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]:

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

where $\varepsilon$ is square size, $N_\varepsilon$ 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/

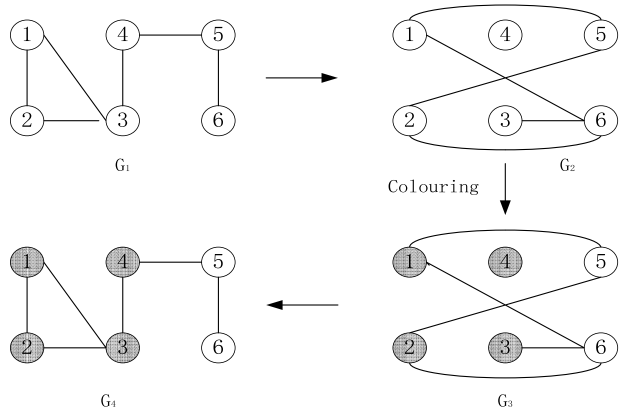


**Fig. 1.** The idea of the classical box-covering algorithm for complex networks, where $l = 3$. The network $G_1$ is original network with 6 nodes and 6 edges. The network $G_2$ is obtained by only connecting to nodes which distance between them not less than 3 in network $G_1$. The network $G_3$ is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