BiWeb : a common framework to facilitate the analysis of bipartite networks

Timothée Poisot*[1,2] , Cesar Flores[3], Sergi Valverde[4], Joshua S Weitz[3,5]

[1] UQAR
[2] Québec Centre for Biodiversity Science
[3] School of Physics, Georgia Institute of Technology, Atlanta, GA 30332
[4] Universitat Pompeu Fabra, Barcelona, Spain
[5] School of Biology, Georgia Institute of Technology, Atlanta, GA 30332

Email: Timothée Poisot*- timothee.poisot@uqar.ca; Cesar Flores - cesar.flores@gatech.edu; Sergi Valverde - svalver@gmail.com; Joshua S Weitz - jsweitz@gatech.edu;

*Corresponding author

Abstract

Almost all ecological interactions can be represented as networks. The structure of this network holds much information about the ecological mechanisms acting in communities, and can help predict their evolutionary responses. However, there is a large number of different metrics for the same property, which makes it difficult for people to get started with these analyses. In this paper, we describe two open-source packages, one in Octave/MatLab and one in Python, to perform the most standard analyses in a straigthforward way.

Background

Viewing species interaction as a network is an efficient way to reduce their inherent complexity to an object that is mathematically analyzable [1, 2]. Among them, bipartite graphs [3] are especially suited to the analysis of systems in which two distinct sets of organisms are interacting. Although an abundant literature on plant–pollinators systems emerged around the use of these networks [4–8], they are with a few exceptions [9,10] less frequently used to analyze antagonistic interactions [11]. Recent work emphasized that viewing antagonistic interactions as being bipartite networks allowed to look at emerging properties of these systems, including how they reacted to changes in environmental quality [12,13], and how their structure relates to eco-logical and evolutionary mechanisms [14]. To some extent, the lack of widespread use of these methods can be explained by two factors. First, thinking in terms of networks requires to gain familiarity with a new literature and nomenclature, which can be daunting, although recent papers made a great effort of "translating" the different notions from one field to another [11]. Second, the analysis of bipartite networks can be an intense task from a computational point of view, and out of the multiple methods existing to measure a given property, not all are publicly available or easily usable. Additionally, several groups often independently developed different metrics for the same property of a network, and choosing the most robust or meaningful one is a difficult task for peopke unfamiliar with the litterature. The goal

of this paper is to provide a toolbox for ecologists wanting to analyze bipartite networks. We propose an implementation of this toolbox in two languages, Python and MatLab.

Implementation

We propose two alternative implementations of the toolbox, `biweb-ml` (MatLab/Octave) and `biweb-py` (Python). Our toolbox allows the estimation of nestedness and modularity, using two robust estimators, along with ways of network visualization allowing to reflect both properties. It is assumed that data come in matrix shape, with organisms of the upper level as rows, and organisms of the lower levels as columns. The matrix can be filled with either 0 and 1 (adjacency matrix), or with numerical values indicating the strength of the pairwise interaction: $\mathbf{M}_{ij}$ is the strength of the interaction between species i and species j. However, some of the available analysis work only with boolean matrices (i.e. nestedness, modularity). In these cases the boolean version of $\mathbf{M}$ is used. For any kind of matrix input, all rows and columns that contain only zeros are deleted before any type of analysis.

For each property, as detailed in the following paragraphs, only one measure was selected. This is intended to facilitate the analysis, as while contrasting the results of two different metrics can be interesting in some situations, it risks confusings people not interested by network theory. The metrics to include were selected based on a litterature review and tests using datasets either public or gathered by the authors.

Measures

Nestedness is measured using NODF [15], which is a robust algorithm, insensitive to network shape (i.e. asymmetry in the number of species at each level) and shape (i.e. number of species in the whole network). Nestedness is measured for the whole network, and separately for each of the levels.

Bipartite modularity is estimated using the LP-BRIM method [16], which optimizes Barber's modularity [17, 18] in a short time. With this method, best results are achieved by selecting the optimal solution of a large number of runs. In addition to reporting Barber's Q_{bip}, we propose an additional measure termed the realized modularity, noted Q_R.

This measure reflects the *a posteriori* goodness of the division of the networks in modules, and is calculated by dividing the number of interactions established by species belonging to the same module by the total number of interactions in the network. If all links are established by species belonging to the same module, then $Q_R = 1$, reflecting the perfect modularity. If there are as many links established between and within modules, then Q_R takes its minimal value of $1/2$. Note that due to this, Q_R can be ranged in $[0, 1]$, with $Q'_R = 2 \times Q_R - 1$.

We also implement several "species-level" metrics. These include generalitty and vulnerability [?], specificity as measured with (PDI) and without (RR) link strength [?], and the Species- Specialization Index [?]. In addition, when the analysis to measure modularity has been done, the module to which each species belong is recorded.

Null models

We propose four null models allowing to test the significancy of the value of nestedness and modularity (see [6, 14, 19] for more details), working by generating random networks through a Bernoulli process, where the probability of interactions are determined following different rules. If $\mathbf{V}$ and $\mathbf{G}$ are vectors with the number of interactions respectively received by lower level species, and established by upper level species, L is the number of interactions in the network, and l and u are respectively the number of species at the lower and upper levels, then the probability of two species having an interaction in the random network, P_{ij}, is

in model 1 , $P_{ij} = L/(ul)$ – the connectance of the network is respected, but not the number. of interactions in which each species is involved.

in model 2 , $P_{ij} = (G_i/l + V_j/u)/2$ – the connectance, and the number of interactions in which each species is involved, are respected

in model 3r , $P_{ij} = G_i/l$ – the connectance, and the number of interactions of species at the upper level, are respected

in model 3c , $P_{ij} = V_j/u$ – the connectance, and the number of interactions of species at the lower level, are respected

We impose the additional constraint that in the random network, no species can be left with no interaction at the end of the randomization process, so that only networks of the same size and shape are generated through this process. For each null model, the original value of each measure (*e.g.* NODF) is measured on the observed network, and assigned a value N. For all of the replicates, the randomized value of the measure (N') is calculated. Once this is done for all replicates, we use the distribution of absolute differences $N - N'$, and use a one-sample t-test to assess whether its mean is significantly different from 0.

Visualization

Results

- FIG 1a convergence time number of runs null models
- FIG 1b convergence time number of runs modularity
    - FIG 2 Different representations web
    - TAB 1 species level
    - TAB 1 network-level

Conclusions

Availability & requirements

The following section lists the informations and modalities of use of the different packages.

Project name: biweb
Project home: http://www.github.com/tpoisot/biweb
Operating system(s): Platform independent
Programming language: MatLab, Octave, Python
Other requirements: Python 2.6 or higher, LaTeX, and the following packages: *pp*, *numpy*, *scipy*, *pyx*
Licence: GNU GPL
Restrictions to use: Papers using this package should cite the present publication

Author's contributions

Developped the software: TP, CF. Conceived the study: TP, CF, SV, JSW. Analyzed the results: TP, CF. All authors contributed to the redaction of the paper.

References

1. Proulx SR, Promislow DEL, Phillips PC: **Network thinking in ecology and evolution.** *Trends in Ecology & Evolution* 2005, **20**(6):345–353, [http://www.ncbi.nlm.nih.gov/pubmed/16701391].

2. Dunne JA: **The network structure of food webs**. In *Ecological Networks: Linking Structure to Dynamics in Food Webs.* Edited by Pascual M, Dunne JA, Oxford: Oxford University Press 2006:27–86.

3. Chartrand G: *Introductory Graph Theory.* Dover 1985, [http://www.springer.com/west/home/generic/order?SGWID=4-40110-22-1256251-0].

4. Bascompte J, Jordano P: **Plant-Animal Mutualistic Networks: the Architecture of Biodiversity**. *Annual Review of Ecology, Evolution, and Systematics* 2007, **38**:567–593.

5. Stouffer DB, Bascompte J: **Compartmentalization increases food-web persistence.** *Proceedings of the National Academy of Sciences of the United States of America* 2011, **108**(9):3648–3652, [http://www.ncbi.nlm.nih.gov/pubmed/21307311].

6. Bascompte J, Jordano P, Melián CJ, Olesen JM: **The nested assembly of plant–animal mutualistic networks**. *Proceedings of the National Academy of Sciences of the United States of America* 2003, **100**(16):9383–9387, [http://dx.doi.org/10.1073/pnas.1633576100].

7. Bastolla U, Fortuna MA, Pascual-García A, Ferrera A, Luque B, Bascompte J: **The architecture of mutualistic networks minimizes competition and increases biodiversity**. *Nature* 2009, **458**(7241):1018–1020.

8. Joppa LN, Bascompte J, Montoya JM, Solé RV, Sanderson J, Pimm SL: **Reciprocal specialization in ecological networks.** *Ecology Letters* 2009, **12**(9):961–969, [http://www.ncbi.nlm.nih.gov/pubmed/19566586].

9. Thebault E, Fontaine C: **Does asymmetric specialization differ between mutualistic and trophic networks?** *Oikos* 2008, **117**(4):555–563.

10. Thebault E, Fontaine C: **Stability of ecological communities and the architecture of mutualistic and trophic networks**. *Science* 2010, **329**(5993):853–856, [http://dx.doi.org/10.1126/science.1188321].

11. Poulin R: **Network analysis shining light on parasite ecology and diversity.** *Trends in Parasitology* 2010, **26**(10):492–498, [http://www.ncbi.nlm.nih.gov/pubmed/20561821].

12. Poisot T, Lepennetier G, Martinez E, Ramsayer J, Hochberg ME: **Resource availability affects the structure of a natural bacteria-bacteriophage community.** *Biology Letters* 2011, **7**(2):201–204, [http://www.ncbi.nlm.nih.gov/pubmed/20961886].

13. Poisot T, Thrall PH, Hochberg ME: **Trophic network structure emerges through antagonistic coevolution in temporally varying environments**. *Proceedings of the Royal Society B: Biological Sciences* 2011, :in press.

14. Flores CO, Meyer JR, Valverde S, Farr L, Weitz JS: **Statistical structure of host-phage interactions.** *Proceedings of the National Academy of Sciences of the United States of America* 2011, **108**(28):E288–297, [http://www.ncbi.nlm.nih.gov/pubmed/21709225].

15. Almeida-Neto M, Guimaraes P, Guimaraes Jr P, Loyola R, Ulrich W: **A consistent metric for nestedness analysis in ecological systems: reconciling concept and measurement**. *Oikos* 2008, **117**(8):1227–1239, [http://onlinelibrary.wiley.com/doi/10.1111/j.0030-1299.2008.16644.x/full].

16. Liu X, Murata T: **Community detection in large-scale bipartite networks**. In *Transactions of the Japanese Society for Artificial Intelligence*, *Volume 5* 2010:184–192.

17. Barber M: **Modularity and community detection in bipartite networks**. *Physical Review E* 2007, **76**(6), [http://link.aps.org/doi/10.1103/PhysRevE.76.066102].

18. Barber MJ, Faria M, Streit L, Strogan O: **Searching for communities in bipartite networks**. *ArXiV* 2008, [http://arxiv.org/abs/0803.2854].

19. Ulrich W, Gotelli NJ: **Null model analysis of species nestedness patterns**. *Ecology* 2007, **88**(7):1824–1831.

Figures

Figure 1 - Sample figure title

A short description of the figure content should go here.

Figure 2 - Sample figure title

Figure legend text.

Tables

Table 1 - Null models and their interpretation

Here is an example of a *small* table in LaTeX using `\tabular{...}`. This is where the description of the table should go.

Model name	Works by	if significant
1	creating a random network with equal connectance	C3
2	...	..
3r	..	.
3c	..	.

Additional Files

Additional file 1 — Sample additional file title

Additional file descriptions text (including details of how to view the file, if it is in a non-standard format or the file extension). This might refer to a multi-page table or a figure.

Additional file 2 — Sample additional file title

Additional file descriptions text.