Physica A 450 (2016) 682–686

Contents lists available at ScienceDirect

Physica A

journal homepage: www.elsevier.com/locate/physa

A method of characterizing network topology based on the breadth-first search tree

Bin Zhou[a,*], Zhe He[b], Nianxin Wang[a], Bing-Hong Wang[b]

[a] School of Economics and Management, Jiangsu University of Science and Technology, ZhenJiang, 212003, China

[b] Department of Modern Physics and Nonlinear Science Center, University of Science and Technology of China, Hefei, 230026, China

HIGHLIGHTS

- A method is proposed to deeply characterize network topology.
- A similarity coefficient is defined to quantitatively distinguish networks.
- The similarity coefficient can quantitatively measure the topology stability of the network generated by a model.
- The network generated by a mode is more and more stable with the increasing of the network scale.
- For a network model, a broader node degree distribution will make the network generated by the model more unstable.

ARTICLE INFO

Article history:
Received 25 February 2015
Received in revised form 21 December 2015
Available online 28 January 2016

Keywords:
complex networks
Network topology
Network similarity

ABSTRACT

A method based on the breadth-first search tree is proposed in this paper to characterize the hierarchical structure of network. In this method, a similarity coefficient is defined to quantitatively distinguish networks, and quantitatively measure the topology stability of the network generated by a model. The applications of the method are discussed in ER random network, WS small-world network and BA scale-free network. The method will be helpful for deeply describing network topology and provide a starting point for researching the topology similarity and isomorphism of networks.

Generally speaking, network is a set of interconnected nodes, where a node is an element of a natural or man-made system. Network science is an emerging, highly interdisciplinary research area that aims to develop theoretical and practical approaches for understanding the natural and man-made systems. The last decade has witnessed the birth of a new movement of interest and research in the study of complex networks [1]. The study of complex networks is pervading all kinds of sciences today, from physical, biological to social science [2,3], and the network application is also studied [4–7]. Many real complex networks have emerged some common characteristics, such as small world [8], scale-free [9]. Therefore, some important network models with real network characteristics have been proposed. For example, BA scale-free network [9], ER random network [10] and WS small-world network [11].

The development of network science depends on the precise anatomy of network topology. The network topology always affects the function and the behavior of a dynamic system [12]. For example, the topology of social networks affects the spread of information and disease [13], and the topology of the power grid affects the robustness and stability of power transmission [14]. The networks in the ensemble with the same degree distribution could have different connection details, which could lead to different dynamics phenomena. The apparent ubiquity of complex networks leads to a fascinating set

* Corresponding author.
E-mail address: binzhou@mail.ustc.edu.cn (B. Zhou).

http://dx.doi.org/10.1016/j.physa.2015.12.160

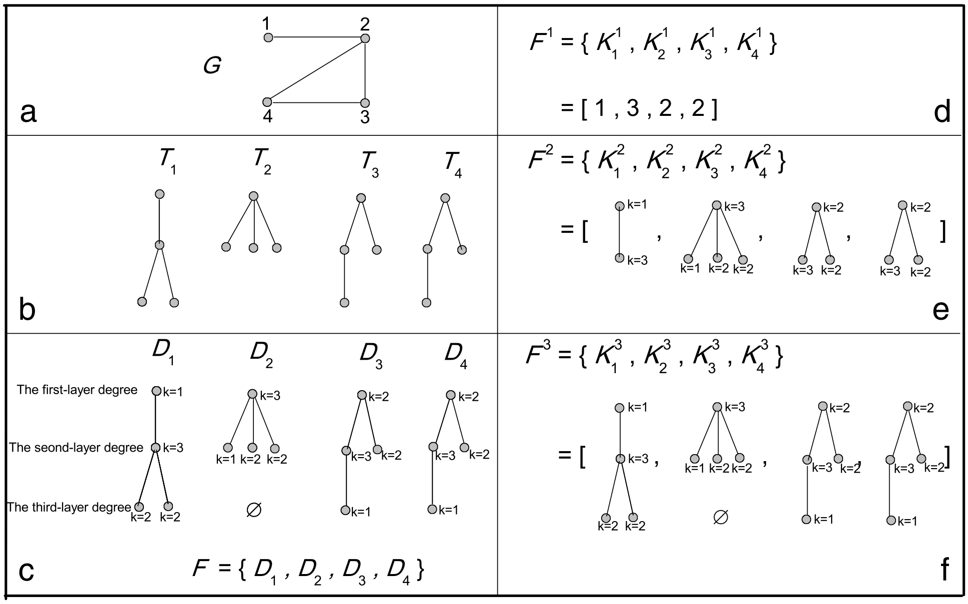

Fig. 1. The schematic illustration of the method. G is a network with four nodes and four edges. The number around a node is the degree of the node in D_1, D_2, D_3 and D_4. D_2 is a two-layer degree-tree, and the third-layer of D_2 is defined null, labeled $\varnothing$. K_2^2, K_3^2, K_3^3 are correspondingly identical to K_2^3, K_4^2, K_4^3, respectively.

of common problems concerning how the network structure facilitates and constraints the network dynamics. Therefore, it is important to characterize the topology of complex networks appropriately.

The research on complex networks begins with the effort of defining concepts and measures to characterize the topology of real networks, such as the degree distributions, degree correlations, average path length, network diameter, clustering coefficient, betweenness and modularity [15,16]. In this paper, a method will be proposed to describe deeply the network topology, and a similarity coefficient is defined to quantitatively distinguish networks and quantitatively measure the topology stability of the network generated by a model. The applications of our method will also be presented on ER random networks, WS small-world networks and BA scale-free networks.

1. Method

In a network, a node can be taken as a root, labeled i. Starting from the root i, a breadth-first search tree can be built, labeled T_i. T_i has a hierarchical structure and contains all the nodes of the network, but the degree of each node cannot be contained in T_i. We make a special provision that the degree of each node in the network is also contained in the breadth-first search tree, and the new breadth-first search tree is called breadth-first search degree-tree(BFSDT), labeled D_i. Therefore, the degrees of all nodes in the network are layered with the breadth-first search tree. The degree-trees of all nodes can be composed to a forest, signed $F = \{D_i, i = 1, \ldots, N\}$, where N is the number of nodes in the network. The forest F can deeply characterize the network topology.

Two concepts of the method are defined as follows:

(i) n-layer degree-tree: The sub-degree-tree which includes the part from the 1th layer to the nth layer in D_i is called n-layer degree-tree, labeled K_i^n, and n is a positive integer. Therefore, K_i^n has a hierarchical structure and is a sub-degree-tree with degrees associated with its nodes.

(ii) Identical n-layer degree-trees: For $\forall K_i^n, K_j^n$, if K_i^n is isomorphic to K_j^n ($K_i^n \cong K_j^n$), and the degrees of any two nodes which meet the relationship of one-to-one isomorphic mapping in K_i^n and K_j^n are equal, then we define K_i^n and K_j^n as the identical n-layer degree-trees.

Because K_i^n is a sub-degree-tree of D_i, the n-layer degree-trees of all degree-trees constitute a set, labeled $F^n = \{K_i^n, i = 1, \ldots, N\}$. F^n can characterize the network topology, and the topology can be characterized better and better with the increasing of n. Specially, for one-layer degree, $n = 1$, $K_i^n = K_i^1$, and $F^n = F^1 = \{K_i^1, i = 1, \ldots, N\}$. One-layer degree-tree K_i^1 is the degree of node i and F^1 is a set of degrees for all nodes. Therefore, general degree is a special case in the method. The schematic illustration of the method is shown in Fig. 1.

For two given networks G_1 and G_2, the n-layer degree-trees of all degree-trees from G_1 constitute a set, labeled F_1^n and the n-layer degree-trees of all degree-trees from G_2 constitute a set, labeled F_2^n. According to the concept of identical n-layer degree-trees, a similarity coefficient can be defined to quantitatively measure the similarity of n-layer degree-trees between

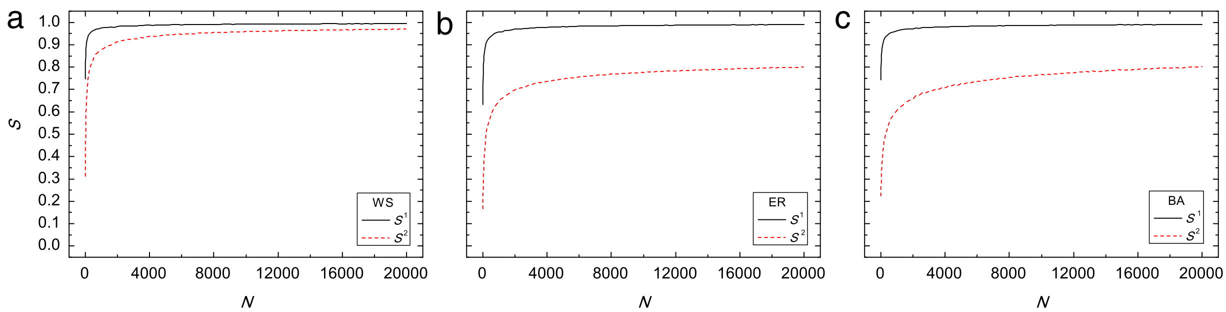

Fig. 2. The comparisons between S^1 and S^2 for ER random networks, WS small-world networks and BA scale-free networks. In the three sub-figures, N represents the number of nodes from 10 to 20 000. The black solid line and the red dashed line represent the corresponding results of S^1 and S^2, respectively. For WS small-world networks, the average degree is 2 and the probability of rewiring each edge is 0.5. For ER random networks, the probability of connecting any two nodes is $2/(N-1)$. For BA scale-free networks, starting with a globally coupled network of three nodes, we add a new node with 1 edge in each time step. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

F_1^n and F_2^n,

$$S^n = \frac{2\alpha^n}{N_1 + N_2}, \quad (0 \le S^n \le 1) \tag{1}$$

where S^n is the similarity coefficient of n-layer degree-trees between F_1^n and F_2^n, and α^n is the pairs of identical n-layer degree-trees between F_1^n and F_2^n. N_1 and N_2 are the total number of nodes in networks G_1 and G_2, respectively. S^n can also be used to quantitatively distinguish the networks G_1 and G_2, and the two networks can be distinguished better for a larger n.

Two corollaries can be obtained from the method.

1 The first corollary is stated as follows:
$S^n \ge S^{n+1} (n \ge 1)$, for $\forall G_1, G_2$.

Proof. For $\forall G_1, G_2, S^n$ and S^{n+1} can be written as

$$S^n = \frac{2\alpha^n}{N_1 + N_2}, \qquad S^{n+1} = \frac{2\alpha^{n+1}}{N_1 + N_2}, \tag{2}$$

where α^n is the pairs of identical n-layer degree-trees between G_1 and G_2, and α^{n+1} is the pairs of identical $(n+1)$-layer degree-trees between G_1 and G_2. The two definitions of both n-layer degree-tree and identical n-layer degree-trees imply that

$$\alpha^n \ge \alpha^{n+1}, \tag{3}$$

then we get

$$S^n \ge S^{n+1}. \tag{4}$$

2 The second corollary is stated as follows:
If $S^{n+1} = 1$, then $S^n = 1 (n \ge 1)$, for $\forall G_1, G_2$.

Proof. If $S^{n+1} = 1$, we invoke the result of Corollary 1 that

$$S^n \ge S^{n+1}, \tag{5}$$

$$\because\ 0 \le S^n \le 1, \qquad 0 \le S^{n+1} \le 1, \tag{6}$$

$$\therefore\ S^n = 1. \tag{7}$$

2. Application of the method

The application of the method is studied in ER random network, WS small-world network and BA scale-free network. In the three network models, when the parameters for each model are fixed, each model can generate a lot of networks with the same degree distribution, but the topologies are different. The reason is that Monte Carlo methods is used in the three network models. According to the method above, we specially calculate S^1 and S^2 to both quantitatively distinguish these networks generated by each model and quantitatively measure the topology stability of the network generated by each model.

Fig. 2 shows the comparisons between S^1 and S^2 for WS small-world network, ER random network and BA scale-free network. In the three sub-figures, for $\forall N$, $S^1 \ge S^2$, and this verifies Corollary 1. S^1 reaches a stable value more quickly than

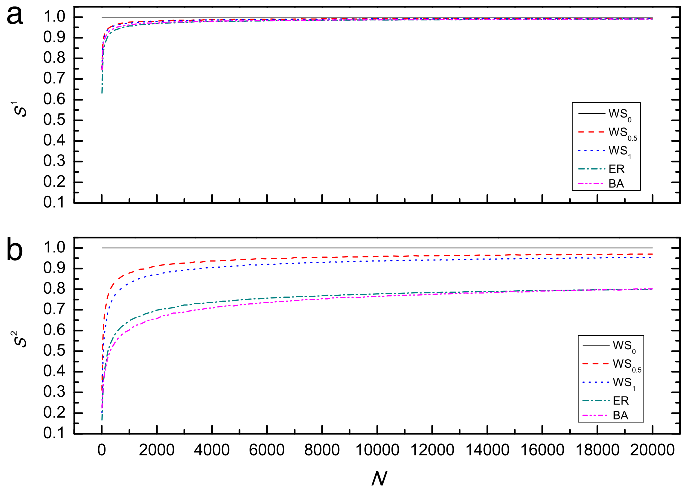

Fig. 3. S^1 and S^2 for ER random networks, WS small-world networks and BA scale-free networks with the same average degree 2. The range of N is from 10 to 20 000. In the two sub-figures, the black solid line represents the results of WS small-world networks with the probability of rewiring each edge 0, labeled WS_0. The red dashed line represents rewiring each edge 0.5, labeled $WS_{0.5}$. The blue dot line represents that of rewiring each edge 1, labeled WS_1. The olive dash-dot line represents the results of ER random networks, labeled ER. The magenta dash-dot-dot line represents the results of BA scale-free networks, labeled BA. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

S^2 with the increasing of N. When $N > 20$, $S^1 \to 1$ means the one-layer degree-trees between these networks are almost identical, and S^1 cannot distinguish these networks well. However, when $N = 20$, $S^2 < 0.5$ means the two-layer degree-trees of these networks are much different, and S^2 can distinguish these networks well. When $N = 20\,000$, $S^2 \to 1$, this means that the two-layer degree-trees of these networks are almost identical, and S^2 cannot distinguish these networks. In this case, we need to calculate S^3 to distinguish these networks. Therefore, S^n can quantitatively distinguish these networks with the same degree distribution, and these networks can be distinguished better for a larger n.

Fig. 3 shows S^1 and S^2 for ER random network, WS small-world network and BA scale-free network. In Fig. 3(a), when $N > 1000$, the four curves of $WS_{0.5}$, WS_1, ER and BA networks approximately tend to $S^1 = 1$. Therefore, S^1 cannot distinguish these networks well. In Fig. 3(b), the four curves of $WS_{0.5}$, WS_1, ER and BA networks tend to be stable very slowly, and S^2 can distinguish these networks well. We easily find that

$$\forall N, \quad S^2_{WS_0} > S^2_{WS_{0.5}} > S^2_{WS_1} > S^2_{ER} > S^2_{BA}. \tag{8}$$

The inequalities indicate that the topology of WS_0 small-world network is the most stable and that the topology of BA scale-free network is the most unstable. Therefore, S^n can quantitatively measure the topology stability of the network generated by a model. The larger n is, the better the network topology stability can be measured, and the larger S^n is, the more stable the network generated by the model is. When $N > 10\,000$, S^2 for $WS_{0.5}$, WS_1, ER and BA networks tend to be stable, and these indicate that if we want to get a stable network generated by a model, the network scale must be large enough. That is to say, the network topology is more and more stable with the increasing of network size.

Fig. 4 shows the influences of the degree distribution on S^1 and S^2 for ER random network, WS small-world network and BA scale-free network. In Fig. 4(a), we can find that when $P = 0.5$, S^1 is minimum, implying the network generated by ER model is the most unstable. The reason is analyzed as follows. The degree distribution of ER random network follows a binomial distribution. The average degree and the variance of degree distribution for ER random network can be written as

$$\bar{k} = p(N-1), \tag{9}$$

$$\sigma_k^2 = p(1-p)(N-1). \tag{10}$$

Given ER random network with size N, when $p = 0.5$, σ_k^2 is maximum. It implies that the node degree distribution of ER network is the broadest and the pairs of identical one-layer degree-trees between two networks generated by ER model are minimal in the case of $p = 0.5$. Therefore, S^1 is minimum in the case of $p = 0.5$, and a broader degree distribution will lead to a smaller S^1. In Fig. 4(b), we can find that S^2 is decreasing with the increasing of the probability p. The reason is that a larger probability of rewiring each edge will make the degree distributed more broadly [17], and a broader degree distribution leads to a smaller S^2. In Fig. 4(c), we can find S^1 is also decreasing with the increasing of m. The reason is also relative to the degree distribution [17] and is the same to Fig. 4(b). Therefore, the degree distribution plays an important role in the topology stability of the network generated by a model. The broader the degree distribution is, the more unstable the network topology is.

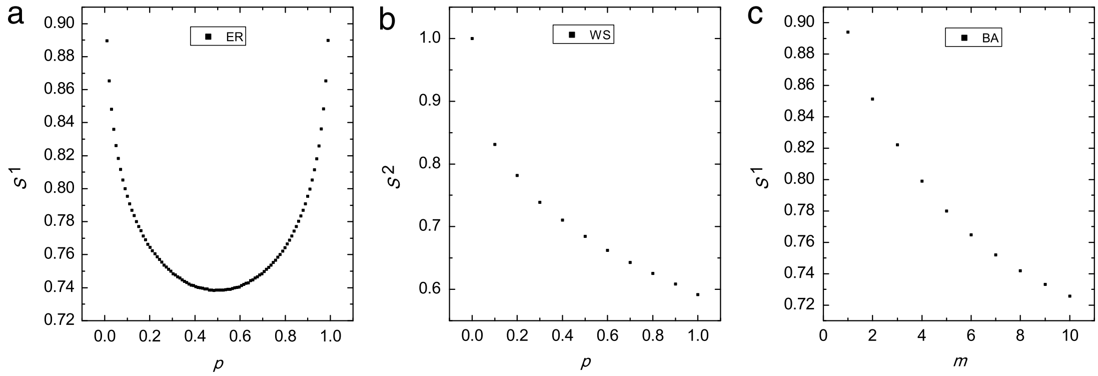

Fig. 4. The influences of the degree distribution on S^1 and S^2 for ER random network, WS small-world network and BA scale-free network. The scale of the network is $N = 100$. For ER random network, p is the connection probability between any two nodes, $p \in (0, 1)$. For WS small world network, p is the probability of rewiring each edge, $p \in [0, 1]$. For BA scale-free network, starting with a globally coupled network of m_0 nodes, we add a new node with m edges at every time step, $m \in [1, 10]$, $m_0 = 2m + 1$.

3. Conclusions

In this paper, we have proposed a method based on the breadth-first search tree to describe deeply network topology. On one hand, a similarity coefficient has been defined to quantitatively measure the similarity of n-layer degree-trees between networks. Two corollaries are given and proved. On the other hand, the applications of the method have been studied. First, the comparisons between S^1 and S^2 are done for ER random networks, WS small-world networks and BA scale-free networks. S^n can quantitatively distinguish these networks with the same degree distribution, and the networks can be distinguished better for a larger n. Secondly, the comparisons of S^1 and the comparisons of S^2 are done among ER random network, WS small-world network and BA scale-free network. S^n can quantitatively measure the topology stability of the network generated by a model. The larger n is, the better the network topology stability can be measured, and the larger S^n is, the more stable the network generated by the model is. In addition, a network generated by a mode is more and more stable with the increasing of the network size. Finally, the influences of the degree distribution on S^1 and S^2 are discussed for ER random network, WS small world network and BA scale-free network. S^1 and S^2 are smaller for a broader node degree distribution, and this implies that a broader degree distribution can make the network topology generated by a model more unstable. In conclusion, the method is helpful for deeply characterizing the network topology, and provides a starting point for researching the similarity and isomorphism between two networks.

Acknowledgments

This work was supported by The National Natural Science Foundation of China (Grant Nos. 61503159, 71331003, 71271104, 11275186 and 91024026). The numerical calculations in this paper have been done on the supercomputing system in the Supercomputing Center of University of Science and Technology of China. We thank Ming Li for useful discussions.

References

[1] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, D.-U. Hwang, Complex networks: Structure and dynamics, Phys. Rep. 424 (4) (2006) 175–308.
[2] S.H. Strogatz, Exploring complex networks, Nature 410 (6825) (2001) 268–276.
[3] C. Orsini, M.M. Dankulov, P. Colomer-de Simón, A. Jamakovic, P. Mahadevan, A. Vahdat, K.E. Bassler, Z. Toroczkai, M. Boguñá, G. Caldarelli, et al., Quantifying randomness in real networks, Nat. Commun. 6 (3) (2015) 175–185.
[4] L. Lü, M. Medo, C.H. Yeung, Y.-C. Zhang, Z.-K. Zhang, T. Zhou, Recommender systems, Phys. Rep. 519 (1) (2012) 1–49.
[5] L.D.F. Costa, O.N. Oliveira Jr., G. Travieso, F.A. Rodrigues, P.R. Villas Boas, L. Antiqueira, M.P. Viana, L.E. Correa Rocha, Analyzing and modeling real-world phenomena with complex networks: a survey of applications, Adv. Phys. 60 (3) (2011) 329–412.
[6] L. Lü, C.-H. Jin, T. Zhou, Similarity index based on local paths for link prediction of complex networks, Phys. Rev. E 80 (4) (2009) 046122.
[7] L. Lü, T. Zhou, Link prediction in complex networks: A survey, Physica A 390 (6) (2011) 1150–1170.
[8] D.J. Watts, Six Degrees, 2003.
[9] A.-L. Barabási, R. Albert, Emergence of scaling in random networks, Science 286 (5439) (1999) 509–512.
[10] P. Erdős, A. Rényi, On the evolution of random graphs, Publ. Math. Inst. Hung. Acad. Sci. 5 (1960) 17–61.
[11] D.J. Watts, S.H. Strogatz, Collective dynamics of 'small-world' networks, Nature 393 (6684) (1998) 440–442.
[12] J. Zhang, M. Small, Complex network from pseudoperiodic time series: Topology versus dynamics, Phys. Rev. Lett. 96 (23) (2006) 238701.
[13] M.E. Newman, Spread of epidemic disease on networks, Phys. Rev. E 66 (1) (2002) 016128.
[14] R. Albert, I. Albert, G.L. Nakarado, Structural vulnerability of the north American power grid, Phys. Rev. E 69 (2) (2004) 025103.
[15] L.D.F. Costa, F.A. Rodrigues, G. Travieso, P.R. Villas Boas, Characterization of complex networks: A survey of measurements, Adv. Phys. 56 (1) (2007) 167–242.
[16] M.E. Newman, The structure and function of complex networks, SIAM Rev. 45 (2) (2003) 167–256.
[17] A.-L. Barabási, R. Albert, H. Jeong, Mean-field theory for scale-free random networks, Physica A 272 (1) (1999) 173–187.