouraging discussions in theoretical ecology.

References

[1] Costa L F Rodrigues F A Travieso G Villas Boas P R 2007 Characterization of complex networks: A survey of measurements Adv. Phys. 56 167–242
[2] Jiang Y Xu T He D 2004 A self-adaptive model of actor-collaboration network Int. J. Mod. Phys. B 18, 2004–2007
[3] Newman M E J 2001 Scientific collaboration networks Phys. Rev. E 64 016131.
[4] Lewinsohn T M Prado P J Jordano P Bascompte J and Olesen J M 2006 Structure in plant-animal interaction assemblages Oikos 113 174–184
[5] Bascompte J Jordano P and Olesen J M 2006 Asymmetric coevolutionary networks facilitate biodiversity maintenance Science 312 431–433
[6] Mac Arthur R H and Wilson E O 1967 The theory of Island Biogeography Monographs in Population Biology. Princeton Univ. Press.
[7] Darlington P J 1957 Zoogeography: the geographical distribution of animals Wiley, New York.
[8] Fleishman E Betrus C J Blair R B MacNally R. and Murphy D 2002 Nestedness analysis and conservation planning: the importance of place, environment, and life history across taxonomic groups Oecologia 133 78–78
[9] Araujo A I L Corso Gilberto Almeida A M and Lewinsohn T M 2010 An analytical approach to the measurement of nestedness in bipartite networks Physica A 389 1405–1411
[10] Almeida-Neto M Guimarães Jr P R Lewinsohn T M 2007 On nestedness analyses: rethinking matrix temperature and anti-nestedness. Oikos 116 716–722.
[11] Magurran A E 2003 Measuring Biological Diversity Blackwell Pub.
[12] Patterson B D and Atmar W 1986 Nested subsets and the structure of insular mammalian faunas and archipelagos Biol. J. Linn. Soc. 28 65–82
[13] Brualdi R A and Sanderson J G 1999 Nested species subsets, gaps and discrepancy *Oecologia* **119** 256–264
[14] Almeida-Neto M, Guimarães P, Guimarães Jr P R, Loyola R D and Ulrich W 2008 A consistent metric for nestedness analysis in ecological systems: reconciling concept and measurement *Oikos* **117** 1227–1239
[15] Araujo A I L., Almeida A M, Cardoso M Z and Corso G 2010. Abundance and nestedness in interaction networks *Ecological Complexity* Preprint doi:10.1016/j.ecocom.2010.02.004
[16] Albert R and Barabási A L 2002 Statistical mechanics of complex networks *Rev Mod Phys* **74** 47–97