Physics Letters A 378 (2014) 1091–1094

Contents lists available at ScienceDirect

Physics Letters A

www.elsevier.com/locate/pla

A new information dimension of complex networks

Daijun Wei [a,b], Bo Wei [a], Yong Hu [c], Haixin Zhang [a], Yong Deng [a,d,*]

[a] *School of Computer and Information Science, Southwest University, Chongqing 400715, China*
[b] *School of Science, Hubei University for Nationalities, Enshi 445000, China*
[c] *Institute of Business Intelligence and Knowledge Discovery, Guangdong University of Foreign Studies, Guangzhou 510006, China*
[d] *School of Engineering, Vanderbilt University, TN 37235, USA*

ARTICLE INFO

Article history:
Received 14 November 2013
Received in revised form 7 February 2014
Accepted 8 February 2014
Available online 17 February 2014
Communicated by C.R. Doering

Keywords:
Fractal
Self-similarity
Information dimension
Complex networks

ABSTRACT

The fractal and self-similarity properties are revealed in many complex networks. The classical information dimension is an important method to study fractal and self-similarity properties of planar networks. However, it is not practical for real complex networks. In this Letter, a new information dimension of complex networks is proposed. The nodes number in each box is considered by using the box-covering algorithm of complex networks. The proposed method is applied to calculate the fractal dimensions of some real networks. Our results show that the proposed method is efficient when dealing with the fractal dimension problem of complex networks.

1. Introduction

Recently, complex networks have attracted a growing interest in many disciplines [1–7]. Several properties of complex networks have been revealed, including small-world phenomena [8], scale-free degree [9] and community structure [10], etc. The fractal theory has been used in the study of various subjects [11–18]. The fractal and self-similarity properties of complex networks are discovered by Song et al. [19]. The fractal and self-similarity of complex networks are extensively studied by many researchers [20–26]. The classical box-covering algorithm is described in detail and applied to demonstrate the existence of self-similarity in some real complex networks [27,28]. From then on, the classical box-covering algorithm for complex networks is extensively studied [29–36] and modified for weighted complex networks [37].

The fractal and self-similarity properties of complex networks are revealed from a different perspective, such as volume dimension [38,39], correlation dimension [40] and information dimension [41]. However, the classical information dimension is only suitable for planar networks, since complex networks must be measured in plane area [41]. It is almost impossible since the distance between nodes of real complex networks cannot be obtained in the plane. We find that most of boxes cover different nodes number for given a box size in the classical box-covering algorithm. It means that boxes contain different information, even in these boxes have same size. In this Letter, by improving the classical information dimension, a new information dimension to characterize the fractal dimension of complex networks is proposed. In what follows, the classical information dimension of complex networks is introduced in Section 2. The proposed model of information dimension for complex networks is depicted in Section 3. In Section 4, the efficiency of the proposed method is illustrated by calculating fractal dimensions of some real complex networks. Some discussions and conclusions are presented in Section 5.

2. The classical information dimension of complex networks

In this section, the classical information dimension of complex networks is briefly introduced. Information dimension is introduced by Renyi based on the probability method [42]. By applying this method to planar networks, the planar network is covered by various-size squares. The information dimension of complex networks is given as follows [41]:

$$I = -\sum_{i=1}^{N_\varepsilon} p_i(\varepsilon) \ln p_i(\varepsilon), \tag{1}$$

where ε is square size, N_ε is number of boxes. $p_i(\varepsilon)$ is ration of number of squares and given as [41],

* Corresponding author at: School of Computer and Information Science, Southwest University, Chongqing 400715, China. Tel.: +86 023 68254555.
E-mail addresses: ydeng@swu.edu.cn, prof.deng@hotmail.com (Y. Deng).

http://dx.doi.org/10.1016/j.physleta.2014.02.010
0375-9601/

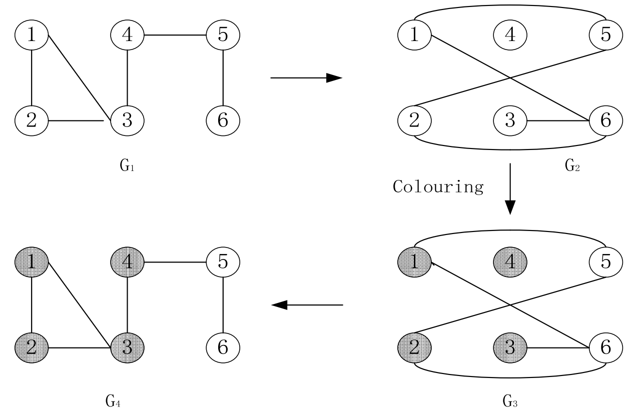


Fig. 1. The idea of the classical box-covering algorithm for complex networks, where $l = 3$. The network G_1 is original network with 6 nodes and 6 edges. The network G_2 is obtained by only connecting to nodes which distance between them not less than 3 in network G_1. The network G_3 is obtained when the greedy algorithm is used for node coloring on G_2.

$$p_i(\varepsilon) = \frac{q_i(\varepsilon)}{Q_i(\varepsilon)}, \tag{2}$$

where $q_i(\varepsilon)$ is number of squares including nodes of complex network, $Q_i(\varepsilon)$ is the number of squares covering the given plane area. The value of information dimension d_I of fractal set is obtained as follows [41]:

$$d_I = -\lim_{\varepsilon \to 0} \frac{I}{\ln \varepsilon} = \lim_{\varepsilon \to 0} \frac{\sum_{i=1}^{N_\varepsilon} p_i(\varepsilon) \ln p_i(\varepsilon)}{\ln \varepsilon}. \tag{3}$$

In planar networks, the distance between nodes is defined by the Euclidean distance of these nodes. The number of squares including nodes of planner network is easily obtained. The classical information dimension of complex network is feasible for planar networks. However, in the most of real complex networks, the distance between nodes is depended on the number of edges of the shortest path between two nodes. It is unreasonable that a real complex network is embedded in the plane area since the Euclidean distance between nodes cannot be obtained. For real complex network, it is almost impossible to obtain the value of $p_i(\varepsilon)$ in Eq. (2).

3. A new information dimension of complex networks

3.1. The classical box-covering algorithm of complex networks

The original definition of box-covering is initially proposed by Hausdorff [43,44]. It is applied in complex networks by Song, et al. [19,28]. For given box size l, randomly assign a unique id from 1 to n to all network nodes, every box is a set of nodes where all distances d_{ij} between any two nodes i and j in the box are smaller than l. The minimum number of boxes $N_b(l)$ must cover the entire network. The core idea of the classical box-covering algorithm of complex networks is shown in Fig. 1 and summarized as three steps [28].

Step 1: For a given network G_1 and a box size l, a new network G_2 is obtained, in which node i is connected to node j when the distance between them d_{ij} is not less than l.

Step 2: By the coloring problem of graph theory, nodes of the network G_2 between directly link each other are painted different colors. And then, a network G_3 is obtained in Fig. 1.

Step 3: Each color in network G_3 represents a different box. And then, the minimum number of box $N_b(l)$ can be counted.

Increase l by one until l is more than network diameter. For fractal complex networks, the relationship $N_b(l)$ and l can be given as follows:

$$N_b(l) \sim l^{-d_b} \tag{4}$$

where d_b is box dimension of complex networks. The value of d_b is obtained as follows [28]:

$$d_b = -\lim_{l \to 0} \frac{\ln N_b(l)}{\ln l}. \tag{5}$$

As described above, the distance between nodes only depend on number of edges, which connect from a node to another. The algorithm has been widely used to calculate fractal dimension of complex networks. In practice, the value of d_b is obtained by the slope of the straight line in the log–log plot, which by fitting the relationship $\ln N_b(l)$ and $\ln l$ [28].

3.2. Definition of proposed information dimension of complex networks

In our proposed information dimension of complex networks, the classical information dimension and the box-covering algorithm are referenced. Most of boxes have different number of nodes for given a box size in the classical box-covering algorithm. The difference number of nodes in boxes is considered in our method. For a give box size l, the probability of information containing the ith box is denoted by $p'_i(l)$ and defined as follows:

$$p'_i(l) = \frac{n_i(l)}{n}, \tag{6}$$

where $n_i(l)$ is the number of nodes in the ith box and n is all number of nodes of complex networks. Similar to Eqs. (1) and (3), information dimension is given as follows:

$$I'(l) = -\sum_{i=1}^{N_b} p'_i(l) \ln p'_i(l). \tag{7}$$

d'_I is obtained as follows:

$$d'_I = -\lim_{l \to 0} \frac{I'(l)}{\ln(l)} = \lim_{l \to 0} \frac{\sum_{i=1}^{N_b} p'_i(l) \ln p'_i(l)}{\ln(l)}, \tag{8}$$

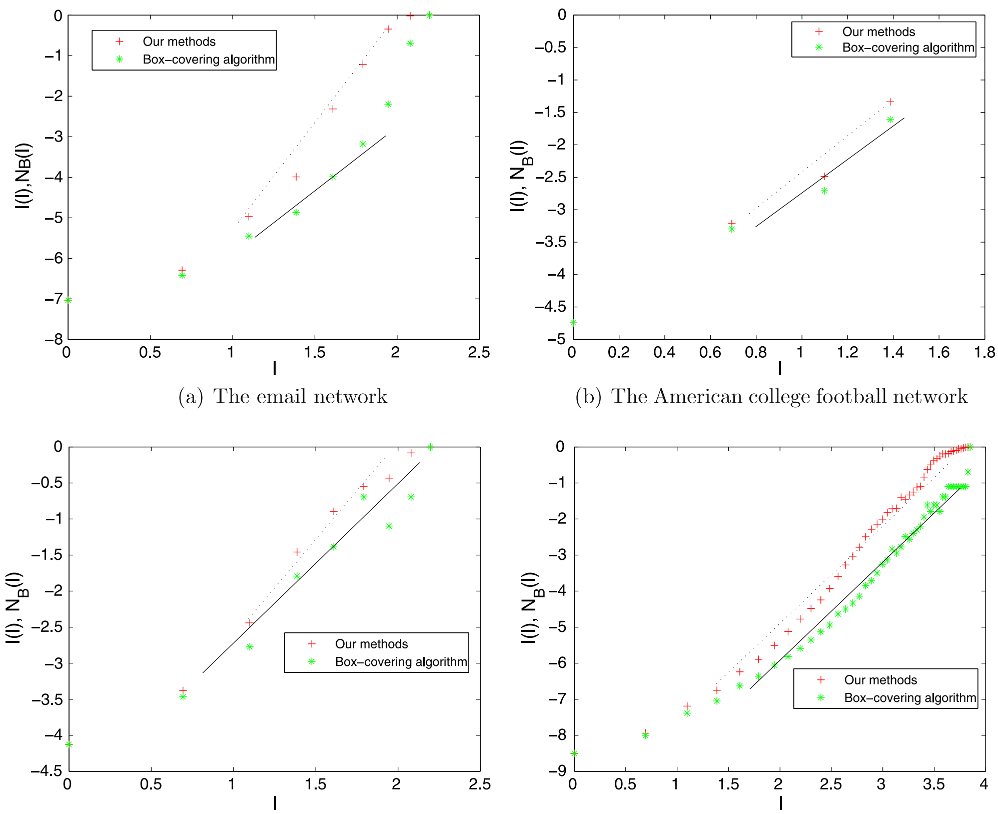
where d'_I is information dimension of complex networks. Using Eqs. (6) and (8), we have

$$d'_I = \lim_{l \to 0} \frac{\sum_{i=1}^{N_b} \frac{n_i(l)}{n} \ln \frac{n_i(l)}{n}}{\ln(l)}. \tag{9}$$

Eqs. (8) and (9) are theoretic formulations. In planar networks, the value of l can be very small. However, the value of l of real complex network cannot be very small since the distance between nodes is not less than one. In our method, the relationship $\ln(l)$ and $I'(l)$ is linear in a log–log plot. Limited number of box size l is considered. And then, the value of $d_I{}'$ is provided by the slope of the straight line in the log–log plot. In proposed definition of information dimension of complex networks, the probability of information is represented the ration of node number. The number of nodes is easily calculated by the classical box-covering algorithm.

4. Applications

In the section, the email network (http://vlado.fmf.unilj.si/pub/networks/data/), the American college football network, the dolphins social network and the power network (http://www-personal.umich.edu/mejn/netdata/) are calculated by using our method and the classical box-covering algorithm [28]. The number of nodes and edges of these complex networks are given in the first and second row of Table 1. The fractal dimension is averaged for 1000 times. Fractal scaling analysis of these real complex networks is shown in Fig. 2. In Fig. 2, asterisk indicates the correlation



(a) The email network (b) The American college football network

(c) The dolphins social network (d) The power network

Fig. 2. Fractal scaling analysis of these real complex networks. Symbols refer to: $*$ indicates the correlation between l and N_b and + indicates the correlation between l and $I'(l)$.

Table 1
General characteristics of sever real networks and the fractal dimension $d_I{}'$ by our proposed method and d_b by Song's method [28]. Q_I and Q_b represent MSE of linear in our method and Song's method [28], respectively.

Network	Size	Edges	d'_I	Q_I	d_b	Q_b
Email	1133	10 902	3.69	0.7685	3.166	0.9841
American college football	115	615	2.766	0.4187	2.688	0.6011
Dolphins social	62	159	2.061	0.2448	1.888	0.323
Power grid	4941	6594	2.694	0.3767	2.4109	0.4641

between $\ln(l)$ and $\ln(N_b)$ by the classical box-covering algorithm and plus sign indicates the correlation between $\ln(l)$ and $I'(l)$ by our method. By means of the least square fit, the slope of straight lines in the log–log plot is given. The solid line and dotted line are obtained by using the classical box-covering algorithm and our method, respectively. In order to check the fitting line, sum square error (SSE) of linear is applied in Ref. [45]. In our method, mean square error (MSE) Q is used and defined as follows:

$$Q = \sqrt{\frac{1}{m}\sum_{k=1}^{m}(y_k - \hat{y}_k)^2} \tag{10}$$

where y_k is value of discrete point, $\hat{y}_k$ is a function of line and m is all number of points. Q_i and Q_b represent MSE of linear by our method and Song's method [28], respectively. The values of d'_I, Q_i, d_b and Q_b for these real networks are shown in Table 1.

From Fig. 2 and Table 1, the fractal properties of these complex networks are revealed. The values of MSE in line 2, 3 and 4 of Table 1 are rather small. And from (b), (c) and (d) of Fig. 2, the American college football, the dolphins social network and power gird have fractal property using our method and the classical box-covering method. From (a) of Fig. 1, these points are more dispersed using the classical box-covering algorithm than the proposed method. And on line 1 of Table 1, the values of Q_b and Q_I are 0.9841 and 0.7685, which are both closer to 1. The email network cannot be revealed significant fractal property. From Table 1, the values of MSE of our method are small. It demonstrates that the information dimension is effective for fractal dimension of complex networks.

5. Discussions and conclusions

The information dimension of complex networks is form of Shannon entropy of complex networks. These information dimensions measure uncertainty of complex networks from different perspectives. In the classical information dimension, the uncertainty of complex networks is represented by ration of number of boxes. These boxes are indiscriminate used, even if different number of nodes is covered in these boxes. The information dimension is modified to real complex networks in our method. These boxes covering nodes are viewed separately in the proposed information dimension since nodes number are different for these boxes. The uncertainty of complex networks is ration of number of nodes. In Ref. [34], the coverage ability for given box size is considered and denoted as N_b/n, although this measure N_b/n familiar to the box-covering dimension N_b. For given box size, N_b/n has unique value. However, for given box size, $N_i(l)/n$ have many data values in our method since every box is treated differently. In the classical box dimension, coverage ability of all boxes $N_b(l)$ are the same. Let $N_i(l)/n = 1/N_b(l)$ in our method, values of d'_I and d_b are equal. In a word, these dimensions are to differ slightly to each other, even if them looks very familiar.

It is an important issue that how to describe the fractal and self-similarity of complex network. Some different physical quantities are considered in the definitions of fractal, such as the box-covering fractal dimension [28] and the classical information dimension [41]. In this Letter, a new and easy-operating algorithm of information dimension is proposed based on the box-covering algorithm for complex networks. Computational results of real networks show that the proposed method is effective and flexible. The proposed method is useful to reveal fractal property of complex networks.

Acknowledgements

We greatly appreciate the editor for the encouragement and the anonymous reviewers for their helpful comments and suggestions. The work is partially supported by National Natural Science Foundation of China (Grant Nos. 61174022, 71271061 and 61364030), National Key Technology R&D Program (Grant No. 2012BAH07B01), National High Technology Research and Development Program of China (863 Program) (No. 2013AA013801), the Fundamental Research Funds for the Central Universities, Grant No. XDJK2012D009.

References

[1] M. Newman, The structure and function of complex networks, SIAM Rev. (2003) 167–256.
[2] W. Yu, G. Chen, J. Lü, On pinning synchronization of complex dynamical networks, Automatica 45 (2009) 429–435.
[3] Q. Song, J. Cao, F. Liu, Synchronization of complex dynamical networks with nonidentical nodes, Phys. Lett. A 374 (2010) 544–551.
[4] C. Hu, J. Yu, H. Jiang, Z. Teng, Synchronization of complex community networks with nonidentical nodes and adaptive coupling strength, Phys. Lett. A 375 (2011) 873–879.
[5] M. Vidal, M. Cusick, A. Barabási, Interactome networks and human disease, Cell 144 (2011) 986–998.
[6] D. Wei, X. Deng, Y. Deng, S. Mahadevan, Identifying influential nodes in weighted networks based on evidence theory, Physica A 392 (2013) 2564–2575.
[7] Q. Zhang, J. Chen, L. Wan, Impulsive generalized function synchronization of complex dynamical networks, Phys. Lett. A 377 (2013) 2754–2760.
[8] D. Watts, S. Strogatz, Collective dynamics of 'small-world' networks, Nature 393 (1998) 440–442.
[9] A. Barabási, R. Albert, Emergence of scaling in random networks, Science 286 (1999) 509–512.
[10] S. Fortunato, Community detection in graphs, Phys. Rep. 486 (2010) 75–174.
[11] B. Mandelbrot, D. Passoja, A. Paullay, Fractal character of fracture surfaces of metals, Nature 308 (1984) 721–722.
[12] M. Locci, G. Concas, R. Tonelli, I. Turnu, Three algorithms for analyzing fractal software networks, WSEAS Trans. Inf. Sci. Appl. 7 (2010) 371–380.
[13] M.A. Moret, L.Q. Antonio, H.B. Pereira, Classical and fractal analysis of vehicle demand on the ferry-boat system, Physica A 391 (2012) 1657–1661.
[14] I.I. Smolyaninov, Metamaterial model of fractal time, Phys. Lett. A 376 (2012) 1315–1317.
[15] F.O. Redelico, A.N. Proto, Empirical fractal geometry analysis of some speculative financial bubbles, Physica A 391 (2012) 5132–5138.
[16] V. Rajagopalan, Z. Liu, D. Allexandre, L. Zhang, X. Wang, E.P. Pioro, G.H. Yue, Brain white matter shape changes in amyotrophic lateral sclerosis (als): a fractal dimension study, PLoS ONE 8 (2013) e73614.
[17] P. Chełminiak, Emergence of fractal scale-free networks from stochastic evolution on the Cayley tree, Phys. Lett. A 377 (2013) 2846–2850.
[18] R.B. de Benicio, T. Stošić, P. de Figueirêdo, B.D. Stošić, Multifractal behavior of wild-land and forest fire time series in Brazil, Physica A 392 (2013) 6367–6374.
[19] C. Song, S. Havlin, H. Makse, Self-similarity of complex networks, Nature 433 (2005) 392–395.
[20] G. Long, X. Cai, The fractal dimensions of complex networks, Chin. Phys. Lett. 26 (2009) 088901.
[21] K. Anand, G. Bianconi, S. Severini, Shannon and von Neumann entropy of random networks with heterogeneous expected degree, Phys. Rev. E 83 (2011) 036109.
[22] K. Zhao, M. Karsai, G. Bianconi, Entropy of dynamical social networks, PLoS ONE 6 (2011) e28116.
[23] G. Vitiello, Fractals, coherent states and self-similarity induced noncommutative geometry, Phys. Lett. A 376 (2012) 2527–2532.
[24] F. Comellas, A. Miralles, H. Liu, Z. Zhang, The number of spanning trees of an infinite family of outerplanar, small-world and self-similar graphs, Physica A 392 (2013) 2803–2806.
[25] S. Hwang, D.-S. Lee, B. Kahng, Origin of the hub spectral dimension in scale-free networks, Phys. Rev. E 87 (2013) 022816.
[26] F.N. Silva, L.d.F. Costa, Local dimension of complex networks, preprint, 2012, arXiv:1209.2476.
[27] C. Song, S. Havlin, H. Makse, Origins of fractality in the growth of complex networks, Nat. Phys. 2 (2006) 275–281.
[28] C. Song, L. Gallos, S. Havlin, H. Makse, How to calculate the fractal dimension of a complex network: the box covering algorithm, J. Stat. Mech. Theory Exp. (2007) P03006.
[29] J. Kim, K.-I. Goh, G. Salvi, E. Oh, B. Kahng, D. Kim, Fractality in complex networks: critical and supercritical skeletons, Phys. Rev. E 75 (2007) 016110.
[30] L. Gallos, C. Song, H. Makse, A review of fractality and self-similarity in complex networks, Physica A 386 (2007) 686–691.
[31] L. Gao, Y. Hu, Z. Di, Accuracy of the ball-covering approach for fractal dimensions of complex networks and a rank-driven algorithm, Phys. Rev. E 78 (2008) 046109.
[32] H. Ng, H. Abderrahmane, K. Bates, N. Nikiforakis, The growth of fractal dimension of an interface evolution from the interaction of a shock wave with a rectangular block of sf6, Commun. Nonlinear Sci. Numer. Simul. 16 (2011) 4158–4162.
[33] C. Schneider, T. Kesselring, J. Andrade Jr., H. Herrmann, Box-covering algorithm for fractal dimension of complex networks, Phys. Rev. E 86 (2012) 016707.
[34] I. Turnu, G. Concas, M. Marchesi, R. Tonelli, The fractal dimension of software networks as a global quality metric, Inf. Sci. 245 (2013) 290–303.
[35] I. Turnu, G. Concas, M. Marchesi, R. TonelliI, Entropy of some CK metrics to assess object-oriented software quality, Int. J. Softw. Eng. Knowl. Eng. 23 (2013) 173–188.
[36] H. Zhang, X. Lan, D. Wei, Y. Deng, Self-similarity in complex networks: from the view of the hub repulsion, Mod. Phys. Lett. A 27 (2013) 1350201.
[37] D. Wei, Q. Liu, H. Zhang, Y. Hu, Y. Deng, S. Mahadevan, Box-covering algorithm for fractal dimension of weighted networks, Sci. Rep. 3 (2013) 3049.
[38] O. Shanker, Defining dimension of a complex network, Mod. Phys. Lett. A 21 (2007) 321–326.
[39] O. Shanker, Graph zeta function and dimension of complex network, Mod. Phys. Lett. A 21 (2007) 639–644.
[40] L. Lacasa, J. Gómez-Gardeñes, Correlation dimension of complex networks, Phys. Rev. Lett. 110 (2013) 168703.
[41] H. Zhu, C. Ji, Fractal Theory and Its Applications, Sciences Press, 2011.
[42] A. Renyi, Dimension, entropy and information, in: Trans. 2nd Prague Conf. Information Theory, 1960, pp. 545–556.
[43] H. Jürgens, H. Peitgen, D. Saupe, Chaos and Fractals: New Frontiers of Science, Springer Verlag, 1992.
[44] A. Bunde, S. Havlin, Fractals in Science: With a MS-DOS Program Diskette, Springer Verlag, 1994.
[45] W. Pedrycz, A. Bargiela, Fuzzy fractal dimensions and fuzzy modeling, Inf. Sci. 153 (2003) 199–216.