Cyclic topology in complex networks

Hyun-Joo Kim[1,2] and Jin Min Kim[2,3,*]
[1]Department of Physics Education, Korea National University of Education, Chungbuk 363-791, Korea
[2]Department of Physics and CAMDRC, Soongsil University, Seoul 156-743, Korea
[3]Asia Pacific Center for Theoretical Physics, POSTEC, Pohang, 790-784, Korea
(Received 14 November 2004; revised manuscript received 21 March 2005; published 9 September 2005)

We introduce a cyclic coefficient R which characterizes the degree of circulation in complex networks. If a network has a perfect treelike structure, then R becomes zero. The larger value of R represents that the network has more cyclic structure. We measure both the cyclic coefficients and the distributions of local cyclic coefficients for various networks and discuss the cyclic structures of them.

DOI: 10.1103/PhysRevE.72.036109 PACS number(s): 89.75.Fb, 89.20.−a, 89.75.Hc

During recent years, there have been considerable efforts in studying the structure of complex networks [1]. They appear in a variety of systems such as biological [2–4], social [5–8], informational [9,10], and economic [11] systems. Such complex networks are characterized by some topological and geometric properties such as small worlds, a high degree of clustering, and scale-free topology. The small-world property denotes that the average shortest path length L between vertex pairs in a network grows logarithmically with network size N. The clustering structure is characterized by the clustering coefficient C which is the fraction of pairs between the neighbors of a vertex that are directly connected to each other. The high degree of clustering indicates that if vertices A and B are linked to vertex C, then A and B are also likely to be linked to each other. These two properties are realized by the small-world network (SWN) model [12] in which randomly selected vertex pairs are linked by shortcuts. The scale-free (SF) topology reflects that the degree distribution $P(k)$ follows a power law $P(k)\sim k^{-\gamma}$, where the degree k is the number of edges of a given vertex and γ is the degree exponent. An evolving model introduced by Barabási and Albert (BA) [13] illustrates the SF property very well.

Recently, much attention has been focused on the structural properties of complex networks. A hierarchical structure appears in some real networks, and it has been clarified by a power-law behavior of the clustering coefficient $C(k)$ as a function of the degree k [14–19]. This indicates that the networks are fundamentally modular. It is the origin of the high degree of clustering of the networks. Also, it was recently found that many real networks include statistically significant subnetworks, so-called motifs, in their structures [20–22]. Especially, recent studies of the topological properties in complex networks have paid much attention to the loop (cycle) structure. In comparison with a treelike topology, loops provide more paths along which information or a virus can propagate. So loops can affect the delivery of information, transport process, and epidemic spreading behavior [23]. In considering the loop structure, a cycle of order k is defined as a closed loop composed of k edges. That is, a triangular structure has a cycle of order 3, and a rectangular structure has a cycle of order 4. If there is no closed loop passing through a vertex, then it is assumed that it has a cycle of infinite order. Actually, the clustering coefficient counts for the triangular structure only. However, there are many other closed loops of higher orders consisting of more than three edges. There have been some previous studies [24–27] about cycles of order 4 or 5. So it is natural to consider loops of all orders to characterize the cyclic structure.

In this paper, we introduce a new quantity R and a local quantity r which characterize the degree of circulation in complex networks in order to consider the loops of all orders from 3 up to infinity. By monitoring R and the distribution of r, we discuss the cyclic topology for both several real networks from technological to social systems and network models such as the SWN and BA models.

We define a local cyclic coefficient r_i for a vertex i as the average of the inverse size of the smallest loop that connects the vertex i and its two neighbor vertices—i.e.,

$$r_i=\frac{2}{k_i(k_i-1)}\sum_{\langle lm\rangle}\frac{1}{S^i_{lm}}, \tag{1}$$

where k_i is the degree of the vertex i and $\langle lm\rangle$ is for all the pairs of neighbors of the vertex i. S^i_{lm} is the smallest size of the closed path that passes through vertex i and its two neighbor vertices l and m. There are $k_i(k_i-1)/2$ of such pairs of neighbors. If vertices l and m are directly linked to each other, then vertices i, l, and m form a triangle. It is a cycle of order 3 and S^i_{lm} has a value 3, which is the smallest value of S. If there does not exist any path that connects vertices l and m except for the path through the vertex i, then vertices i, l, and m have a tree structure. In this case, there is no closed loop passing through the three vertices i, l, and m where S^i_{lm} is infinity. The vertex i has a cycle of infinite order. For an example shown in Fig. 1(a) the local cyclic coefficient of the vertex ● is $r_\bullet=0.13$ with $S^\bullet_{12}=3$, $S^\bullet_{23}=4$, $S^\bullet_{13}=5$, and $S^\bullet_{14}=S^\bullet_{24}=S^\bullet_{34}=\infty$.

We define a cyclic coefficient R as the average of r_i over all the vertices, $R=\langle r_i\rangle$. It has a value between zero and 1/3. $R=0$ means that the network has a perfect treelike structure without having any loops. Meanwhile, if all the neighbor pairs of the vertices have direct links to each other, then the

*Electronic address: jmkim@physics.ssu.ac.kr

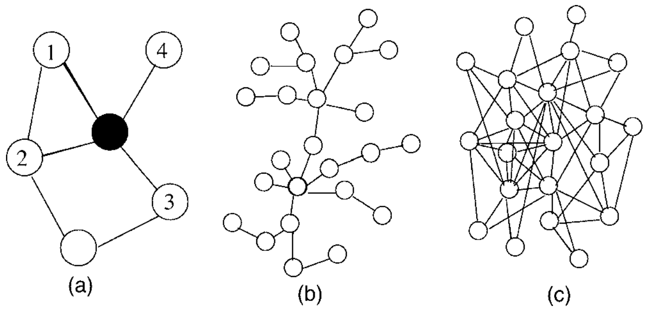

FIG. 1. (a) Typical examples which have different cyclic coefficients are shown. The local cyclic coefficient of the solid circle is $r_{\bullet}=0.13$. The two sample networks with the same network size $N=25$ and the different cyclic coefficients are shown in (b) and (c) where $R=0$ and $R=0.29$, respectively.

cyclic coefficient becomes $R=1/3$. Figures 1(b) and 1(c) show two examples of $R=0$ and $R=0.29$, respectively, for the same network size $N=25$. Thus the larger the cyclic coefficient R is, the more cyclic the network is. The cyclic coefficient R is a good quantity to identify the degree of circulation in complex networks.

In order to characterize the cyclic topology, we measure the cyclic coefficient R and the distribution of the local cyclic coefficient $P(r)$ for several real networks [28] appearing in biological, technological, and social systems. In this measurement, we exclude the isolated vertices and focus on the entirely connected part of the network.

First, we consider a protein network [3] which is composed of 1458 proteins. It has 1948 identified direct physical interactions. The proteins and direct interactions are considered as vertices and edges, respectively. Figure 2(a) shows the histogram of the distribution $P(r)$ of the local cyclic coefficient. About 60% of the total vertices have $r=0$, and $P(r)$ has a small value for the other range $0<r\leqslant 1/3$. We obtain a small value of the cyclic coefficient $R\approx 0.06$. The network has a treelike structure with few loops. Thus the treelike topology of the protein network visualized in [3] is well quantified by the distribution function $P(r)$.

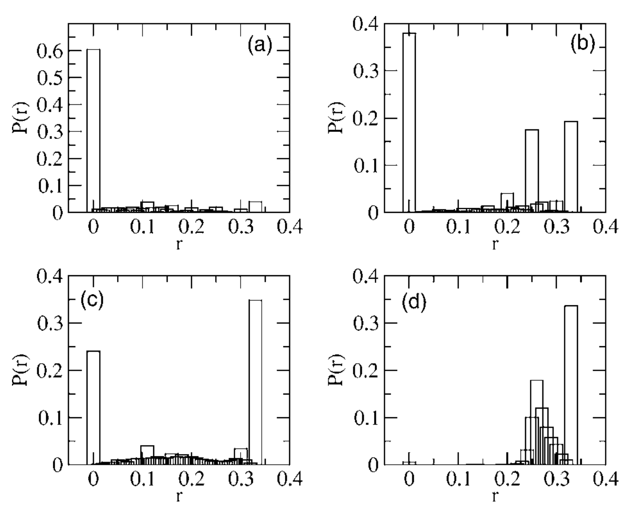

FIG. 2. The probability distribution of local cyclic coefficient for four real networks, (a) protein network, (b) internet network, (c) math coauthorship network, and (d) movie actor collaboration network.

Second, a physical Internet network [9] at the interdomain [autonomous system (AS)] level is considered. Each domain, composed of hundreds of routers and computers, acts as a vertex. An edge is drawn between two domains if there is at least one route that connects them directly. The network at the AS level of 15 September 1999 is composed of both 5746 vertices and 11 017 edges. We obtain $R\approx 0.16$ in the network. The distribution of the local cyclic coefficient is shown in Fig. 2(b); it has big three peaks at $r=0$, $r=0.25$, and $r=1/3$. That is, many vertices have tree structures ($r=0$) and the rest of the vertices have loops of small sizes (three or four).

Third, we consider a network of scientific collaborations in the field of mathematics published in the period 1991–1998 [6], in which the vertices are the scientists. They are connected if they write a paper together. The total number of vertices and edges are 57 516 and 143 778, respectively. We obtain $R\approx 0.19$ in the network. Figure 2(c) shows the probability distribution of r. It has a strong peak at $r=1/3$, which indicates that cycles of order 3 are very dominant in the networks. It is quite different from the characteristics of $P(r)$ in the protein and Internet networks where treelike structures are more dominant.

Finally, we consider a movie actor collaboration network [5] which has 9865 vertices and 273 412 edges. The actors are treated as vertices, and two vertices are linked if the corresponding actors have acted in the same movie together. As shown in Fig. 2(d), the probability distribution $P(r)$ has a maximum value at $r=1/3$, which reflects the high degree of clustering in the social network. Meanwhile, there are almost no vertices having $r=0$ in contrast to the case of the other networks. This explains that the movie actor network is more cyclic with large values of the cyclic coefficient $R\approx 0.29$.

From the results of the above four examples, we have found that both clustered parts and nonclustered parts coexist in real networks especially in the math coauthorship network. The probability distribution $P(r)$ is not uniform at all. Instead, there are a few peaks at certain values of r such as $r=0$ or $r=1/3$. It means that most of vertices have either triangle structure or tree structure with the neighbor vertices. That is, the neighbor vertices in the well-clustered parts have high connections with each other while the vertices in the nonclustered parts have tree structures. Thus by measuring the distribution of the local cyclic coefficient we can understand the details of the cyclic structure in the complex networks.

We have also considered the cyclic coefficient R for two representative models of complex networks, the SWN [12] and BA [13] models. The algorithm of the SWN model is the following: Consider a one-dimensional lattice of N vertices with periodic boundary conditions—i.e., a ring—and connect each vertex to its nearest m neighbors. The small-world model is then created by randomly rewiring each edge of the lattice with probability p, moving one end of the edge to a new vertex chosen randomly from the lattice, except that self-connections and duplicate edges are created. This rewiring process introduces $pNm/2$ shortcuts which connect the

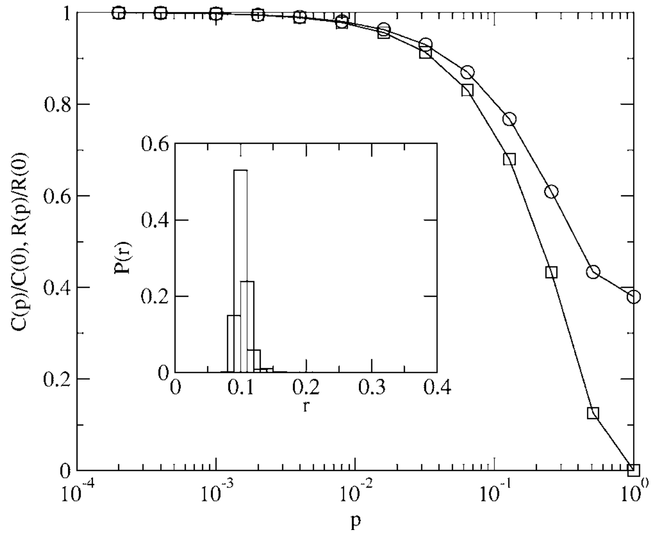

FIG. 3. The normalized cyclic coefficient $R(p)/R(0)$ (circles) and the normalized clustering coefficient $C(p)/C(0)$ (squares) are given for the SW network model where $R(0)$ and $C(0)$ are 0.283 and 0.5, respectively, for the regular network. The distribution of local cyclic coefficient for the random network ($p=1$) is shown in the inset.

vertices at long distance. The transition between regular lattice network ($p=0$) and random network ($p=1$) [29] can be shown by varying p.

Figure 3 shows the plot of the normalized clustering coefficient $C(p)/C(0)$ and the normalized cyclic coefficient $R(p)/R(0)$ as a function of the rewiring probability p with the network size $N=10\,000$ and $m=4$. The clustering coefficient stays almost unchanged for $p<0.01$ and drops to zero at $p=1$. It is the characteristics of the SWN with a high degree of clustering for $p<0.01$. The cyclic coefficient $R(p)$ also has almost the same value of $R(0)$ for $p<0.01$ while it decreases to a finite value with increasing p. The finite value of $R(1)$ comes from the contribution of loops of all orders. The inset of Fig. 3 shows a Poisson-like distribution of $P(r)$ for the random network ($p=1$). The cyclic probability distribution has a peak at $r=0.11$ with $R\approx 0.11$. It is interesting that $P(r)$ is almost zero for both $r=0$ and $r=1/3$ in the random network.

We also measure the cyclic coefficient for the BA model

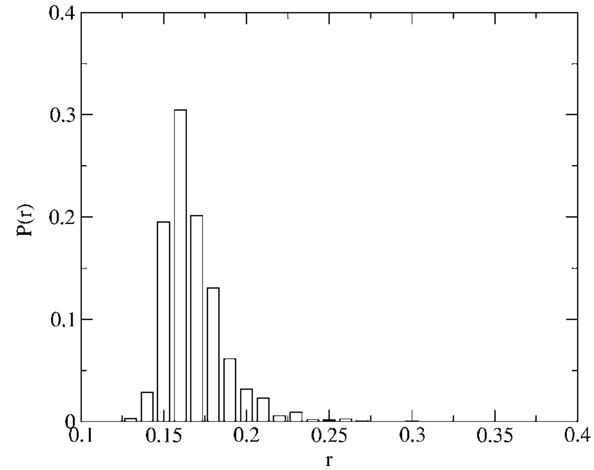

FIG. 4. The distribution of local cyclic coefficient for the BA model.

[13]. The BA model is carried out as follows: Start from a small number N_0 of vertices and no edges. At every time step, a new vertex with m ($\leqq N_0$) edges is added where the m edges link the new vertex to m different existing vertices in the system. The m vertices to which the new vertex is connected are chosen with the preferential attachment rule in which the probability Π for a vertex i to be connected with a new vertex depends on the degree k_i of the vertex i, such that $\Pi(k_i)=k_i/\Sigma_j k_j$. We have obtained $R\approx 0.17$ with the network size $N=10\,000$ for the BA model. As shown in Fig. 4, the distribution of the local cyclic coefficient in the BA model shows a Poisson-like shape having a peak at $r=0.16$. However, in the real networks given above, the probability distributions $P(r)$ do not follow a Poisson-like shape and have a peak at either $r=0$ or $r=1/3$. There exists neither triangle nor tree structure in the BA model, in contrast to the case of the real networks. The Poisson-like distribution of $P(r)$ is one of the specific characteristics of the BA networks.

We summarized the various data of the network size N, the mean degree $\langle k\rangle$, the clustering coefficient C, the cyclic coefficient R, the cyclic probability distribution $P(0)$ with $r=0$ (tree structure), and $P(1/3)$ with $r=1/3$ (cyclic structure of loops with length 3) in Table I for various networks.

In conclusion, we introduced a cyclic coefficient R to evaluate the degree of circulation and measured R in various

TABLE I. For both four real networks and two network models, we summarized the various data of the network size N, the mean degree $\langle k\rangle$, the clustering coefficient C, the cyclic coefficient R, the probability distribution $P(0)$ with $r=0$ (tree structure), and $P(1/3)$ with $r=1/3$ (cyclic structure of loops with length 3).

Network	N	⟨k⟩	C	R	$P(0)$	$P(1/3)$
Protein interactions	1458	2.67	0.07	0.06	0.60	0.04
Internet	5746	3.83	0.24	0.16	0.38	0.19
Math coauthorship	57516	5.00	0.48	0.19	0.24	0.35
Movie actor collaborations	9853	54.95	0.58	0.29	0.01	0.34
Random network ($p=1$)	10000	4	0.0003	0.11	0	0
BA network	10000	6	0.006	0.17	0	0

networks. It includes the effects of all sizes of loops. If a network has a perfect treelike structure, R becomes zero. The value of cyclic coefficient is in between zero and $1/3$. The larger the cyclic coefficient is, the more cyclic the network becomes. We measured the cyclic coefficients for various real complex networks and the representative network models. For the protein network of a biological system the cyclic coefficient is small. It reflects the fact that the network is tree like. For the movie actor collaboration network social system, we found that its structure is more cyclic with a large cyclic coefficient. Also by measuring the probability distribution of the local cyclic coefficient, we could classify the cyclic structures of the networks. Thus the cyclic coefficient and the distribution of the local cyclic coefficient help us to understand the structures of complex networks. It would be interesting to measure the cyclic coefficient for various other networks and compare it with our results.

One of us (J.M.K.) thanks the Korea Institute for Advanced Study for its hospitality during his stay. This work was supported by Korea Research Foundation Grant No. KRF-2003-015-C00003.

[1] M. E. J. Newman, SIAM Rev. **45**, 167 (2003); R. Albert and A. L. Barabábasi, Rev. Mod. Phys. **74**, 47 (2002); S. N. Dorogovtsev and F. F. Mendes, Adv. Phys. **51**, 1079 (2002).
[2] H. Jeong, B. Tombor, R. Albert, Z. N. Oltvai, and A.-L. Barabási, Nature (London) **407**, 651 (2000).
[3] H. Jeong, S. Mason, R. Albert, A.-L. Barabási, and Z. N. Oltvai, Nature (London) **411**, 41 (2001).
[4] A.-L. Barabási and Z. N. Oltvai, Nature (London) **5**, 101 (2004).
[5] M. E. J. Newmann, S. H. Strogatz, and D. J. Watts, e-print cond-mat/0007235; A.-L. Barabási and R. Albert, Science **286**, 509 (1999).
[6] M. E. J. Newman, Proc. Natl. Acad. Sci. U.S.A. **98**, 404 (2001).
[7] A.-L. Barabási, H. Jeong, Z. Néda, E. Ravasz, A. Schbert, and T. Vicsek, e-print cond-mat/0104162.
[8] F. Liljeros, C. R. Edling, L. A. N. Amaral, H. E. Stanley, and Y. Averg, Nature (London) **411**, 907 (2001).
[9] M. Faloutsos, P. Faloutsos, and C. Faloutsos, in *Proceedings of ACM SIGCOMM '99* [Comput. Commun. Rev.29, 251 (1999)]; R. Pastor-Satorras, A. Vázquez, and A. Vespignani, Phys. Rev. Lett. **87**, 258701 (2001).
[10] R. Albert, H. Jeong, and A.-L. Barabási, Nature (London) **401**, 130 (1999).
[11] H.-J. Kim, Y. Lee, B. Kahng, and I.-m. Kim, J. Phys. Soc. Jpn. **71**, 2133 (2002).
[12] D. J. Watts and S. H. Strogatz, Nature (London) **393**, 440 (1998); D. J. Watts, *Small Worlds: The Dynamics of Networks between Order and Randomness* (Princeton University Press, New York, 1999).
[13] A.-L. Barabási and R. Albert, Science **286**, 509 (1999); A.-L. Barabási, R. Albert, and H. Jeong, Physica A **272**, 173 (1999).
[14] E. Ravasz and A.-L. Barabási, Phys. Rev. E **67**, 026112 (2003).
[15] S. N. Dorogovtsev, A. V. Goltsev, and J. F. F. Mendes, Phys. Rev. E **65**, 066122 (2002).
[16] G. Szabo, M. Alava, and J. Kertész, e-print cond-mat/0208551.
[17] M. E. J. Newman, Phys. Rev. E **68**, 026121 (2003).
[18] K. Klemm and V. M. Eguiluz, Phys. Rev. E **65**, 036123 (2002).
[19] A. Vazquez, R. Pastor-Satorras, and A. Vespignani, Phys. Rev. E **65**, 066130 (2002).
[20] R. Milo, S. Shen-Orr, S. Itzkovitz, N. Kashtan, D. Chklovskii, and U. Alon, Science **298**, 824 (2002).
[21] N. Kashtan, S. Itzkovitz, R. Milo, and U. Alon (unpublished).
[22] V. P. Zhigulin, e-print cond-mat/0311330.
[23] T. Petermann and P. De Los Rios, Phys. Rev. E **69**, 066116 (2004).
[24] G. Bianconi and A. Capocci, Phys. Rev. Lett. **90**, 078701 (2003).
[25] G. Caldarelli, R. Pastor-Satorras, and A. Vespignani, e-print cond-mat/0212026.
[26] H. D. Rozenfeld, J. E. Kirk, E. M. Bolt, and D. ben-Avraham, e-print cond-mat/0403536.
[27] A. Fronczak, J. A. Holyst, M. Jedynak, and J. Sienkiewicz, Physica A **316**, 688 (2002).
[28] The data of the real networks used in this paper are publicly available at http://www.nd.edu/~networks and http://stat.kaist.ac.kr
[29] P. Erdós and A. Rényi, Publ. Math. (Debrecen) **6**, 290 (1959); B. Bollobás, *Random Graphs* (Academic Press, London, 1985).