Network nestedness as generalized core-periphery structures

Sang Hoon Lee (상훈)\textsuperscript{1}

\textsuperscript{1}School of Physics, Korea Institute for Advanced Study, Seoul 02455, Korea

The concept of nestedness, in particular for ecological and economical networks, has been introduced as a structural characteristic of real interacting systems. We suggest that the nestedness is in fact another way to express a mesoscale network property called the core-periphery structure. With real ecological mutualistic networks and synthetic model networks, we reveal the strong correlation between the nestedness and core-periphery-ness (likeness to the core-periphery structure), by defining the network-level measures for nestedness and core-periphery-ness in the case of weighted and bipartite networks. However, at the same time, via more sophisticated null-model analysis, we also discover that the degree (the number of connected neighbors of a node) distribution poses quite severe restrictions on the possible nestedness and core-periphery-ness parameter space. Therefore, there must exist structurally interwoven properties in more fundamental levels of network formation, behind this seemingly obvious relation between nestedness and core-periphery structures.

PACS numbers: 87.23.-n, 89.75.Fb, 89.75.Hc, 92.40.Oj

I. INTRODUCTION

Since the pioneering work by May \cite{May1972}, the concept of nestedness indicating the systematically included structure composed of generalists and specialists has been assumed to be one of the most characteristic structures of ecological networks \cite{Martinez2001,Dunne2006}. These obviously evolved, not designed, systems must have reasons to be formed as such, and the candidates for the reasons include dynamical stability \cite{May1972,Dunne2006,Dunne2008} and biodiversity \cite{Balev2007}. Once such a structural property is expressed as a purely mathematical form, it is possible to study the network systems in general without involving the intrinsic properties of ecosystems. Indeed, compared to when the concept was first conceived, such networked systems in general have been widely investigated since the turn of the century and so on \cite{Balev2007}, which naturally enables us to connect the nestedness to possibly more network properties in more general contexts. For instance, the concept has been used to describe economic systems \cite{Balev2007,Balev2011} such as industrial ecosystems \cite{Balev2008} as well.

In such a general setting, nestedness is one of the examples of the mesoscale structure of networks. The term mesoscale means somewhere between the microscale structure such as the degree (the number of neighbors a node has) and the macroscopic structure such as the average edge density (the ratio of the number of existing edges to the number of node pairs). In this paper we suggest that another mesoscale structure, called the core-periphery structure of networks \cite{Barabasi2013,Mehran2015}, is closely related to nestedness. To be more specific, the gradual change from generalists to specialists corresponds to the gradual change in the coreness of the nodes, which makes the nested structure a generalized version of a clear-cut core versus periphery structure. In fact, the connection between the two concepts may seem to be too obvious from the adjacency matrix form (\( V \) shape for the nested and \( U \) shape for the core-periphery structure) to report, but we also find that it reveals a more fundamental property of real networks constrained by the degree distribution, which is widely used as the keystone of network ensembles.

To systematically investigate the relation between the two concepts, we first define the measures of nestedness and core-periphery-ness (likeness to the core-periphery structure) in the weighted and bipartite network level in Sec. \textsuperscript{II}. The mutualistic ecological networks and synthetic network model used in our study are presented in Sec. \textsuperscript{III}. Using the measures and data introduced, we present the result in Sec. \textsuperscript{IV} with a summary and discussion.

II. MEASURES FOR NESTEDNESS AND CORE-PERIPHERY STRUCTURES

A. Nestedness

We use the basic nestedness metric based on overlap and decreasing fill (NODF) \cite{Balev2007}, denoted by \( \nu \) in this paper for simplicity, although we note that there are other measures \cite{Balev2008,Balev2011}. The NODF counts the number of pairs of rows satisfying the nested structure for each column pair and the number of columns satisfying the nested structure for each row pair, after sorting the rows and columns in descending order of degree, or strength (the sum of weights on the edges attached to the node) in the case of weighted networks \cite{Balev2008}. Suppose \( W_{ij} \) is the weighted adjacency matrix \cite{Balev2008} representing the network with \( (W_{ij} > 0 \) represents the interaction between nodes \( i \) and \( j \), while \( W_{ij} = 0 \) represents the absence of interaction between \( i \) and \( j \), where both sets of indices are sorted by the descending order of nodes’ strength. Mathematically,

\[
\nu = \frac{\sum_{r=1}^{m(m-1)/2} f_{\text{col}}(r)}{n(n-1)/2} + \frac{\sum_{c=1}^{n(m-1)/2} f_{\text{row}}(c)}{m(m-1)/2},
\]

where \( f_{\text{col}}(r) \in [0,1] \) and \( f_{\text{row}}(c) \in [0,1] \) are the fraction of pairs of columns satisfying the nested inclusion structure for the row pair index \( r \) and the fraction of pairs of rows satisfying the nested inclusion structure for the column pair index \( c \).
for each type of node ω ∈ (animal, plant). References [16,17] also discuss the use of alternative transition functions to the one in Eq. [3].

We wish to determine the core-vector elements in [3] so that the core quality in Eq. [2] is maximized. This yields a CS value denoted by Ξ for node i of

$$\Xi_\omega(i) = Z_\omega \sum_{(\alpha, \beta)} C_i(\alpha, \beta) R(\alpha, \beta),$$  

(4)

where the normalization factor $Z_\omega$ is determined so that the maximum value of $\Xi$ over the entire set of nodes i is 1 for each type $\omega$ of node separately as $\omega \in \{\text{animal, plant}\}$. As in Refs. [16,17], we use simulated annealing [21] (with the same cooling schedule as in that paper). The core-quality landscape tends to be less sensitive to $\alpha$ than it is to $\beta$; for computational tractability, we fix the value of $\alpha_{\text{animal}} = \alpha_{\text{plant}} = 1/2$ and take the same value of $\Delta \beta_{\text{animal}} = \Delta \beta_{\text{plant}} = 0.01$ and thus consider $10^2$ evenly spaced points in the $\beta = (\beta_{\text{animal}}, \beta_{\text{plant}})$ plane. Finally, to define the coreness of an entire network, we define the normalized core quality inspired by Refs. [16,17], denoted by $\xi$ in this paper, as

$$\xi = \frac{\sum_{i,j} W_{ij} \Xi_{\text{animal}}(i) \Xi_{\text{plant}}(j)}{\sum_{i,j} W_{ij}},$$  

(5)

for the animal nodes $i \in \{1, \ldots, N_{\text{animal}}\}$ and the plant nodes $j \in \{1, \ldots, N_{\text{plant}}\}$, which we use for the coreness measure throughout this paper.

Sec. [IV] the nested structures of adjacency matrices are in fact such a gradual change of coreness. Of course, one can always define certain objective functions analogous to the community identification [20] to actually find the core-periphery separation, e.g., as presented in Ref. [18].

The method to calculate the edge-density-based coreness, called a core score (CS), is the modified version of the one introduced in Refs. [16–18] to fully consider the bipartivity of ecological mutualistic networks. There are other ways to quantify the coreness such as backup-path-based one in Refs. [17,18], but we use CS in this analysis because there is a natural way to quantify the overall core-periphery structure in the formalism of CS. Again, suppose $W_{ij}$ is the weighted adjacency matrix representing the mutualistic network, in particular, among a given set of animals $\{1, \ldots, N_{\text{animal}}\}$ and plants $\{1, \ldots, N_{\text{plant}}\}$. The network has $N = N_{\text{animal}} + N_{\text{plant}}$ nodes in total, and the value $W_{ij}$ indicates the weight of the connection between the animal node $i \in \{1, \ldots, N_{\text{animal}}\}$ and the plant node $j \in \{1, \ldots, N_{\text{plant}}\}$. We insert the core-matrix elements $C_{ij}$ into the core quality

$$R(\alpha, \beta) = \sum_{i,j} W_{ij} C_{ij}(\alpha, \beta),$$  

(2)

where the components of the parameter vector $\alpha = (\alpha_{\text{animal}}, \alpha_{\text{plant}}) \in [0, 1]$ determines the sharpness of the core-periphery division and $\beta = (\beta_{\text{animal}}, \beta_{\text{plant}}) \in [0, 1]$ determines the fraction of core nodes for animals and plants, respectively. We decompose the core-matrix elements into a product form, $C_{ij}(\alpha, \beta) = C_i(\alpha_{\text{animal}}, \beta_{\text{animal}}) C_j(\alpha_{\text{plant}}, \beta_{\text{plant}})$, where the elements of the core vector

$$C_i(\alpha_{\text{animal}}, \beta_{\text{animal}}) = \frac{i(1 - \alpha_{\text{animal}})}{2[\beta_{\text{animal}} N_{\text{animal}}]}$$

$$+ \frac{1 + \alpha_{\text{animal}}}{2}$$

$$C_j(\alpha_{\text{plant}}, \beta_{\text{plant}}) = \frac{\beta_{\text{plant}} N_{\text{plant}}}{2(\beta_{\text{plant}} N_{\text{plant}} - N_{\text{animal}})}$$

$$\times \left(1 - \beta_{\text{plant}} N_{\text{plant}}\right)$$

for $i \leq \lfloor \alpha_{\text{animal}} N_{\text{animal}} \rfloor$, for $i > \lfloor \alpha_{\text{animal}} N_{\text{animal}} \rfloor$,

(3)
To get rid of any effect from the degree distribution or sequence, for a given network structure, we present an additional randomization scheme called the edge-pair-shuffling (EPS) process, illustrated in Fig. 3. Since there does not exist a possible pair for swapping in the perfectly nested structure illustrated in Fig. 2(a) (suppose that A and B are more "generalist" than C and D, without loss of generality, then there should always be the edges A–D and B–C in that case), we start from our synthetic network model with given \( \eta \) values and apply the EPS process for selected edge pairs (A–B and C–D, when both A–D and B–C do not exist, in Fig. 3) uniformly at random repeatedly. T Monte Carlo steps in the unit of the number of edges. Figure 4 shows some examples with various T values for the 100 animals and 100 plants. Note that the EPS process cannot destruct the nested structure, because there is a fundamental constraint of graphicality for a given degree sequence in a bipartite network [24]. In fact, it is quite the opposite. Somewhat counterintuitively, the average \( \nu \) value is slightly increased as we increase the number of Monte Carlo steps T, as presented in Sec. IV.

### III. DATA AND SYNTHETIC MODEL

#### A. Ecological Network Data

We use 89 mutualistic network data introduced in Ref. [8] that can be downloaded in Ref. [22], consisting of 59 pollinator networks and 30 seed dispersal networks. Some networks are weighted by the interaction strength, while the others are unweighted. The size of networks varies greatly, from the smallest one composed of 6 nodes (3 animals and 3 plants) to the largest one composed of 997 nodes (883 animals and 114 plants). Such size diversity provides us with a nice opportunity to cross-check the correlation between various measures for the system size varying across the two orders of magnitude. Figure 1 shows one example of a network with the core score [16,18] defined in Eq. (4).

#### B. Synthetic Network Model

To control the various effects of other properties of real networks that will be discussed in Sec. IV, we construct the series of synthetic unweighted networks with tunable nestedness. First, we construct the perfectly nested structure shown in Fig. 2(a) with given numbers of animal and plants. One can see the similarity between these nested structures represented in the adjacency matrix \( A \) and the core-periphery structure \( D \) as shown in Fig. 1.1(b) in Ref. [16]. In Ref. [18], the possibility of generalization of such a step structure is presented, and the finest step structure would be the perfectly nested structure in Fig. 2(a), indeed. In this respect, we regard the nested structure as a generalized core-periphery structure. Starting from this perfectly nested structure, we add noise with a certain probability \( \eta \), for each existing edge, the edge is removed, and a randomly chosen node pair that is currently not connected is connected with probability \( \eta \). Figure 2 shows some examples with various \( \eta \) values for 100 animals and 100 plants. However, even this model does not preserve the degree sequence (thus the effect of degree heterogeneity), which yields nontrivial correlations in regard to degree, as presented in Sec. IV.
Therefore, starting from . . .

edge-pair-shuffling (EPS)

A
B
C
D
A
B
C
D
animals
plants
animals
plants

FIG. 3. Illustration of the EPS method, where the pair of edges (the edges in red) A–B and C–D are switched to A–D and B–C.

V. CONCLUSIONS AND DISCUSSION

We have explored the seemingly obvious connection between the nestedness and core-periphery structure of networks, despite the fact that it looks obvious solely from the shape of adjacency matrix with proper ordering. We have shown the actual correlation between the two using a set of ecological mutualistic networks and model networks, by clearly addressing the nestedness and core-periphery-ness in the weighted and bipartite network level. In addition, in the process of generating null-model networks, we have found that given degree distributions set large restriction on both nestedness and core-periphery-ness, which hinders the investigation on the parameter space.

In any case, it is clear that the nestedness can be considered as a generalized (or finer) version of the core-periphery structure based on our observation. We may even consider other variants such as sorting the nodes with respect to the core-periphery-ness values instead of degree values to calculate the nestedness, e.g., modification of Eq. (1), where the nodes are sorted based on ξ in Eq. (4) instead of degree. Once we accept that those two concepts are closely related, albeit not equivalent, we can map many problems in regard to nestedness, such as its origin, and effects on the system of interest, to those of the core-periphery structures where we may be able to find answers more easily.

Finally, we would like to remark on the work on the nestedness and the community structure [20] measured by the modularity function [33], where the authors indeed found some degree of correlation between the two concepts. However,
FIG. 4. Examples of adjacency matrices for the EPS method illustrated in Fig. 3 with different shuffling step length $T$ in the unit of the number of edges, starting from our synthetic network model with $\eta = 0.2$: (a) $T = 0$ (an original synthetic network with $\eta = 1/5$), (b) $T = 1/2$, and (c) $T = 1$. The purple squares and empty sites represent 1 and 0, respectively.

FIG. 5. For the mutualistic network data, (a) $\nu$ in Eq. (1) versus $\xi$ in Eq. (5), with the correlation coefficients 0.634 with the $p$-value less than $10^{-10}$ (Pearson), 0.751 with the $p$-value less than $10^{-16}$ (Spearman), and 0.559 with the $p$-value less than $10^{-14}$ (Kendall), (b) $\nu$ in Eq. (1) versus the edge density, with the correlation coefficients 0.875 with the $p$-value less than $10^{-28}$ (Pearson), 0.878 with the $p$-value less than $10^{-28}$ (Spearman), and 0.721 with the $p$-value less than $10^{-22}$ (Kendall), and (c) $\xi$ in Eq. (5) versus the edge density, with the correlation coefficients 0.595 with the $p$-value less than $10^{-19}$ (Pearson), 0.798 with the $p$-value less than $10^{-19}$ (Spearman), and 0.614 with the $p$-value less than $10^{-18}$ (Kendall).

FIG. 6. For the synthetic network model with 100 animals and 100 plants, and the noise parameter $\eta$, as shown in Fig. 2 (a) $\nu$ in Eq. (1) versus $\xi$ in Eq. (5), (b) $\nu$ in Eq. (1) versus the standard deviation of degree, and (c) $\xi$ in Eq. (5) versus the standard deviation of degree. For all panels, the points from the lower (left) correspond to $\eta = 1, 0.95, 0.9, 0.85, \ldots, 0.1, 0.05$, and 0 [perfect nestedness: Fig. 2(a)]. The error bars, most of which are smaller than the symbol size, in both the horizontal and vertical directions present the standard deviation for each $\eta$ (100 network realizations for each $\eta$), except for $\eta = 0$ (undefined because there is only one possible realization).
the correlations reported there are much weaker than the ones we report in this work. It is worth looking at the mesoscale properties of networks, but we believe that the core-periphery structure is the correct measure to compare, rather than the community structure. The phrase “two sides of the same coin” included the title of Ref. [33] should be attached in fact to the nestedness versus core-periphery structure, not the community structure.

ACKNOWLEDGMENTS

The author thanks Puck Rombach for the MATLAB code to calculate \( \xi \) values and Young-Ho Eom (임영호), Daniel Kim (김영호), and Petter Holme for fruitful discussions.

[1] R. M. May, Will a large complex system be stable?, Nature 238, 413 (1972).
[2] A. Roberts, Stability of a feasible random ecosystem, Nature 251, 607 (1974).
[3] M. Almeida-Neto, P. Guimarães Jr., R. D. Loyola, and W. Ul- rich, A consistent metric for nestedness analysis in ecological systems: Reconciling concept and measurement, Oikos 117, 1227 (2008).
[4] G. Corso, A. I. L. Araujo, and A. M. Almeida, A new nestedness estimator in community networks, e-print arXiv:0803.0007.
[5] J. Galeano, J. M. Pastor, and J. M. Iriondo, Weighted-interaction nestedness estimator (WINE): A new estimator to calculate over frequency matrices, Environ. Modell. Softw. 24, 1342 (2009).
[6] U. Bastolla, M. A. Fortuna, A. Pascual-García, A. Ferrera, B. Luque, and J. Bascompte, The architecture of mutualistic networks minimizes competition and increases biodiversity, Nature 458, 1018 (2009).
[7] S. Allesina and S. Tang, Stability criteria for complex ecosystems, Nature 483, 205 (2012).
[8] R. P. Rohr, S. Saavedra, and J. Bascompte, On the structural stability of mutualistic systems, Science 345, 1253497 (2014).
[9] S. N. Dorogovtsev and J. F. F. Mendes, Evolution of networks, Adv. Phys. 51, 1079 (2002); R. Albert and A.-L. Barabási, Statistical mechanics of complex networks, Rev. Mod. Phys. 74, 47 (2002); M. E. J. Newman, The structure and function of complex networks, SIAM Rev. 45, 167 (2003); S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D.-U. Hwang, Complex networks: Structure and dynamics, Phys. Rep. 424, 175 (2006); M. E. J. Newman, Networks: An Introduction (Oxford University Press, Oxford, 2010).

[10] S. Saavedra, F. Reed-Tsochas, and B. Uzzi, A simple model of bipartite cooperation for ecological and organizational networks, Nature 457, 463 (2009).

[11] S. Saavedra, D. B. Stouffer, B. Uzzi, and J. Bascompte, Strong contributors to network persistence are the most vulnerable to extinction, Nature 478, 233 (2011).

[12] S. Bustos, C. Gomez, R. Hausmann, and C. A. Hidalgo, The dynamics of nestedness predicts the evolution of industrial ecosystems, PLOS ONE 7, e49393 (2012).

[13] S. P. Borgatti and M. G. Everett, Models of core/periphery structures, Soc. Networks 21, 375 (2000).

[14] P. Holme, Core-periphery organization of complex networks, Phys. Rev. E 72, 046111 (2005).

[15] F. Csermely, A. London, L.-Y. Wu, and B. Uzzi, Structure and dynamics of core-periphery networks, J. Complex Netw. 1, 93 (2013).

[16] M. P. Rombach, M. A. Porter, J. H. Fowler, and P. J. Mucha, Core-periphery structure in networks, SIAM J. Appl. Math. 74, 167 (2014).

[17] S. H. Lee, M. Cucuringu, and M. A. Porter, Density-based and transport-based core-periphery structures in networks, Phys. Rev. E 89, 032810 (2014).

[18] M. Cucuringu, P. Rombach, S. H. Lee, and M. A. Porter, Detection of core-periphery structure in networks using spectral methods and geodesic paths, e-print arXiv:1410.6572.

[19] A. Barrat, M. Barthélemy, R. Pastor-Satorras, and A. Vespignani, The architecture of complex weighted networks, Proc. Natl. Acad. Sci. USA 101, 3747 (2004).

[20] M. A. Porter, J.-P. Onnela, and P. J. Mucha, Communities in networks, Not. Am. Math. Soc. 56, 1082 (2009); S. Fortunato, Community detection in graphs, Phys. Rep. 486, 75 (2010).

[21] S. Kirkpatrick, C. D. Gelatt, Jr., and M. P. Vecchi, Optimization by simulated annealing, Science 220, 671 (1983).

[22] web of life: ecological network database. http://www.web-of-life.es/

[23] T. Kamada and S. Kawai, An algorithm for drawing general undirected graphs, Inform. Process. Lett. 31, 7 (1989).

[24] D. Gale, A theorem on flows in networks, Pacific J. Math. 7, 1073 (1957); H. J. Ryser, Combinatorial properties of matrices of zeros and ones, Canad. J. Math. 9, 371 (1957).

[25] P. Holme and J. Zhao, Exploring the assortativity-clustering space of a network’s degree sequence, Phys. Rev. E 75, 046111 (2007).

[26] C. Orsini, M. M. Dankulov, P. Colomer-de-Simón, A. Jamakovic, P. Mahadevan, A. Vahdat, K. E. Bassler, Z. Toroczkai, M. Boguñá, G. Caldarelli, S. Fortunato, and D. Krioukov, Quantifying randomness in real networks, Nat. Commun. 6, 8627 (2015).

[27] R. Fischer, J. C. Leitão, T. P. Peixoto, and E. G. Altmann, Sampling motif-constrained ensembles of networks, Phys. Rev. Lett. 115, 188701 (2015).

[28] C. I. Del Genio, T. Gross, and K. E. Bassler, All scale-free networks are sparse, Phys. Rev. Lett. 107, 178701 (2011).

[29] Y. Baek, D. Kim, M. Ha, and H. Jeong, Fundamental structural constraint of random scale-free networks, Phys. Rev. Lett. 109, 118701 (2012).

[30] J. Bascompte, P. Jordano, C. J. Melián, and J. M. Olesen, The nested assembly of plant-animal mutualistic networks, Proc. Natl. Acad. Sci. U.S.A. 100, 9383 (2003).

[31] E. Burgos, H. Ceva, R. P. J. Perazzo, M. Devoto, D. Medan, M. Zimmermann, and A. M. Delbue, Why nestedness in mutualistic networks?, J. Theor. Biol. 249, 307 (2007).

[32] O. Pastravanu and M. Voicu, Generalized matrix diagonal stability and linear dynamical systems, Linear Algebra Appl. 419, 299 (2006).

[33] M. A. Fortuna, D. B. Stouffer, J. M. Olesen, P. Jordano, D. Mouillot, B. R. Krasnov, R. Poulin, and J. Bascompte, Nestedness versus modularity in ecological networks: Two sides of the same coin?, J. Anim. Ecol. 79, 811 (2010).