

Introducing the bipartite Package: Analysing Ecological Networks

Carsten F. Dormann, Bernd Gruber and Jochen Fründ

Introduction

Interactions among species in ecological communities have long fascinated ecologists. Prominent recent examples include pollination webs (Memmott et al., 2004), species-rich predator-prey systems (Tylianakis et al., 2007) and seed dispersal mutualisms (all reviewed in Blüthgen et al., 2007). Many of the topological descriptors used in food webs since the 1960s have limited ecological meaning when only two trophic levels are investigated (for example chain length: Pimm, 1982/2002; Montoya et al., 2006). Here, the network becomes *bipartite*, i.e. every member of one trophic level is only connected to members of the other trophic level: direct interactions *within* trophic levels are regarded as unimportant. For bipartite ecological networks very many, more or less different, indices have been proposed to capture important features of the interactions and of the species. For example, species degrees (number of species the target species is linked to) or species strength (sum of level of dependencies of other species on the target) are supposed to quantify the importance of a species in a web.

The new R-package *bipartite*, introduced here, provides functions to visualise webs and calculate a series of indices commonly used to describe pattern in ecological webs. It focusses on webs consisting of only two trophic levels, e.g. pollination webs or predator-prey-webs. We had three types of ecological bipartite webs in mind when writing this package: seed-disperser, plant-pollinator and predator-prey systems.

Bipartite networks, as analysed and plotted in the package *bipartite*, can be represented by a matrix, in which, in our definition, columns represent species in the higher trophic level, and rows species in the lower trophic level. Entries in the matrix represent observed links, either quantitatively (with one to many interactions per link) or qualitatively (binary). Usually such matrices are very sparse, marginal sums (i.e. abundance distributions) highly skewed and average number of interactions per link are low (around 2: Blüthgen et al., 2007).

With the package *bipartite*, presented here, we wanted to overcome two main deficiencies in the field: 1. Lack of software to calculate various indices and topological descriptors of bipartite networks. And 2. No convenient plotting tool for bipartite networks. This article aims to briefly present the two visualisation functions (*plotweb* and *visweb*),

then present an example output from the calculation of network-level descriptors (using function *networklevel*) and finally address some miscellaneous issues to do with fitting degree distributions, secondary extinction slopes and null models for bipartite webs.

Along with several functions we also include 19 data sets on pollinator networks, taken from the National Center for Ecological Analysis and Synthesis webpage devoted to this topic (www.nceas.ucsb.edu/interactionweb). There are several other bipartite data sets at this repository, and our data include only quantitative plant-pollinator networks.

Plotting ecological networks

The function *plotweb* draws a bipartite graph, in which rectangles represent species, and the width is proportional to the sum of interactions involving this species. Interacting species are linked by lines, whose width is again proportional to the number of interactions (but can be represented as simple lines or triangles pointing up or down). An example is given in Fig. 1 for the data set *mosquin1967*, which is included in the package and can be generated using *plotweb(mosquin1967)*.

Alternatively, the function *visweb* plots the data matrix with shading representing number of interactions per link. As default, this gives an idea about the filling of the matrix. With option *type="diagonal"*, however, *visweb* depicts compartments for easy perception (Fig.2). The same compartments are visible in Fig.1, too, due to the default sequence of species here being the arrangement used in *visweb(., type="diagonal")*. This sequence is determined, as suggested by (Lewinsohn et al., 2006), using correspondence analysis.

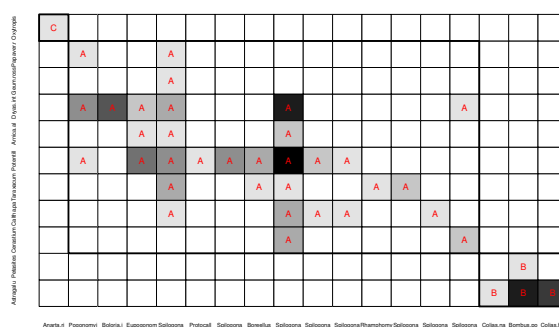


Figure 2: A plot of the network matrix produced by *visweb(mosquin1967, type="diagonal")*.

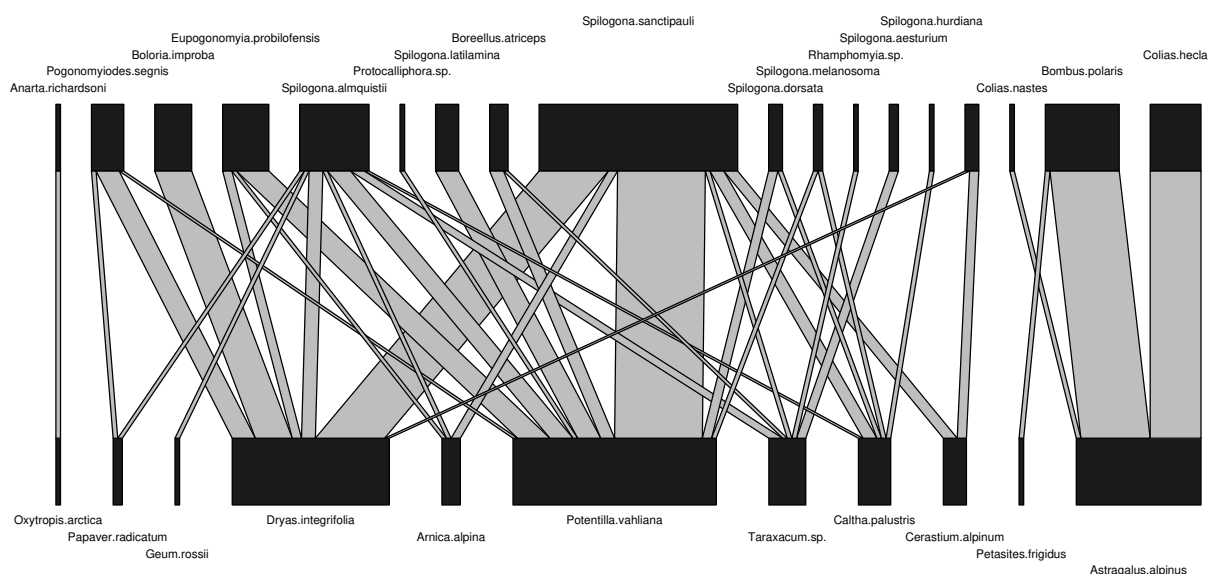


Figure 1: A bipartite graph produced by default settings of function plotweb.

Calculating network metrics

bipartite features two main functions to calculate indices: `specieslevel` and `networklevel`. The former returns various values describing, e.g., the specialisation or the dependence asymmetry of each species. Since its output is bulky, we omit it here and only present the function `networklevel` in detail, which also comprises a larger set of different indices:

```
>networklevel(mosquin1967)
$'number of higher trophic species'
[1] 18

$'number of lower trophic species'
[1] 11

$'number of links'
[1] 1.310345

$generality
[1] 2.677306

$vulnerability
[1] 4.114345

$'interaction evenness'
[1] 0.8512671

$'Alatalo interaction evenness'
[1] 0.6587369

$'number of compartments'
[1] 3

$'compartment diversity'
[1] 2.00787

$'cluster coefficient'
[1] 0.1363636

$H2
[1] 0.4964885

$'web asymmetry'
```

```
[1] 0.2413793

$'interaction strength asymmetry'
[1] 0.1607192

$'specialisation asymmetry'
[1] -0.1755229

$'extinction slope lower trophic level'
[1] 2.286152

$'extinction slope higher trophic level'
[1] 2.9211

$'degree distribution lower trophic level'
      Estimate Std.Err Pr(>|t|)  R2  AIC
exponential    0.207527 0.02905 0.000834 0.992 -7.11
power law       0.701034 0.08856 0.000517 0.967 -8.09
trunc. power law [slope] 0.431810 0.31661 0.244321 0.987 -7.15

$'degree distribution higher trophic level'
      Estimate Std.Err Pr(>|t|)  R2  AIC
exponential    0.221084 0.04283 0.006691 0.999 -3.21
power law       0.744383 0.12834 0.004394 0.960 -4.34
trunc. power law [slope] 0.511777 0.43347 0.322823 0.980 -2.82

$'higher trophic level niche overlap'
[1] 0.2237163

$'lower trophic level niche overlap'
[1] 0.2505869

$'mean number of shared hosts'
[1] 0.8545455

$togetherness
[1] 0.1050109

$'C-score'
[1] 0.6407096

$'V-ratio'
[1] 11.11811

$nestedness
[1] 44.28693
```

We opted for a list structure to be able to accommodate tables in the output, and because the option

index allows specification of the indices the user is interested in (defaults to all).

All indices are explained and/or referenced, in the help pages, so a detailed description is omitted here. Among our personal favourites are the network-wide specialisation H'_2 (Blüthgen et al., 2006), generality and vulnerability (Tylianakis et al., 2007) and the small-world measure 'clustering coefficient' (Watts and Strogatz, 1998). Furthermore, we took the liberty to modify the index 'dependence asymmetry', because it has been shown to be biased (Blüthgen et al., 2007). The original formulation is available as a special case of 'interaction strength asymmetry' and can be called using `networklevel(mosquin1967, index="ISA", ISAMethod="Bascompte")`.

Miscellaneous

Three list entries may warrant particular mentioning: *Web asymmetry* is simply the ratio of matrix dimensions. In a recent paper, Blüthgen et al. (2007) showed that some indices may be particularly influenced by the matrix dimensions, and hence web asymmetry may serve as a kind of correction index. *Extinction slopes* (for lower and higher level) are hyperbolic fits to a simulated extinction sequence of the network, which causes secondary extinctions in the other trophic level (only for networks with strong dependence). The idea was presented by Memmott et al. (2004) and we include this rather rough measure as a simple implementation (see `?second.extinct` for specification of simulations and `?slope.bipartite` for details on fitting of the hyperbolic curve). Finally, *degree distributions* (for both trophic levels) have been heralded by Jordano et al. (2003) and Montoya et al. (2006) as being best described by truncated power laws, rather than exponential of simple power law functions. We fit all three, but also notice that many networks provide only 4 to 5 levels of degrees, so that a non-linear fit to so few data points gives us little confidence in its meaning, and often the fitting does not converge, due to singular gradients.

Some of the indices calculated by `networklevel` may be ecologically uninformative because they are driven by either constraints in web dimensions or are a consequence of the lognormal distribution of species abundances (e.g. nestedness). This is not to say that for specific questions these indices are not important, just to caution that in some cases statistical artefacts may confound an index' intended meaning. We can investigate this by constructing random webs, e.g. by employing Patefield's `r2dtable` algorithm (Fig. 3).

These findings give us some hope that the observed pattern are not a mere artefact of species distributions and web dimensions. There are many dif-

ferent ways to construct a null model, and this is only one of them (Vázquez and Aizen, 2003). We provide two further null models, `shuffle.web` and `swap.web`. The former simply shuffles the observed values within the matrix under the constraint of retaining dimensionality; the latter constrains both marginal sums (as does `r2dtable`) and connectance (i.e. number of non-zero entries in the matrix). As observed connectance is much lower than connectance in random marginal-sum-constrained networks, maintaining connectance implies an ecological mechanism (such as co-evolution of pollinator and plant, body size-relationships between prey and predator, and so forth).

Although we tried to incorporate many descriptors of networks, there are certainly several missing. For example, the social literature has put much emphasis on betweenness and centrality, concepts that we find difficult to interpret in the context of bipartite (=two-mode) ecological networks. Some of these functions are implemented in the R-package `sna` (Social Network Analysis Butts, 2007), which can be accessed after transforming the bipartite data to one-mode graphs using `bipartite`'s function `as.one.mode`. Others can be calculated using the freeware Pajek (<http://vlado.fmf.uni-lj.si/pub/networks/pajek/>; Batagelj and Mrvar, 2003). We made, as yet, no attempt to include indices that use sophisticated optimisation algorithms (e.g. Guimerà and Amaral 2005's modularity index or Clauset et al. 2008's hierarchical structure functions), mainly because of time limitations, but also because they draw heavily on computer resources. Contributions to `bipartite` are welcomed, in order to make further progress in the issues awaiting networks in ecology (Bascompte, 2007).

Bibliography

- J. Bascompte. Networks in ecology. *Basic and Applied Ecology*, 8:485–490, 2007.
- V. Batagelj and A. Mrvar. Pajek - analysis and visualization of large networks. In M. Jünger and P. Mutzel, editors, *Graph Drawing Software*, pages 77–103. Springer, Berlin, 2003.
- N. Blüthgen, F. Menzel, and N. Blüthgen. Measuring specialization in species interaction networks. *BMC Ecology*, 6(9):12, 2006.
- N. Blüthgen, F. Menzel, T. Hovestadt, B. Fiala, and N. Blüthgen. Specialization, constraints and conflicting interests in mutualistic networks. *Current Biology*, 17:1–6, 2007.
- C. T. Butts. *sna: Tools for Social Network Analysis*, 2007. URL <http://erzuli.ss.uci.edu/R.stuff>. R package version 1.5.

```
>null.t.test(mosquin1967, index=c("generality", "vulnerability", "cluster coefficient", "H2", "ISA",
+ "SA"), nrep=2, N=20)
```

	obs	null mean	lower CI	upper CI	t	P
generality	2.6773063	4.22916358	4.15161471	4.30671245	41.88423	3.497206e-20
vulnerability	4.1143452	7.21792806	7.07648249	7.35937363	45.92490	6.178095e-21
cluster coefficient	0.1363636	0.24545455	0.22667908	0.26423001	12.16108	2.067759e-10
H2	0.4964885	0.14030524	0.12975531	0.15085517	-70.66398	1.800214e-24
interaction strength asymmetry	0.1607192	0.05951281	0.05190249	0.06712313	-27.83425	7.291115e-17
specialisation asymmetry	-0.1978587	-0.16827052	-0.20874785	-0.12779319	1.52996	1.425060e-01

Figure 3: Illustration of the `null.t.test`-function with selected network indices.

- A. Clauset, C. Moore and M. E. J. Newman. Hierarchical structure and the prediction of missing links in networks. *Nature*, 453:98–101, 2008.
- R. Guimerà and L. A. N. Amaral. Functional cartography of complex metabolic networks. *Nature*, 433: 895–900, 2005.
- P. Jordano, J. Bascompte, and J. M. Olesen. Invariant properties in coevolutionary networks of plant-animal interactions. *Ecology Letters*, 6:69–81, 2003.
- T. M. Lewinsohn, P. I. Prado, P. Jordano, J. Bascompte, and J. M. Olesen. Structure in plant-animal interaction assemblages. *Oikos*, 113(1):174–184, 2006.
- J. Memmott, N. M. Waser, and M. V. Price. Tolerance of pollination networks to species extinctions. *Proceedings of the Royal Society*, 271:2605–2611, 2004.
- J. M. Montoya, S. L. Pimm, and R. V. Solé. Ecological networks and their fragility. *Nature*, 442:259–264, 2006.
- S. Pimm. *Food Webs*. Chicago University Press, Chicago, 1982/2002.
- J. M. Tylianakis, T. Tscharrntke, and O. T. Lewis. Habitat modification alters the structure of tropical host-parasitoid food webs. *Nature*, 445:202–205, 2007.
- D. P. Vázquez and M. A. Aizen. Null model analyses of specialization in plant-pollinator interactions. *Ecology*, 84(9):2493–2501, 2003.
- D. J. Watts and S. Strogatz. Collective dynamics of ‘small-world’ networks. *Nature*, 393:440–442, 1998.

Carsten F. Dormann & Bernd Gruber
 Department of Computational Landscape Ecology
 UFZ Centre for Environmental Research
 Permoserstr. 15
 04318 Leipzig, Germany
 Email: carsten.dormann@ufz.de

Jochen Fründ
 Agroecology
 University of Göttingen
 Waldweg 26
 37073 Göttingen
 Germany