Methods in Ecology and Evolution

British Ecological Society

Methods in Ecology and Evolution 2015 doi: 10.1111/2041-210X.12334

Clustering methods differ in their ability to detect patterns in ecological networks

Jean-Benoist Leger[1,2,3,4], Jean-Jacques Daudin[1,2,†] and Corinne Vacher[3,4*,†]

[1]*INRA, UMR 518 MIA, Paris, France;* [2]*AgroParisTech, UMR 518 MIA, Paris, France;* [3]*INRA, UMR 1202 BIOGECO, F-33610 Cestas, France; and* [4]*University of Bordeaux, BIOGECO, UMR 1202, F-33615 Pessac, France*

Summary

1. Network ecology has been an extraordinarily fertile field of research over the last 20 years. Its ultimate goal is to understand how the complex systems of interdependent species assemble, function and evolve. Here, we aimed to help ecologists to select the best methods for detecting subgroups of highly interacting species (usually referred to as compartments or modules) in bipartite networks (e.g. plant–pollinator networks, host–parasite networks), because these subgroups may reveal the processes underlying the assembly of the network and may influence its stability.

2. We simulated several thousand bipartite ecological networks and we compared seven methods of network clustering in terms of their ability to retrieve the number and the composition of species subgroups.

3. Among the seven methods compared, we found that the edge-betweenness algorithm was the best option for binary networks. The stochastic block model was the best method for weighted networks. Modularity maximization, the most popular clustering method in ecology, was among the three best methods in both cases.

4. We thus provide ecology researchers with precise advice concerning the best choice of network clustering method, according to the type of data collected. We also provide the code for simulating bipartite networks and clustering them, in order to facilitate future methodological comparisons.

Key-words: clustering, compartment, ecological network, module

Introduction

Networks are an extraordinary tool for portraying the complexity of biological systems. In ecology, networks are used for representing the complex interactions (*network links*) between species (*network nodes*). Food webs are the ecological networks representing predator/prey interactions, whereas bipartite ecological networks usually represent long-lasting, intimate interactions between two sets of species (hereafter referred to as the basal species and the top species; e.g. plant species and their pollinators, or host species and their parasites). Both food webs and bipartite ecological networks usually contain several subgroups of species, within which the similarity in interaction patterns is higher (Allesina & Pascual 2009) or interactions are denser (Newman & Girvan 2004). This latter kind of subgroup is usually referred to as compartment (Krause *et al.* 2003) or module (Olesen *et al.* 2007).

The detection of species subgroups in real ecological networks is important to ecologists for three reasons. First, their visual representation (i.e. a plot with nodes pertaining to the same subgroup having the same colour) provides a simplified picture of the network (Allesina & Pascual 2009). Second, they may suggest the processes underlying the assembly of the network (Vacher, Piou & Desprez-Loustau 2008; Rezende *et al.* 2009; Krasnov *et al.* 2012). Third, as suggested by mathematical models (Thébault & Fontaine 2010; Stouffer & Bascompte 2011), they may influence the functioning (including stability, in particular) of the network. If their influence is confirmed by experimental studies (Rip *et al.* 2010), they could thus become a relevant target for ecosystem conservation and restoration programmes.

Since the publication of the seminal article by Olesen *et al.* (2007), modularity maximization with a simulated annealing algorithm (Guimerà & Amaral 2005; Guimerà, Sales-Pardo & Amaral 2007) has been the most widely used method for detecting subgroups of highly interacting species in ecological networks. It was adapted for application to bipartite networks (Barber 2007; Guimerà, Sales-Pardo & Amaral 2007; Thébault 2013; Dormann & Strauß 2014). Other methods for detecting modules in unipartite or bipartite networks (Fortunato 2010; Leger, Vacher & Daudin 2013), originating from various fields of research (physics, mathematics, statistics, computer science), have been less used in ecology.

Here, we compare the ability of seven graph clustering methods to detect subgroups of highly interacting species in bipartite ecological networks. These methods, which correspond to a subset of the methods described in detail by Leger, Vacher & Daudin (2013), are representative of the various approaches that can be used for clustering networks. The seven methods are the modularity maximization method (Newman 2006), the edge-betweenness algorithm (Girvan &

*Correspondence author. E-mail: corinne.vacher@pierroton.inra.fr
†These authors contributed equally to this work.

Newman 2002), two variants of the Markov cluster algorithm (Leger, Vacher & Daudin 2013), two spectral clustering methods [Ng-normalized spectral clustering (Luxburg 2007) and absolute eigenvalues spectral clustering (Rohe, Chatterjee & Yu 2011)] and the stochastic block model (Mariadassou, Robin & Vacher 2010). We compared the performances of these methods, by applying them to simulated ecological networks containing known subgroups. We compared the ability of each of the seven methods to retrieve the number of subgroups and the composition of subgroups for both weighted and binary networks. The simulation algorithm, derived from that of Thébault & Fontaine (2010), was not related to any clustering method and therefore did not favour any one method over the others. We first compared the clustering methods for intermediate ecological networks (simulated by fixing the network properties to their mean ecological value). We then assessed the robustness of our comparison to variations in network properties, by allowing these properties to vary within a range slightly larger than the ecological range.

Methods

CHOICE OF NETWORK CLUSTERING METHODS AND GENERAL IMPLEMENTATION FRAMEWORK

The seven methods compared are of two kinds (Allesina & Pascual 2009; Leger, Vacher & Daudin 2013): those that detect sets of highly interacting nodes (hereafter called *communities*) and those that detect sets of nodes with similar interaction patterns (hereafter called *structurally homogeneous subsets*). The two kinds of methods belong to different lines of research, but the differences between them are not entirely clear-cut, because nodes within a community tend to have similar interaction patterns. Both kinds of methods may thus be used to detect subgroups of highly interacting species in bipartite ecological networks (Leger, Vacher & Daudin 2013).

Four of the seven methods tested were initially developed for the detection of communities: the modularity maximization method (Newman 2006), the edge-betweenness algorithm (Girvan & Newman 2002), the Markov cluster algorithm (van Dongen 2000) and the Ng-normalized spectral clustering method (Luxburg 2007). The other three were initially developed for the detection of structurally homogeneous clusters: a reparameterized version of the Markov cluster algorithm (Leger, Vacher & Daudin 2013), the absolute eigenvalues spectral clustering method (Rohe, Chatterjee & Yu 2011) and the stochastic block model (Mariadassou, Robin & Vacher 2010). Due to the large number of simulated networks, the modularity maximization method was implemented here with a fast algorithm (Newman 2006). Modularity maximization with a simulated annealing algorithm (like in Guimerà & Amaral 2005; Guimerà, Sales-Pardo & Amaral 2007) is more time-consuming (a couple thousand more) and was thus only applied to a subset of the networks.

We implemented the seven methods see Appendix S1 for details, based on their original descriptions, by using a combination of m-code (MATLAB language) and C++ (both used in GNU OCTAVE), and some functions of the C-library IGRAPH (Csardi & Nepusz 2006). The code is available online (https://gitlab.crans.org/leger/clustering_methods_comparison). Two types of implementation were used for each method. We first allowed the method (or the associated criteria) to select the optimal number of subgroups. We then forced the method to perform the clustering analysis with the true number of subgroups. The adjustments required to perform the two types of implementation are described in Appendix S1.

SIMULATION OF ECOLOGICAL BIPARTITE NETWORKS

We adapted the algorithm developed by Thébault & Fontaine (2010), to simulate weighted bipartite networks (Appendix S1). The algorithm parameters were the number of top species n_B, the number of basal species n_T, the total number of links n_L, the total weight of all links n_W, the number of subgroups g, the degree of compartmentalization p_{comp} and the degree of nestedness p_{nest}.

An analysis of 47 ecological networks taken from the Interaction Web Database (Table S1) revealed that some pairs of parameters (from the list n_B, n_T, n_L and n_W) were highly correlated. Thus, had we allowed these parameters to vary independently, the simulated networks generated would have been different from real networks. We, therefore, performed a reparameterization (Appendix S1). The four new parameters, which were almost independent, were: the number of all possible links $k_S = n_B n_T$, the ratio $k_R = n_B/n_T$ of the number of basal species to the number of top species, the mean weight of edges $k_W = n_W/n_L$ and the parameter $k_L = n_L/k_S^{0.63}$. The parameter k_L is linked to the network connectance $C = n_L/k_S$ (i.e. the proportion of possible links between species that actually occur) by the relation $C = k_L/k_S^{0.37}$. For simplicity, the number of all possible links k_S is referred to as network size and the parameter k_L is referred to as network connectance.

Based on the 47 real ecological networks (Table S1), we estimated the ecological ranges for these four new parameters (Table S2). The ranges used for simulations were slightly larger than the ecological ranges. It was not possible to give an ecological range for g, p_{comp} and p_{nest}, because these three parameters cannot be measured in real networks. We therefore simply varied them over a large range of values.

We also used the 47 ecological networks taken from the Interaction Web Database (Table S1) to calculate a geometric mean value for the four new parameters (Table S2). These values are referred to as intermediate ecological values. An intermediate network simulated with such values had 64 top species, 18 basal species, 120 edges and a total weight of 1018. We arbitrarily chose $g = 4$, $p_{comp} = 0{\cdot}6$ and $p_{nest} = 0{\cdot}5$ as intermediate values. This choice corresponds to a strongly compartmentalized and nested network with four subgroups.

To validate our simulation approach, we then compared simulated networks to real ecological networks for five topological properties (Table S1): the cumulative distribution of degrees for each level of the network (Fig. S1), the frequency distribution of dependence for each level of the network (Fig. S2) and the frequency distribution of the asymmetry values of dependences (Fig. S3), as defined by Bascompte, Jordano & Olesen (2006). The distributions in simulated networks were similar to those observed in previous studies (Jordano, Bascompte & Olesen 2003; Bascompte, Jordano & Olesen 2006) for real ecological networks.

Finally, we compared the performance of clustering methods for intermediate ecological networks by first simulating 1000 weighted networks, keeping all the parameters fixed at their intermediate value. We then assessed the robustness of ranking by the methods to variations in network properties, by simulating networks with all the parameters fixed to the intermediate value other than one, which was allowed to vary within the simulated range. This variable was allowed to take 10 different values (except for the number of groups, which took only four values). We simulated 100 networks for each parameter combination. This simulation design resulted in 6400 weighted networks. We then

obtained the binary versions of all networks, by replacing positive weights by unitary weights, and repeated the comparison of clustering methods.

CRITERIA USED FOR COMPARISONS OF THE EFFICACY OF THE CLUSTERING METHODS

We first compared the ability of the methods to retrieve the true number of subgroups, by calculating the ratio of the number of groups estimated by the clustering method to the expected number of subgroups. A ratio >1 indicates that the clustering method overestimates the number of subgroups, whereas a ratio below 1 indicates that the method underestimates the number of subgroups. The expected number of subgroups was g in the case of methods initially developed for the detection of communities. The expected number of subgroups was $2 \times g$ in the case of methods initially developed for the detection of structurally homogeneous subsets. Indeed, for bipartite networks, this type of method splits communities into two subgroups (Leger, Vacher & Daudin 2013), one containing the basal species and the other, the top species.

We then assessed the ability of each clustering method to retrieve the composition of subgroups. For each simulated network, we compared the composition of subgroups delimited by the clustering method with that of the true subgroups, using the adjusted Rand index (Hubert & Arabie 1985). The Rand index is a measure of the similarity between two partitions. Let P_{obs} be the partition obtained by a given clustering method and P_{true} the true partition. Two species may be (a) in the same subgroup according to both P_{obs} and P_{true}, (b) in the same subgroup according to P_{obs} but not P_{true}, (c) in the same subgroup according to P_{true} but not P_{obs} or (d) in different subgroups according to both P_{obs} and P_{true}. Let a–d be the corresponding numbers of cases. The Rand index is defined by $R = (a + d)/(a + b + c + d)$, the proportion of cases in agreement. It can be adjusted to take its values in the range $[-1, 1]$. A value of 0 corresponds to an expectation of random clustering. A positive value indicates clustering better than would be expected by chance, and a value of 1 indicates perfect clustering. A negative value indicates clustering that is worse than would be expected by chance. We calculated two adjusted Rand indices, one indicating the quality of clustering for the basal network level and the other indicating the quality of clustering for the higher network level. All simulated networks and raw results are available online (https://gitlab.crans.org/leger/clustering_methods_comparison).

Results

COMPARISON OF THE GRAPH CLUSTERING METHODS FOR WEIGHTED NETWORKS

For weighted ecological networks with intermediate properties, the stochastic block model was the clustering method giving the most precise estimate of the number of subgroups (Fig. 1a). According to the adjusted Rand index (Hubert & Arabie 1985), it also retrieved their composition very effectively (Fig. 2a,b). It retrieved the composition of basal species subgroups (Fig. 2a) more effectively than that of top species subgroups (Fig. 2b), and this trend was also observed for all other methods.

The edge-betweenness algorithm (with the modularity criterion for selecting subgroup number) greatly overestimated the number of subgroups (by a factor of 2–6, on average) in weighted networks (Fig. 1a), but was otherwise highly effective for retrieving their composition (Fig. 2a,b). These findings indicate that the edge-betweenness algorithm generates split subgroups in weighted networks. Disappointingly, constraining the algorithm to produce a number of subgroups closer to the expected number (to facilitate the interpretation of clustering results, for instance), greatly decreases the ability of the edge-betweenness algorithm to retrieve subgroup composition (Fig. S4).

Modularity maximization and one of the two variants of the Markov Chain Clustering algorithm ($MCL_{1/10}$) were the models with the best performance after the stochastic block model, in terms of their ability to retrieve species subgroups. Both methods satisfactorily retrieved the composition of species subgroups (Fig. 2a,b). $MCL_{1/10}$ also effectively retrieved the number of subgroups (Fig. 1a). Modularity maximization implemented with a fast algorithm overestimated the number of subgroups (by a factor of 2–4, on average; Fig. 1a), and constraining the method to retrieve fewer subgroups did not decrease its ability to retrieve the composition of species subgroups (Fig. S4). Modularity maximization with a simulated annealing algorithm also overestimated the number of groups

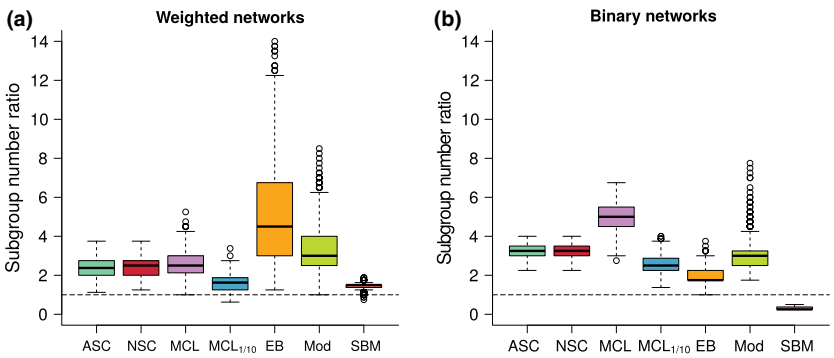

Fig. 1. Ratio of the estimated number of subgroups to the expected number of subgroups in the case of (a) weighted networks and (b) binary networks. In each case, the results are based on 1000 ecological networks simulated with intermediate values for ecological properties. The seven clustering methods compared here were the modularity maximization method (Mod), the edge-betweenness algorithm (EB), two variants of the Markov cluster algorithm (MCL and $MCL_{1/10}$), the Ng-normalized spectral clustering method (NSC), the absolute eigenvalues spectral clustering method (ASC) and the stochastic block model (SBM). The horizontal dashed line indicates the value of 1. Ratios >1 indicate that the clustering method overestimates the number of subgroups and ratios below 1 indicate that the method underestimates the number of subgroups.

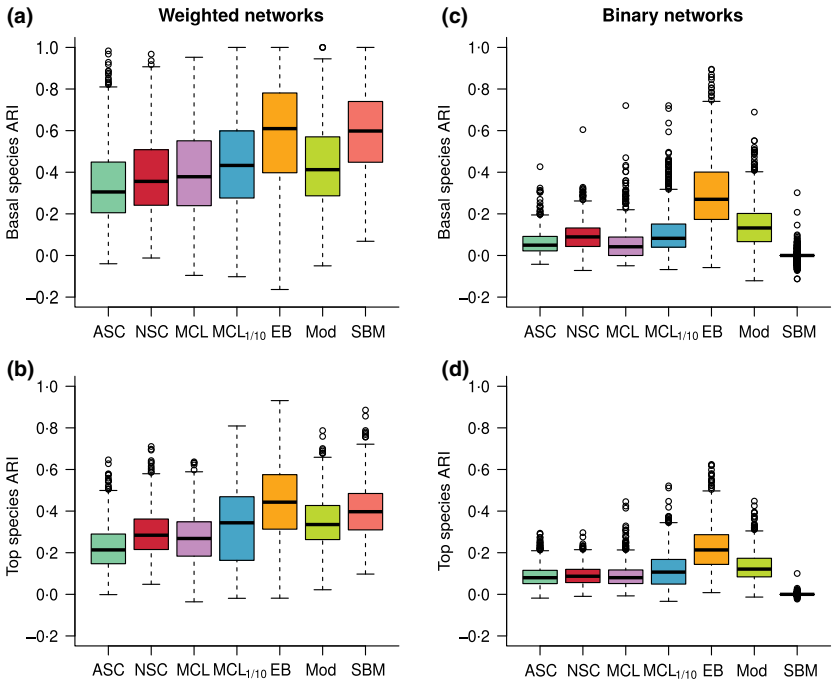

Fig. 2. Ability of the clustering methods to retrieve the composition of species subgroups, as assessed by the adjusted Rand index (ARI), in the case of (a) basal species in weighted networks, (b) top species in weighted networks, (c) basal species in binary networks and (d) top species in binary networks. In each case, the results are based on 1000 ecological networks simulated with intermediate ecological properties. The abbreviations for the clustering methods are as in Fig. 1. An adjusted Rand index of 1 indicates a perfect retrieval of the subgroup composition and an adjusted Rand index close to 0 indicates that the clustering is close to random.

and retrieved the composition of subgroups slightly better than modularity with a fast algorithm (Fig. S5).

The variation of network properties modified the ability of the methods to retrieve the number of subgroups. All methods other than the stochastic block model and one variant of the Markov cluster algorithm ($MCL_{1/10}$) tended to overestimate the number of subgroups, particularly for large networks (Fig. 3) or if the true number of groups was low (Fig. S6). This was particularly true for the edge-betweenness algorithm and the modularity maximization method. The edge-betweenness algorithm was also extremely sensitive to the degree of compartmentalization and nestedness of the network (Fig. 3).

The variation of network properties also modified the ability of the methods to retrieve subgroup composition. As expected, all methods performed better in terms of subgroup retrieval if the degree of compartmentalization was high (Fig. 3). Very high levels of nestedness modified the ranking of the methods. In this situation, the modularity maximization method and Ng-normalized spectral clustering performed slightly better than the other methods (Fig. 3). Finally, for most clustering methods, the ability to retrieve subgroups increased with the true number of subgroups and network connectance, but decreased with network size (Fig. 3).

On the basis of these results, we conclude that the three best clustering methods for weighted networks are, in descending order of performance: the stochastic block model, one variant of the Markov Chain Clustering algorithm ($MCL_{1/10}$) and the modularity maximization method. The edge-betweenness algorithm (with the modularity criterion for selecting subgroup number) appears to be a poorer choice in this context, because it overestimates the number of subgroups, cannot retrieve the composition of subgroups when the number of subgroups is constrained to lower values and is highly sensitive to network properties.

COMPARISON OF THE GRAPH CLUSTERING METHODS FOR BINARY NETWORKS

The binarization of the simulated ecological networks greatly decreased the ability of all methods to retrieve the composition of species subgroups (Fig. 2). For binary ecological networks with average properties, the edge-betweenness algorithm (with the modularity criterion for selecting subgroup number) was the best method for retrieving the number of subgroups (Fig. 1b) and their composition, for both basal species (Fig. 2c) and top species (Fig. 2d). The variation of network properties did not affect the top ranking of the edge-betweenness algorithm for binary networks. This algorithm gave the best results in all situations except that of high network connectance (Fig. S7).

Modularity maximization and one variant of the Markov Chain Clustering algorithm ($MCL_{1/10}$) also gave satisfactory results for retrieval of the number of subgroups (Fig. 1b) and the retrieval of subgroup composition (Fig. 2c,d). Modularity maximization was slightly better than $MCL_{1/10}$ for retrieving the composition of subgroups for a wide range of parameters (Fig. S7).

The stochastic block model had the poorest performance with binary networks, as it was able to delimit only the two trophic levels. Its performance was improved by entering the true number of subgroups as an input parameter (Fig. S4), suggesting that the integrated completed likelihood (ICL) criterion (Daudin, Picard & Robin 2008) used for the selection of group number in this method is not suitable for studies of bipartite binary ecological networks.

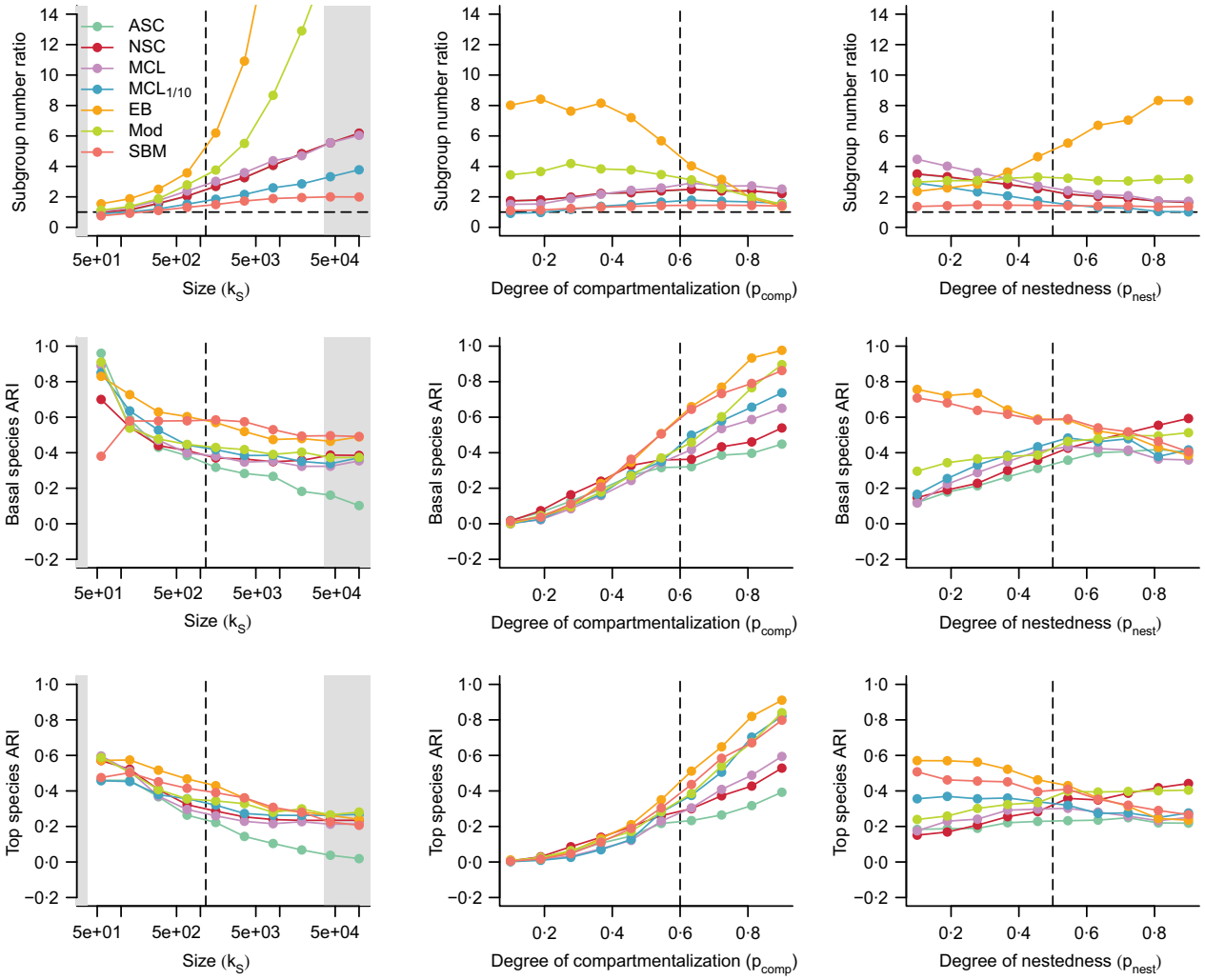

Fig. 3. Effect of the variation of three network properties (size, degree of compartmentalization and nestedness) on the performance of clustering methods in the case of weighted networks. The performance of each method was estimated by determining its ability to retrieve the true number of subgroups (as in Fig. 1) and its ability to retrieve the composition of basal and top species subgroups (as in Fig. 2). For each parameter value, the results are based on 100 simulated ecological networks. The vertical dashed line indicates the parameter value used to obtain the results presented in Figs 1 and 2. The abbreviations used for the clustering methods are as in Fig. 1. For network size, the white area indicates the size range for real ecological networks (based on the real networks listed in Table S1).

On the basis of these results, we conclude that the three best clustering methods for binary networks are the edge-betweenness algorithm (with the modularity criterion for selecting subgroup number), followed by the modularity maximization method and one variant of the Markov Chain Clustering algorithm ($MCL_{1/10}$). The stochastic block model, for which the number of subgroups is defined by the ICL criterion, should not be used in this situation.

COMPARISON OF THE GRAPH CLUSTERING METHODS IN TERMS OF COMPUTATION TIME

The clustering methods also differed in terms of their computer running time (Table 1). This is an important criterion for method selection, particularly when large sets of networks are to be analysed. The modularity maximization method, when implemented with a fast algorithm (Newman 2006), was the fastest method. The stochastic block model was the slowest.

Discussion

We obtained a clear ranking of the methods in terms of their ability to retrieve subgroups of highly interacting species from ecological bipartite networks, making it possible to provide ecology researchers with precise advice concerning the best choice of graph clustering method.

Among the seven methods compared, the stochastic block model emerged as the best option for weighted networks. It precisely retrieved the composition of species subgroups and precisely estimated their number. This second feature is important because a clustering method that generates split subgroups by overestimating the number of subgroups, while retrieving their composition accurately may lead the researcher to

Table 1. Mean computer running time for the clustering of a single intermediate ecological network (left) and 1000 intermediate ecological networks (right)

	1 Network		1000 Networks	
	Weighted	Binary	Weighted	Binary
Mod	13 ms	11 ms	13 s	11 s
$MCL_{1/10}$	31 ms	33 ms	31 s	33 s
MCL	32 ms	31 ms	32 s	31 s
NSC	83 ms	85 ms	1 min 23 s	1 min 25 s
ASC	90 ms	95 ms	1 min 30 s	1 min 35 s
EB	779 ms	31 ms	12 min 59 s	31 s
SBM	43 s	1 min 21 s	11 h 49 min	22 h 26 min

The computer running time is the time taken to execute the analysis on a single-processor computer running only one job at a time. The clustering methods are classified from the fastest to the slowest on the basis of the results obtained for weighted networks. The abbreviations for the clustering methods are the same as in Fig. 1.

interpret ecological discontinuities that do not actually exist. This would be a considerable waste of time and would give rise to false ecological theories. The only drawback of the stochastic block model was its slowness, particularly for large networks. This may not be an issue when analysing a couple of real ecological networks, but it is more problematic for the analysis of several thousand simulated networks.

The edge-betweenness algorithm (with the modularity criterion for selecting subgroup number) consistently gave the best results for binary networks of any of the clustering methods tested here. It retrieved the precise composition of the subgroups, and only moderately overestimated the number of subgroups.

Modularity maximization, the most popular clustering method in ecology, gave good results for both weighted and binary networks, although it also moderately overestimated the number of subgroups. It was among the three best methods for a wide range of ecological networks. With the efficient algorithmic version of modularity maximization (Newman 2006) used here, this method also turned out to be the fastest of the seven methods tested here. Modularity maximization should thus be considered as an alternative to the stochastic block model in situations in which several thousand weighted networks must be analysed in a reasonable amount of time. The new MODULAR (Marquitti *et al.* 2013) software suite could be used in such situations, for example.

One variant of the Markov Chain Clustering algorithm ($MCL_{1/10}$) also emerged as a good option on the basis of our results, but this method has to prove its worth because it is new (developed by Leger, Vacher & Daudin 2013) and has never been applied to real networks until now.

Our results also showed that all clustering methods performed less well with binary networks than with weighted networks. We thus recommend the collection and analysis of quantitative interaction data, to ensure the discovery of the real network structure.

Finally, our methodological comparison generated an unexpected finding, as it revealed that all clustering methods performed asymmetrically. Species subgroups were generally more accurately retrieved for basal species than for top species. The clusters of basal species identified in weighted bipartite networks are thus the most robust and can be interpreted with more confidence than other clusters.

This asymmetry may be accounted for by the smaller number of basal species in ecological networks. The amount of information per basal species within a network is thus greater than the amount of information per top species, leading to a better classification of basal species. Interestingly, an asymmetric influence of evolutionary history has often been found in bipartite antagonistic networks, with phylogeny better accounting for the composition of basal species subgroups than for that of top species subgroups (Ives & Godfray 2006; Vacher, Piou & Desprez-Loustau 2008; Krasnov *et al.* 2012; Elias, Fontaine & Van Veen 2013). For instance, in a tree-parasitic fungus network, a significant association between species subgroups and phylogeny was found for tree species, but not for parasite species (Vacher, Piou & Desprez-Loustau 2008). Similar results have also been obtained for a couple of dozen mammal–flea networks. Closely related hosts tend to occur in the same subgroups of highly interacting species, whereas the distribution of parasite lineages between subgroups is rarely anything other than random (Krasnov *et al.* 2012). Ecological and evolutionary processes have been proposed to account for this asymmetric pattern. For instance, it has been hypothesized that exploitative competition shapes antagonistic interactions more strongly than apparent competition (Krasnov *et al.* 2012; Elias, Fontaine & Van Veen 2013). Our results suggest that the observed pattern may also be, at least partly, a methodological artefact, due to the less accurate classification of antagonist species into subgroups. Further studies are thus required to assess the true effect of phylogeny on the compartmentalized structure of bipartite antagonistic networks.

In conclusion, according to our results, ecologists should favour the edge-betweenness algorithm (with the modularity

Table 2. Advice concerning the best choice of clustering method for bipartite ecological networks

	Binary network	Weighted network
Best clustering method	*Edge-betweenness algorithm*, with the modularity criterion for selecting group number	*Stochastic Block Model*, with the ICL criterion for selecting group number
Second best clustering method	*Modularity maximization method*, implemented with a fast algorithm	*Modularitymaximizationmethod*, implemented with a fast algorithm

criterion for selecting group number) when they wish to retrieve subgroups of highly interacting species from binary bipartite networks (Table 2). For weighted bipartite networks, they should prefer the stochastic block model (Table 2), which accurately estimates the number of subgroups and retrieves their composition effectively in this context. Unfortunately, this method is slow, particularly for large networks. Modularity maximization is a good alternative (Table 2), given that recent developments have made it possible to analyse thousands of ecological bipartite networks in a reasonable amount of time. These three best clustering methods will have to be compared in future studies to other existing methods that have not been included in the present comparison. We provide the code for simulating bipartite networks and clustering them (https://gitlab.crans.org/leger/clustering_methods_comparison), in order to facilitate future methodological comparisons.

Acknowledgements

We thank the INRA and the international interdisciplinary PhD program 'Frontiers in Life Sciences' (ED474 FDV, Universities Paris Descartes and Paris Diderot) for funding Jean-Benoist Leger through a *Contrat Jeune Scientifique* (CJS). We also thank Stéphane Robin and Bastien Castagneyrol and two anonymous reviewers for their helpful comments on the manuscript and Julie Sappa from Alex Edelman & Associates for English proofreading.

Data accessibility

Code and simulated data are available at: https://gitlab.crans.org/leger/clustering_methods_comparison.

References

Allesina, S. & Pascual, M. (2009) Food web models: a plea for groups. *Ecology Letters*, **12**, 652–662.

Barber, M. (2007) Modularity and community detection in bipartite networks. *Physical Review E*, **76**, 1–9.

Bascompte, J., Jordano, P. & Olesen, J.M. (2006) Asymmetric coevolutionary networks facilitate biodiversity maintenance. *Science*, **312**, 431–433.

Csardi, G. & Nepusz, T. (2006) The igraph software package for complex networks. *InterJournal Complex Systems*, **1695**, 1–9.

Daudin, J.J., Picard, F. & Robin, S. (2008) A mixture model for random graphs. *Statistics and Computing*, **18**, 173–183.

van Dongen, S.M. (2000). Graph clustering by flow simulation. Available at: http://igitur-archive.library.uu.nl/dissertations/1895620/UUindex.html

Dormann, C.F. & Strauß, R. (2014) A method for detecting modules in quantitative bipartite networks. *Methods in Ecology and Evolution*, **5**, 90–98.

Elias, M., Fontaine, C. & Van Veen, F.J. (2013) Evolutionary history and ecological processes shape a local multilevel antagonistic network. *Current Biology*, **23**, 1355–1359.

Fortunato, S. (2010) Community detection in graphs. *Physics Reports*, **486**, 75–174.

Girvan, M. & Newman, M.E.J. (2002) Community structure in social and biological networks. *Proceedings of the National Academy of Sciences of the United States of America*, **99**, 7821–7826.

Guimerà, R. & Amaral, L.A.N. (2005) Functional cartography of complex metabolic networks. *Nature*, **433**, 895–900.

Guimerà, R., Sales-Pardo, M. & Amaral, L. (2007) Module identification in bipartite and directed networks. *Physical Review E*, **76**, 036102.

Hubert, L. & Arabie, P. (1985) Comparing partitions. *Journal of Classification*, **2**, 193–218.

Ives, A.R. & Godfray, H.C.J. (2006) Phylogenetic analysis of trophic associations. *The American Naturalist*, **168**, E1–E14.

Jordano, P., Bascompte, J. & Olesen, J.M. (2003) Invariant properties in coevolutionary networks of plant–animal interactions. *Ecology Letters*, **6**, 69–81.

Krasnov, B.R., Fortuna, M.A., Mouillot, D., Khokhlova, I.S., Shenbrot, G.I. & Poulin, R. (2012) Phylogenetic signal in module composition and species connectivity in compartmentalized host–parasite networks. *The American Naturalist*, **179**, 501–511.

Krause, A.E., Frank, K.A., Mason, D.M., Ulanowicz, R.E. & Taylor, W.W. (2003) Compartments revealed in food-web structure. *Nature*, **426**, 282–285.

Leger, J.-B., Vacher, C. & Daudin, J.-J. (2013). Detection of structurally homogeneous subsets in graphs. *Statistics and Computing*, online. Available at: http://link.springer.com/10.1007/s11222-013-9395-3

Luxburg, U. (2007) A tutorial on spectral clustering. *Statistics and Computing*, **17**, 395–416.

Mariadassou, M., Robin, S. & Vacher, C. (2010) Uncovering latent structure in valued graphs: a variational approach. *The Annals of Applied Statistics*, **4**, 715–742.

Marquitti, F.M.D., Roberto, P., Guimaraes, P.R. Jr, Pires, M.M. & Bittencourt, L.F. (2013) MODULAR: Software for the autonomous computation of modularity in large network sets. *Ecography*, **37**, 221–224.

Newman, M.E.J. (2006) Modularity and community structure in networks. *Proceedings of the National Academy of Sciences of the United States of America*, **103**, 8577–8582.

Newman, M. & Girvan, M. (2004) Finding and evaluating community structure in networks. *Physical Review E*, **69**, 026113.

Olesen, J.M., Bascompte, J., Dupont, Y.L. & Jordano, P. (2007) The modularity of pollination networks. *Proceedings of the National Academy of Sciences of the United States of America*, **104**, 19891–19896.

Rezende, E.L., Albert, E.M., Fortuna, M.A. & Bascompte, J. (2009) Compartments in a marine food web associated with phylogeny, body mass, and habitat structure. *Ecology Letters*, **12**, 779–788.

Rip, J.M.K., McCann, K.S., Lynn, D.H. & Fawcett, S. (2010) An experimental test of a fundamental food web motif. *Proceedings of the Royal Society B*, **277**, 1743–1749.

Rohe, K., Chatterjee, S. & Yu, B. (2011) Spectral clustering and the high-dimensional stochastic blockmodel. *Annals of Statistics*, **39**, 1878–1915.

Stouffer, D.B. & Bascompte, J. (2011) Compartmentalization increases food-web persistence. *Proceedings of the National Academy of Sciences of the United States of America*, **108**, 3648–3652.

Thébault, E. (2013) Identifying compartments in presence–absence matrices and bipartite networks: insights into modularity measures. *Journal of Biogeography*, **40**, 759–768.

Thébault, E. & Fontaine, C. (2010) Stability of ecological communities and the architecture of mutualistic and trophic networks. *Science*, **329**, 853–856.

Vacher, C., Piou, D. & Desprez-Loustau, M.-L. (2008) Architecture of an antagonistic tree/fungus network: the asymmetric influence of past evolutionary history. *PLoS One*, **3**, e1740.

Received 17 October 2014; accepted 15 December 2014
Handling Editor: Tamara Münkemüller

Supporting Information

Additional Supporting Information may be found in the online version of this article.

Fig. S1. Cumulative distribution of degrees in simulated networks for (a) basal species and (b) top species.

Fig. S2. Frequency distribution of dependence in simulated networks for (a) basal species and (b) top species.

Fig. S3. Frequency distribution of asymmetry values in simulated networks.

Fig. S4. Performance of the clustering methods for networks simulated with mean values for ecological parameters, with the methods forced to perform the clustering with the true number of subgroups. See Fig. 2 for legend details.

Fig. S5. Performance of the clustering methods for networks simulated with mean values for ecological parameters. Modularity maximization with a simulated annealing algorithm (ModG) was added to the comparison. See Figs 1 and 2 for legend details.

Fig. S6. Effect of the variation of all network properties on the performance of clustering methods in the case of weighted networks. See Fig. 3 for legend details.

Fig. S7. Effect of the variation of all network properties on the performance of the clustering methods for binary networks. See Fig. 3 for legend details.

Table S1. List of real ecological bipartite networks used to choose the parameter values for network simulation.

Table S2. Parameter value ranges used to simulate networks.

Appendix S1. Algorithm and parameters used to simulate weighted ecological bipartite networks, and implementation details for the network clustering methods.