Article

How to construct the statistic network? An association network of herbaceous plants constructed from field sampling

WenJun Zhang
School of Life Sciences, Sun Yat-sen University, Guangzhou, China; International Academy of Ecology and Environmental Sciences, Hong Kong
E-mail: zhwj@mail.sysu.edu.cn, wjzhang@iaees.org

Received 8 November 2011; Accepted 6 January 2012; Published online 1 June 2012
IAEES

Abstract
In present study I defined a new type of network, the statistic network. The statistic network is a weighted and non-deterministic network. In the statistic network, a connection value, i.e., connection weight, represents connection strength and connection likelihood between two nodes and its absolute value falls in the interval (0, 1]. The connection value is expressed as a statistical measure such as correlation coefficient, association coefficient, or Jaccard coefficient, etc. In addition, all connections of the statistic network can be statistically tested for their validity. A connection is true if the connection value is statistically significant. If all connection values of a node are not statistically significant, it is an isolated node. An isolated node has not any connection to other nodes in the statistic network. Positive and negative connection values denote distinct connection types (positive or negative association or interaction). In the statistic network, two nodes with the greater connection value will show more similar trend in the change of their states. At any time we can obtain a sample network of the statistic network. A sample network is a non-weighted and deterministic network. The statistic network, in particular the plant association network that constructed from field sampling, is mostly an information network. Most of the interspecific relationships in plant community are competition and cooperation. Therefore in comparison to animal networks, the methodology of statistic network is more suitable to construct plant association networks. Some conclusions were drawn from this study: (1) in the plant association network, most connections are weak and positive interactions. The association network constructed from Spearman rank correlation has most connections and isolated taxa are fewer. From net linear correlation, linear correlation, to Spearman rank correlation, the practical number of connections and connectance in the constructed network increases. Network compactness also follows the trend. In addition, as the increase of network compactness and connectance, the portion and number of negative association declines dramatically. (2) In an association (interaction) network, only a few connections follow the linear relationship. Most connections follow the quasi-linear or non-linear relationships. (3) The association networks constructed from partial linear correlation and linear correlation measures are generally scale-free complex networks. The degree of these networks is power low distributed. (4) Isolated species (families, etc.) are likely important in the statistic network. They are the sink species for shaping new network after a community is seriously disturbed. (5) Beween-taxa connections at higher taxonomic level are generally weaker than that at lower taxonomic level.

Keywords statistic network; association; correlation measures; grassland community; herbaceous plants; field sampling.

1 Introduction

Interspecific association always refers to the spatial association between species. Habitat variation causes the variation of species distribution, which results in the interspecific association (Greig-Smith, 1983). Interspecific association occupies an important position in community ecology. It is one of the quantitative and structural characteristics of communities. Interspecific association is the comprehensive representation for interspecific interaction in various habitats. It is important for classification and structure identification of communities. Moreover, interspecific association is the basis for constructing biological networks.

Interspecific association can be determined by field sampling. It has been suggested sampling some designated area and document all such association to construct a network within the area and the sites sampled (Butts, 2009; Schoenly and Zhang, 1999). Moreover, the methods used will completely affect the network structure, and the robustness of the network. The definition and classification of taxa and environment are necessary. Therefore simultaneous analysis of the same system at multiple levels of aggregation is suggested (Butts, 2009).

Association (interaction) strength was always omitted in the past ecological networks (Paine, 1980, 1988, 1992). Actually many forms of association occur at variable rates (Whitehead and Dufault, 1999). Network structures with different connection strengths can vary greatly. It is necessary to determine whether the relationship under study is stable enough over the period of interest (Butts, 2006, 2009). For those relations known to be highly heterogeneous, no single threshold may suffice and a weighted graph representation will frequently be more appropriate. To assess the effectiveness of such approximations and provide concrete, empirically validated guidelines for practice within particular problem domains would be a welcome addition to the literature (Butts, 2006, 2009). The static networks are not able to represent actual ecological networks. Dichotomization of such data not only obscures such variation but also requires selecting a threshold level, the choice of which can substantially alter the properties of the resulting network (Faust, 2007; Onnela et al., 2007).

In this study, I tried to define a new network type and present the methodologies for constructing plant association networks from field sampling.

2 Material and Methods

2.1 Field sampling

The field survey for constructing statistic network was conducted from April to May of 2007 and 2008, at the Zhuhai Campus of Sun Yat-sen University, Zhuhai, China (Fig. 1). The grassland in the campus was selected for field survey, which was naturally shaped through species invasion and establishment since late 1990s.

Totally 150 samples (50 samples in 2007, and 100 samples in 2008), with 1m^2 (1m×1m) each sample, were surveyed. Species and coverage (%) of herbaceous plants were carefully recorded for each sample.

2.2 Methods

In present study I define a new type of network, the statistic network. The statistic network is a weighted and non-deterministic network. In the statistic network, a connection value (connection weight) represents both connection strength and connection likelihood between two nodes and its absolute value falls in the interval (0, 1]. The connection value is expressed as a statistical measure such as correlation coefficient, association coefficient, or Jaccard coefficient, etc. In addition, all connections of a statistic network can be statistically tested for their validity. A connection is true if the connection value is statistically significant. If all connection values of a node are not statistically significant, it is an isolated node. An isolated node has not any connection to other nodes in the statistic network. Positive and negative connection values denote distinct connection types (positive or negative association or interaction). At any time we can obtain a sample network of the statistic network. A sample network is a non-weighted and deterministic network.

In the statistic network, two nodes with the greater connection value will show more similar trend in the change of their states. All connections in a deterministic network are actually those connections occurred with connection values 1 or -1. Therefore in a sense, the deterministic network is a special case of statistic network.

The association (interaction) networks constructed from field sampling, as done in this study and previous studies (Zhang, 2011b), are statistic networks.

Fig. 1 A profile of the grassland

2.2.1 Correlation measures

To calculate between-node connection values, we need to find a statistic measure. A useful measure is correlation coefficient (Zhang, 2007, 2011b). The statistically significant correlations may represent the true (direct or indirect) association or interaction between two nodes.

In present study, a positive correlation (association, interaction) means the two taxa (species, family, etc.) tend to jointly occur in a sample, and a negative one means the two taxa tend to exclusively occur in a sample.

The following three correlation measures were used in present study.

(1) Linear correlation

Linear (Pearson) correlation is the most used correlation measure

$$r=\Sigma(x_i-X)(y_i-Y)/[\Sigma(x_i-X)^2\,\Sigma(y_i-Y)^2]^{1/2}$$

where X and Y are the mean of coverage for taxon x and taxon y respectively. The statistic significance of linear correlation can be tested using t-test.

(2) Partial linaer correlation

Partial (pure) linear correlation is based on linear correlation, which reflects between-taxon direct interaction (Zhang, 2011b). The statistic significance of partial linear correlation can be tested using t-test. Using partial correlation will yield a statistic network with all connections as direct association (interaction).

The network derived from (partial) linear correlation is a linear network (Zhang, 2011b).

(3) Spearman rank correlation

Spearman rank correlation (Spearman, 1904; Schoenly and Zhang, 1999) is a correlation measure to denote a weak linear (quasi-linear) relationship. Using Spearman rank correlation will create a quasi-linear network (Zhang, 2011b).

In a linear network, the states of two connected taxa will show a linear dependent relationship whereas for a quasi-linear network, the states of two linked taxa will show a quasi-linear dependent relationship (Zhang, 2011b). Taxa that never follow linear and quasi-linear relationships are excluded from the two networks respectively. In a sense, taxa in the network are relatively predictable but isolated taxa are hard to be predicted (Zhang, 2011b).

2.2.2 Determination of network type and network structure

The calculation of network type and network structure based on power law distribution, binomial distribution, Poisson distribution, exponential distribution, skewness, coefficient of variation, entropy, and aggregation can be found in Zhang (2011b, 2012b), Zhang and Zhan (2011). Here I define and use a new index for network compactness (N_c) based on the definition of Zhang (2011b), i.e., N_c=the average of absolute values of correlation coefficients of all connections × number of practical connections/maximum number of connections. If there are s taxa in the network, then the maximum number of connections is $s(s$-1)/2. For a deterministic network with full connections, its network compactness is 1. Therefore the range of N_c is (0,1].

In this study, connectance=the number of practical connections/s^2, where s is the number of taxa.

2.2.3 Determination of sample homogeneity and sampling completeness

Samples were tested for homogeneity and sampling completeness. See Zhang (2011a, 2012b) for detailed description.

2.2.4 Drawing network

Networks were drawn using the software, netGenerator (Zhang (2012a). In the data file of netGenerator, I used data value, -1, to represent negative correlation, and 1 to represent positive correlation. However, other software is also available for use (Arnold et al., 2012).

3 Results

3.1 Composition of grassland community

The community composition, i.e., species (family) composition and coverage are indicated in Table 1 and 2. At the level of family, Gramineae is most dominant family in the community with the coverage of 45.62%, second by Leguminosae (13.38%) and Asteraceae (11.49%). In addition, Malvaceae (2.41%) and Cyperaceae (1.96%) accounted for a certain portion in the community. Other families are sparsely distributed in the grassland.

For species composition, *Cynodon dactylon* is the most dominant species (15.8%), followed by *Panicum repens* (9.01%), *Desmodium triflorum* (6.93%), *Ischaemum aristatum* (5.05%), *Paspalum scrobiculatum* (5.02%), *Axonopus compressus* (4.02%), *Ageratum conyzoides* (3.93%), and *Rhynchelytrum repens* (3.44%), etc. Gramineae is a most popular family in the most grassland, and *Cynodon dactylon* is a strong invasive and pioneer species, so their dominant coverage in the community is expected.

Table 1 Mean coverage (%) of families for all samples

Euphorbiaceae	0.12	Asteraceae	11.49	Cyperaceae	1.96
Leguminosae	13.38	Onagraceae	0.05	Lycopodiaceae	0.04
Pteridaceae	0.19	Connaraceae	0.01	Araceae	0.45
Lygodiaceae	0.09	Verbenaceae	0.31	Amaranthaceae	0.32
Gramineae	45.62	Anacardiaceae	0.13	Convolvulaceae	0.54
Apocynaceae	0.14	Rubiaceae	0.07	Commelinaceae	0.63
Malvaceae	2.41	Umbelliferae	0.28	Oxalidaceae	0.11

Table 2 Mean coverage (%) of species for all samples

Euphorbia hirta	0.04	*Lophatherum gracile*	0.38	*Wedelia triloba*	0.74
Sapium sebiferum	0.08	*Panicum repens*	9.01	*Jussiaea linifolia*	0.05
Cajanus scarabaeoides	2.56	*Paspalum scrobiculatum*	5.02	*Pteroloma triquetrum*	0.01
Desmodium gangeticum	0.03	*Rhynchelytrum repens*	3.44	*Verbena officinalis*	0.31
Desmodium triflorum	6.93	*Vetiveria zizanioides*	0.11	*Rhus chinensis*	0.13
Lespedeza bicolor	0.09	*Catharanthus roseus*	0.14	*Paederia scandens*	0.07
Mimosa pudica	0.58	*Sida acuta*	0.51	*Centella asiatica*	0.28
Moghania philippinensis	0.17	*Sida rhombifolia*	1.39	*Cayerus exaltatus*	0.10
Atylosia spp	2.80	*Urena lobata*	0.51	*Cyperus compressus*	0.00
Sesbania cannabina	0.22	*Ageratum conyzoides*	3.93	*Cyperus globosus*	1.59
Pteris vittata	0.19	*Artemisia indica*	0.13	*Eleocharis plantagineiformis*	0.01
Lygodium japonicum	0.09	*Aster subulatus*	0.44	*Kyllinga brevifolia*	0.26
Axonopus compressus	4.02	*Bidens bipinnata*	1.82	*Lycopodium cernuum*	0.04
Gramineae spp	1.21	*Bidens pilosa*	1.58	*Syngonium podophyllum*	0.45
Cymbopogon citratus	0.04	*Emilia sonchifolia*	0.52	*Alternanthera philoxeroides*	0.02
Cynodon dactylon	15.80	*Erigeron acer*	0.34	*Alternanthera sessilis*	0.30
Echinochloa crusgalli	0.01	*Eupatorium chinense*	1.55	*Ipomoea triloba*	0.16
Eleusine indica	0.01	*Ixeris polycephala*	0.18	*Pharbitis purpurea*	0.38
Eragrostis minor	1.10	*Mikania micrantha*	0.19	*Chrysopogon aciculatus*	0.56
Eragrostis pilosa	0.43	*Pterocypsela indica*	0.01	*Zebrina pendula*	0.07
Ischaemum aristatum	5.05	*Vernonia cinerea*	0.06	*Oxalis corniculata*	0.11

Gramineae spp is an unidentified Gramineae species. *Atylosia spp* is an unidentified Atylosia species.

3.2 Sample homogeneity and sampling completeness

Samples were conducted 500 bootstrap randomizations, from smallest sample size to maximum sample size (Fig. 2). Samples were homogeneous and sampling was complete at both species and family levels. It is thus reasonable to construct the statistic network from these sampling data.

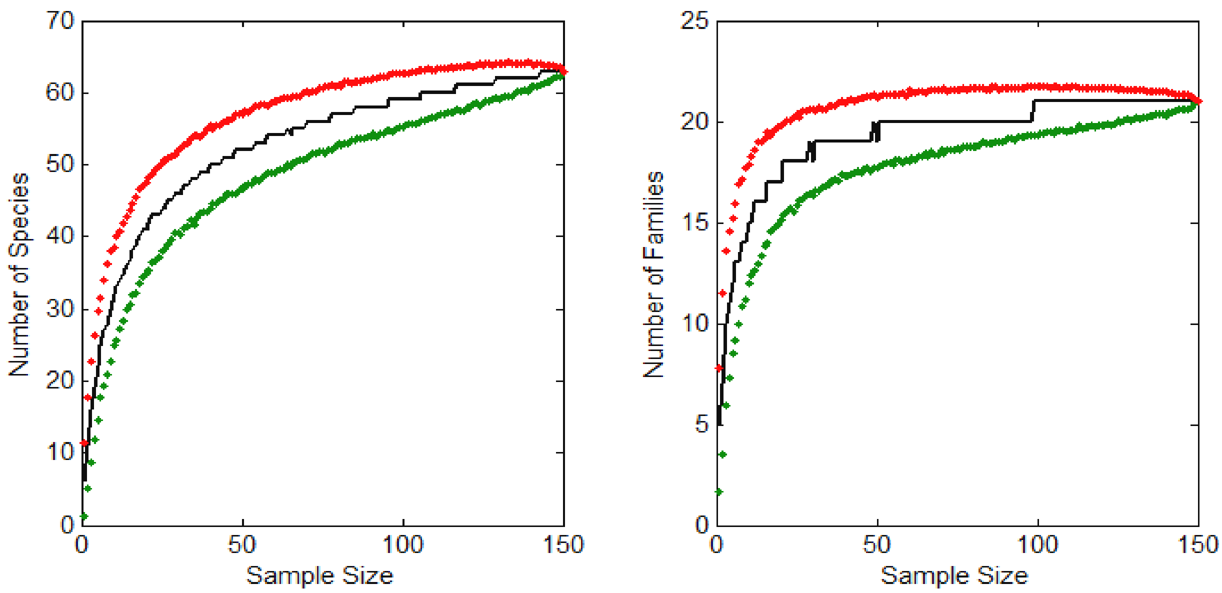

Fig. 2 Five hundreds of bootstrap randomizations of samples for different sample sizes. 95% upper and lower bound of confidence interval of observed number of taxa are indicated.

3.3 Construction of statistic network

Using the three correlation measures above, we obtained connection values and all significant connections. Some results from partial linear correlation are indicated in Tables 3-4 and Fig. 3.

Table 3 Partial linear correlations between families

Family	Family	Association Type	Partial Linear Corr.
Pteridaceae	Gramineae	-	-0.228
Pteridaceae	Oxalidaceae	-	-0.318
Gramineae	Apocynaceae	+	0.230
Gramineae	Rubiaceae	+	0.316
Malvaceae	Araceae	+	0.349
Malvaceae	Commelinaceae	+	0.296
Malvaceae	Oxalidaceae	+	0.359
Verbenaceae	Lycopodiaceae	+	0.339
Rubiaceae	Oxalidaceae	+	0.238
Euphorbiaceae	Euphorbiaceae	*	
Leguminosae	Leguminosae	*	
Lygodiaceae	Lygodiaceae	*	
Asteraceae	Asteraceae	*	
Onagraceae	Onagraceae	*	
Connaraceae	Connaraceae	*	
Anacardiaceae	Anacardiaceae	*	
Umbelliferae	Umbelliferae	*	
Cyperaceae	Cyperaceae	*	
Amaranthaceae	Amaranthaceae	*	
Convolvulaceae	Convolvulaceae	*	

+: positive association; -: negative association;
*: isolated species

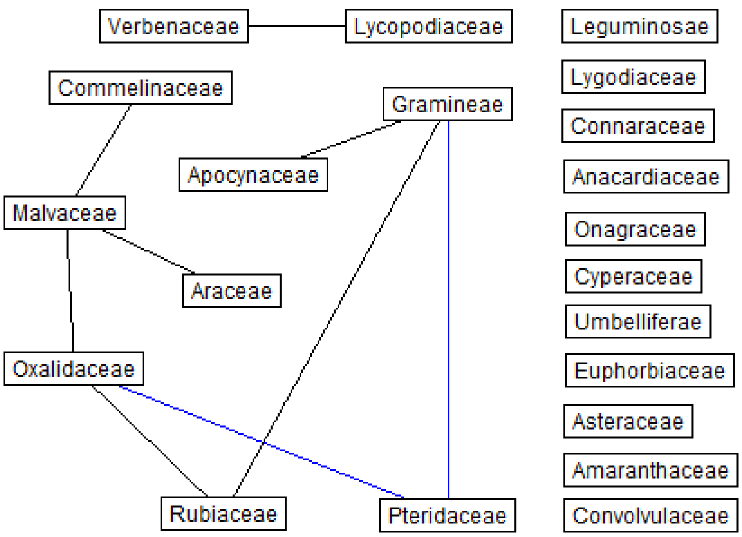

Fig. 3 Association network for families (constructed from partial linear correlation measure; see Table 3)

From results in Table 3, we may find that more than half of the families are isolated families. In the association network of families, most connections are positive interactions and they are all weak links (average of absolute partial linear correlations=0.297).

Table 4 Pure (partial) linear correlations between species

Species	Species	Association Type	Partial Linear Corr.	Species	Species	Association Type	Partial Linear Corr.
Euphorbia hirta	*Lespedeza bicolor*	+	0.492	*Ischaemum aristatum*	*Mikania micrantha*	-	-0.293
Euphorbia hirta	*Eragrostis pilosa*	+	0.534	*Lophatherum gracile*	*Bidens bipinnata*	+	0.617
Euphorbia hirta	*Jussiaea linifolia*	+	0.798	*Lophatherum gracile*	*Kyllinga brevifolia*	+	0.448
Euphorbia hirta	*Syngonium podophyllum*	+	0.506	*Paspalum scrobiculatum*	*Vetiveria zizanioides*	+	0.304
Sapium sebiferum	*Bidens pilosa*	+	0.469	*Paspalum scrobiculatum*	*Artemisia indica*	+	0.447
Cajanus scarabaeoides	*Rhynchelytrum repens*	+	0.299	*Rhynchelytrum repens*	*Chrysopogon aciculatus*	+	0.280
Cajanus scarabaeoides	*Urena lobata*	-	-0.303	*Vetiveria zizanioides*	*Chrysopogon aciculatus*	+	0.456
Cajanus scarabaeoides	*Bidens bipinnata*	+	0.431	*Catharanthus roseus*	*Emilia sonchifolia*	+	0.325
Cajanus scarabaeoides	*Jussiaea linifolia*	+	0.308	*Catharanthus roseus*	*Vernonia cinerea*	+	0.578
Desmodium gangeticum	*Lycopodium cernuum*	+	0.999	*Catharanthus roseus*	*Zebrina pendula*	-	-0.303
Desmodium triflorum	*Panicum repens*	-	-0.275	*Sida acuta*	*Sida rhombifolia*	+	0.329
Desmodium triflorum	*Cyperus globosus*	+	0.360	*Sida acuta*	*Syngonium podophyllum*	+	0.284
Lespedeza bicolor	*Jussiaea linifolia*	-	-0.403	*Sida rhombifolia*	*Wedelia triloba*	+	0.343
Lespedeza bicolor	*Centella asiatica*	+	0.513	*Sida rhombifolia*	*Cyperus globosus*	+	0.379
Lespedeza bicolor	*Chrysopogon aciculatus*	+	0.275	*Urena lobata*	*Bidens bipinnata*	+	0.532
Mimosa pudica	*Cyperus globosus*	-	-0.278	*Ageratum conyzoides*	*Aster subulatus*	-	-0.273
Moghania philippinensis	*Alternanthera sessilis*	+	0.347	*Ageratum conyzoides*	*Erigeron acer*	+	0.478
Atylosia spp	*Sesbania cannabina*	+	0.273	*Aster subulatus*	*Erigeron acer*	+	0.348
Atylosia spp	*Kyllinga brevifolia*	+	0.442	*Aster subulatus*	*Centella asiatica*	+	0.341
Atylosia spp	*Alternanthera sessilis*	+	0.411	*Emilia sonchifolia*	*Mikania micrantha*	+	0.339
Pteris vittata	*Pteris vittata*	*		*Emilia sonchifolia*	*Zebrina pendula*	+	0.356
Lygodium japonicum	*Ischaemum aristatum*	+	0.520	*Eupatorium chinense*	*Eupatorium chinense*	*	
Axonopus compressus	*Axonopus compressus*	*		*Ixeris polycephala*	*Ixeris polycephala*	*	
Gramineae spp	*Cynodon dactylon*	-	-0.274	*Mikania micrantha*	*Zebrina pendula*	-	-0.403
Cymbopogon citratus	*Cymbopogon citratus*	*		*Mikania micrantha*	*Oxalis corniculata*	+	0.290
Cynodon dactylon	*Eragrostis minor*	-	-0.379	*Pterocypsela indica*	*Vernonia cinerea*	+	0.330
Cynodon dactylon	*Panicum repens*	-	-0.378	*Wedelia triloba*	*Paederia scandens*	+	0.343
Cynodon dactylon	*Paspalum scrobiculatum*	-	-0.398	*Jussiaea linifolia*	*Syngonium podophyllum*	-	-0.401
Echinochloa crusgalli	*Ischaemum aristatum*	+	0.282	*Pteroloma triquetrum*	*Pteroloma triquetrum*	*	
Echinochloa crusgalli	*Mikania micrantha*	+	0.773	*Verbena officinalis*	*Verbena officinalis*	*	
Echinochloa crusgalli	*Zebrina pendula*	+	0.506	*Rhus chinensis*	*Rhus chinensis*	*	
Echinochloa crusgalli	*Oxalis corniculata*	-	-0.381	*Paederia scandens*	*Alternanthera sessilis*	+	0.379
Eleusine indica	*Eleusine indica*	*		*Centella asiatica*	*Kyllinga brevifolia*	+	0.365
Eragrostis minor	*Panicum repens*	-	-0.280	*Cayerus exaltatus*	*Cayerus exaltatus*	*	
Eragrostis minor	*Paspalum scrobiculatum*	-	-0.283	*Cyperus compressus*	*Cyperus compressus*	*	
Eragrostis pilosa	*Jussiaea linifolia*	-	-0.49	*Eleocharis plantagineiformis*	*Eleocharis plantagineiformis*	*	
Eragrostis pilosa	*Syngonium podophyllum*	-	-0.339	*Alternanthera philoxeroides*	*Alternanthera philoxeroides*	*	
Eragrostis pilosa	*Pharbitis purpurea*	+	0.300	*Ipomoea triloba*	*Pharbitis purpurea*	+	0.294
Ischaemum aristatum	*Emilia sonchifolia*	+	0.341	*Zebrina pendula*	*Oxalis corniculata*	+	0.611

Gramineae spp is an unidentified Gramineae species. *Atylosia spp* is an unidentified Atylosia species. +: positive association; -: negative association; *: isolated species.

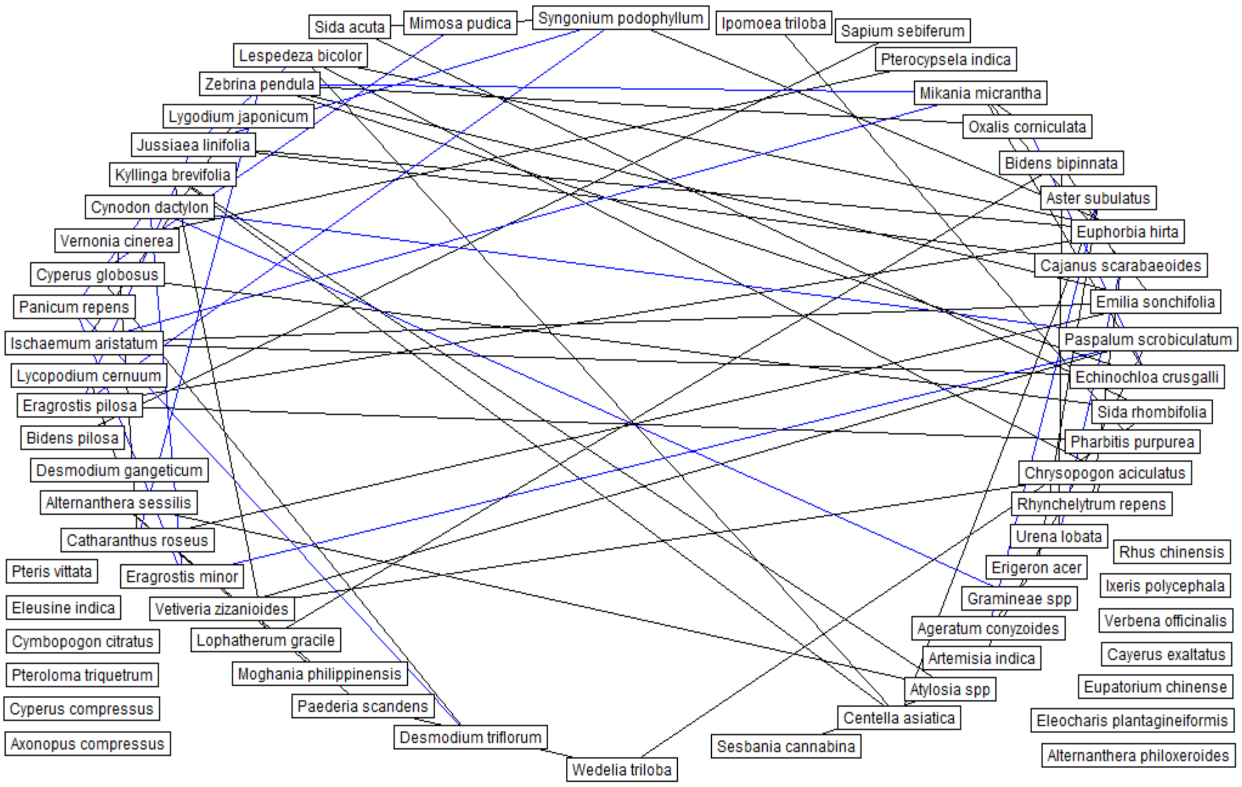

Fig. 4 Association network for species, constructed from partial linear correlation measure (see Table 4). Blue lines denote negative association and black lines denote positive association.

Gramineae, Malvaceae and Oxalidaceae have the most connections to other families. It should be noted that the abundant families, Leguminosae and Asteraceae are isolated families.

As indicated in Table 4 and Fig. 4, in the association network of species, the most connections are positive interactions and they are mostly weak links (average of absolute partial linear correlations=0.402). In addition, isolated species (13) are few in the network.

The state of a taxon in the network is the coverage of the taxon, as indicated in Tables 1 and 2.

The above results demonstrated that between-taxa connections at higher taxonomic level are weaker than that at lower taxonomic level.

3.3.1 Comparison of various networks

Using different correlation measures, we constructed three types of association networks (Figures 5, 6, and 7). The comparison of three association networks is shown in Table 5. It is found that the association network constructed from Spearman rank correlation is most connected (Fig. 5). No isolated family or species in the corresponding networks. From partial linear correlation, linear correlation, to Spearman rank correlation, the practical number of connections and connectance increases. Network compactness generally follows this trend also.

It is interesting that as the increase of network compactness and connectance, the portion and number of negative association declines dramatically.

Table 5 Comparison of three association networks

		Partial Linear Corr. Based Network	Linear Corr. Based Network	Spearman Rank Corr. Based Network
Total nodes	Family	21	21	21
	Species	63	63	63
Isolated nodes	Family	11	11	0
	Species	13	8	0
No.maximum connections	Family	210	210	210
	Species	1953	1953	1953
No. practical connections	Family	9	9	204
	Species	65	76	1946
Connectance	Family	0.0204	0.0204	0.9714
	Species	0.0333	0.0389	0.9964
Network Compactness	Family	0.0127	0.0129	0.0323
	Species	0.0134	0.0142	0.0119
No. positive interactions	Family	7	7	204
	Species	47	74	1946
No.negative interactions	Family	2	2	0
	Species	18	2	0

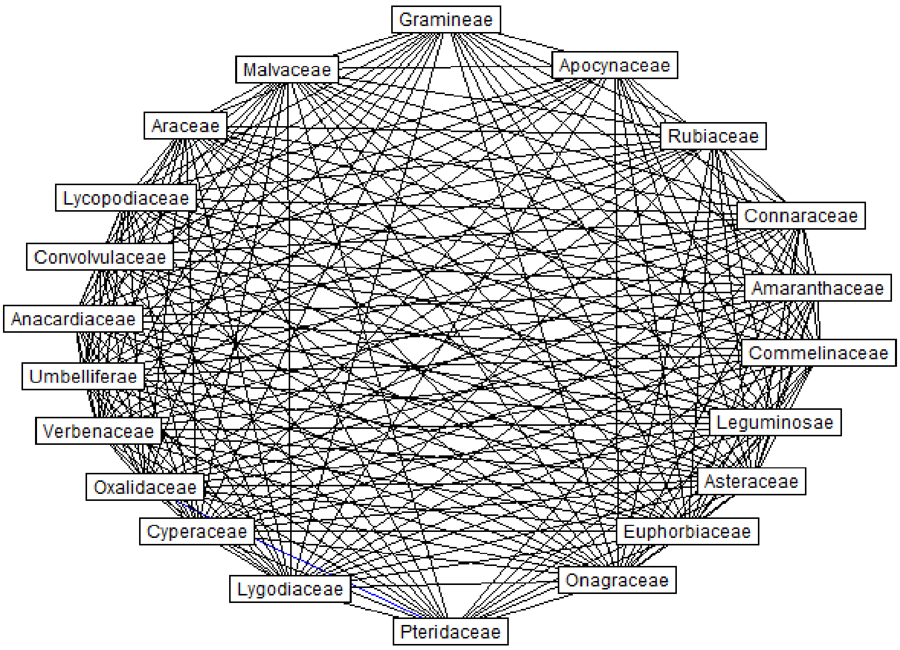

Fig. 5 Association network of families, constructed from Spearman rank correlation measure.

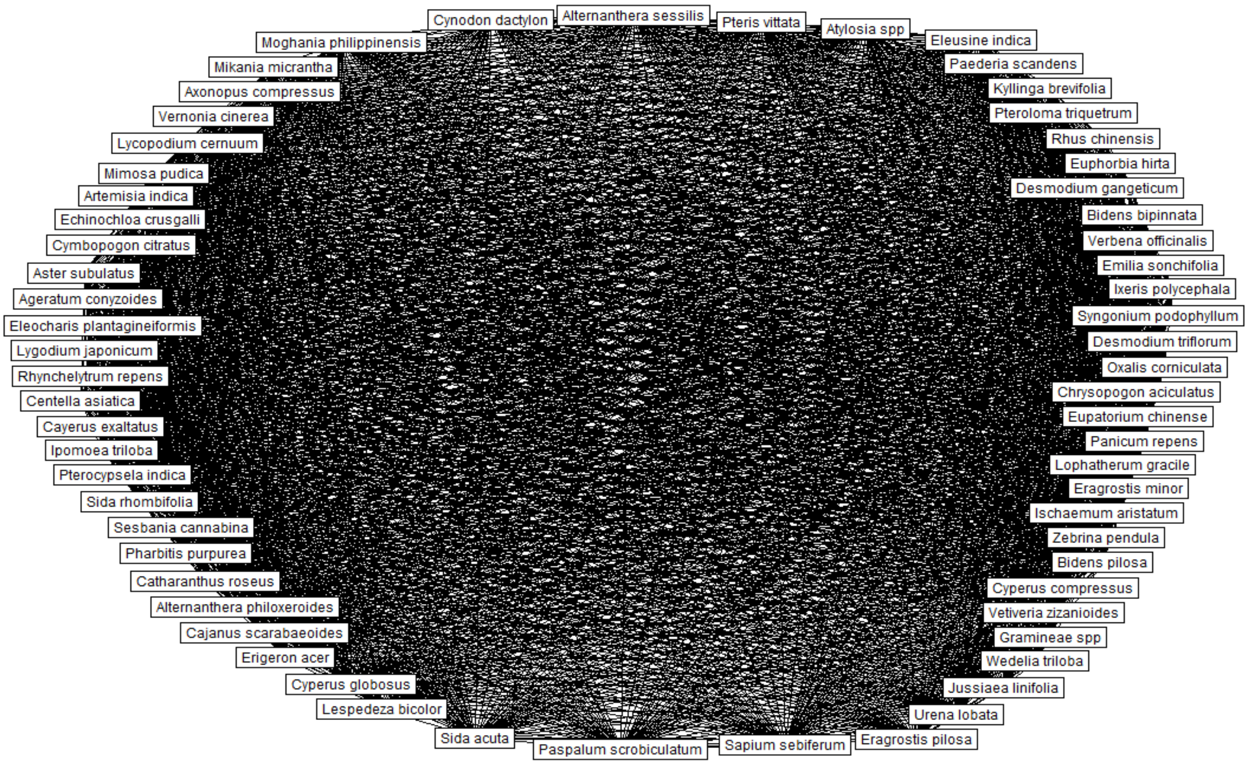

Fig. 6 Association network for species, constructed from Spearman rank correlation measure. Blue lines denote negative association and black lines denote positive association.

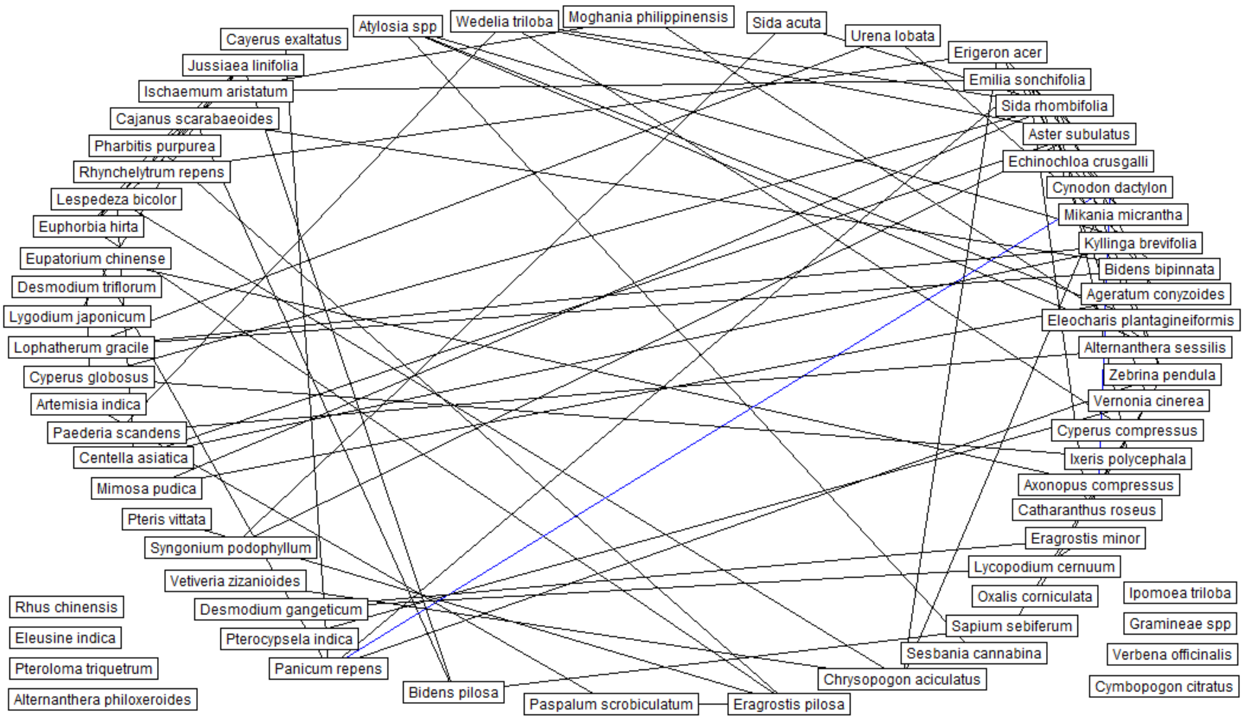

Fig. 7 Association network for species, constructed from linear correlation measure. Blue lines denote negative association and black lines denote positive association.

3.3.2 Network structure and network types

As indicated in Table 6, the networks constructed from partial linear correlation and linear correlation measures are generally (scale-free) complex networks. The degree of these networks is power low distributed.

Table 6 Network structure and network types

Interaction index	Partial Linear Corr.	Linear Corr.	Spearman Rank Corr.
Taxa	Species	Species	Family
Skewness of degree distribution	0.678	0.621	0.986
Aggregation index of network	1.059	1.059	0.954
	Complex network	Complex network	Random network
Variation coefficient H of network	1.124	1.145	0.064
Entropy E of network	0.255	0.349	-18.171
	Complex network	Complex network	Random network
Binomial distri.	NO	NO	NO
Poisson distri.	NO	NO	NO
Exponential distri.	NO	NO	NO
Power law distri.	YES	YES	NO
	K-S D value=0	K-S D value=0	K-S D value=0.445
	Scale-free complex network	Scale-free complex network	-

4 Conclusions and Discussion

(1) The statistic network, particularly the plant association network, constructed from field sampling, is mostly an information network. Interspecific relationships are always simple in plant communities. Most of the interspecific relationships in plant community are competition and cooperation. Compared to animal networks, the methodology of statistic network is more suitable to construct plant association networks.

(2) In the plant association network, most connections are weak and positive interactions. The association network constructed from Spearman rank correlation has most connections and isolated taxa are few. From partial linear correlation, linear correlation, to Spearman rank correlation, the practical number of connections and connectance in the constructed network increases. Network compactness also follows the trend. As the increase of network compactness and connectance, the portion and number of negative association declines dramatically.

(3) In an association (interaction) network, only a few connections follow the linear relationship. Most connections follow the quasi-linear or non-linear relationships.

(4) The association networks constructed from partial linear correlation and linear correlation measures are generally scale-free complex networks. The degree of these networks is power low distributed.

(5) Isolated species (families, etc.) are likely important in the statistic network. They are the sink species for shaping new network after a community is seriously disturbed. Isolated species (families) are always ignored in the studies of biological networks.

(6) Between-taxa connections at higher taxonomic level are generally weaker than that at lower taxonomic level.

Interspecific association (connection) may be measured with a variety of statistical measures, such as Chi-test, variance ratio test, the coupling coefficient, point correlation, the similarity percentage of distribution, Spearman rank correlation, Pearson linear correlation, etc (Schoenly and Zhang, 1999; Zhang, 2011b). In this study three correlation measures were used, of which Pearson linear correlation requires that the distribution of species follows a normal distribution (Leegendre, 1998). However this is not true if sample size is small. Therefore the sample size should be large enough in order to construct statistic network based on linear correlation measure. Spearman rank correlation is a nonparametric measure. It does not require stricter statistical assumption, and thus has not fewer limitations in the application. Other measures are suggested being tested and used in the future studies.

Acknowledgements
I am grateful to undergraduate students of Department of Ecology, Sun Yat-sen University, for their help in the field survey, conducted in 2007 and 2008.

References

Arnold J, Taha TR, Deligiannidis L. 2012. GKIN: a tool for drawing genetic networks. Network Biology, 2(1): 26-37

Butts CT. 2006. Exact bounds for degree centralization. Social Networks, 28: 283-296

Butts CT. 2009. Revisiting the foundations of network analysis. Science, 325: 414-416

Faust K. 2007. Very local structure in social networks. Sociological Methodology, 37: 209-256

Greig-Smith P. 1983. Quantitative Plant Ecology (3rd Ed). Blackwell Science Publication, Oxford, UK

Legendre P. 1998. Numerical Ecology. Elsevier, Amsterdam, Netherlands

Onnela JP, et al. 2007. Structure and tie strengths in mobile communication networks. Proceedings of the National Academy of Sciences of USA, 104: 7332

Paine RT. 1980. Food webs: linkage, interaction strength and community infrastructure. Journal of Animal Ecology, 49: 667-686

Paine RT. 1988. Food webs: road maps of interactions or grist for theoretical deveplopment? Ecology, 69: 1648–1654

Paine RT. 1992. Food-web analysis through field measurement of per capita interaction strength. Nature, 355: 73-75

Schoenly KG, Zhang WJ.1999. IRRI Biodiversity Software Series. V. RARE, SPPDISS, and SPPANK: programs for detecting between-sample difference in community structure[M]. IRRI Technical Bulletin No. 5. Manila (Philippines): International Rice Research Institute

Spearman C. 1904. The proof and measurement of association between two things. American Journal of Psychology, 15: 72-101

Whitehead H, Dufault S. 1999. Techniques for analyzing vertebrate social structure. Advances in the Study of Behavior, 28:33-74

Zhang WJ. 2007. Computer inference of network of ecological interactions from sampling data. Environmental Monitoring and Assessment, 124: 253-261

Zhang WJ. 2011a. A Java program to test homogeneity of samples and examine sampling completeness. Network Biology, 1(2): 127-129

Zhang WJ. 2011b. Constructing ecological interaction networks by correlation analysis: hints from community sampling. Network Biology, 1(2): 81-98

Zhang WJ. 2012a. A Java software for drawing graphs. Network Biology, 2(1): 38-44

Zhang WJ. 2012b. Computational Ecology: Graphs, Networks and Agent-based Modeling. World Scientific, Singapore

Zhang WJ, Zhan CY. 2011. An algorithm for calculation of degree distribution and detection of network type: with application in food webs. Network Biology, 1(3-4): 159-170