Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
Physica A 351 (2005) 671–679
PHYSICA A
www.elsevier.com/locate/physa

Modelling hierarchical and modular complex networks: division and independence

D.-H. Kim[a], G.J. Rodgers[b], B. Kahng[a,*], D. Kim[a]

[a]School of Physics and Center for Theoretical Physics, Seoul National University NS50, Seoul 151-747, Republic of Korea
[b]Department of Mathematical Sciences, Brunel University, Uxbridge, Middlesex, UB8 3PH, UK

Received 6 June 2004
Available online 8 January 2005

Abstract

We introduce a growing network model which generates both modular and hierarchical structure in a self-organized way. To this end, we modify the Barabási–Albert model into the one evolving under the principles of division and independence as well as growth and preferential attachment (PA). A newly added vertex chooses one of the modules composed of existing vertices, and attaches edges to vertices belonging to that module following the PA rule. When the module size reaches a proper size, the module is divided into two, and a new module is created. The karate club network studied by Zachary is a simple version of the current model. We find that the model can reproduce both modular and hierarchical properties, characterized by the hierarchical clustering function of a vertex with degree k, $C(k)$, being in good agreement with empirical measurements for real-world networks.

PACS: 89.65.−s; 89.75.Hc; 89.75.Da

Keywords: Networks; Social networks; Hierarchical and modular structure

Recently, considerable effort has been made to understand complex systems in terms of random graphs, consisting of vertices and edges [1–5]. Such complex

*Corresponding author. Tel.: +82 2 880 1326; fax: +82 2 884 3002.
E-mail address: kahng@phya.snu.ac.kr (B. Kahng).

0378-4371/$ - see front matter
doi:10.1016/j.physa.2004.12.026

networks exhibit many interesting emerging patterns as follows: first, the degree distribution follows a power-law, $P(k) \sim k^{-\gamma}$, where the degree is the number of edges connecting to a given vertex [6]. Such networks, called scale-free (SF), are ubiquitous in the real world. To illustrate such SF behavior in the degree distribution, Barabási and Albert (BA) [6] introduced an in silico model: initially, fully-connected m_0 vertices exist in a system. At each time step, a vertex is newly added and connects to m existing vertices, which are chosen with a probability linearly proportional to the degree of target vertex. Such a selection rule is called the preferential attachment (PA) rule.

Second, the degree–degree correlation in real-world networks is nontrivial. The nontrivial behavior is measured in terms of the mixing coefficient r [7], a Pearson correlation coefficient between the degrees of the two vertices on each side of an edge. Complex networks can be classified according to the mixing coefficient r into three types, having $r<0$, $r \approx 0$, and $r>0$, called the disassortative, the neutral, and the assortative network, respectively [7]. Such classifications can also be identified by a quantity, denoted by $\langle k_{\rm nn}\rangle(k)$, the average degree of a neighboring vertex of a vertex with degree k [8]. For the assortative (disassortative) network, $\langle k_{\rm nn}\rangle(k)$ increases (decreases) with increasing k, i.e., a power law $\langle k_{\rm nn}\rangle(k) \sim k^{-\nu}$ is satisfied where ν is negative (positive) for the assortative (disassortative) network [8].

Third, many real-world networks have modular structures within them. Modular structures form geographically in the Internet [9], functionally in metabolic [10] or protein interaction networks [11], or following social activities in social networks [12,13]. Such modular structures are characterized in terms of the clustering coefficient. Let C_i be the local clustering coefficient of a vertex i, defined as $C_i = 2e_i/k_i(k_i - 1)$, where e_i is the number of edges present among the neighbors of vertex i, out of its maximum possible number $k_i(k_i - 1)/2$. The clustering coefficient, C, is the average of C_i over all vertices. $C(k)$ means the clustering function, the average of C_i over the vertices with degree k. When a network is modular and hierarchical, $C(k) \sim k^{-\beta}$ and C remains finite for large system size N [10,14]. In the BA model with $\gamma = 3$, however, $C(k)$ is independent of k, but decreases with N [2,14], because the BA model does not contain modules.

In this paper, we are interested in modelling complex networks including both modular and hierarchical structure not in a deterministic way, but in a self-organized way. In real-world networks, modules represent communities which may evolve as time passes and such modules form hierarchical structure. The karate club (KC) network, originally proposed by Zachary [15], is a simple example of a real-world social network containing community structure. Recently, Newman and Girvan [12] studied the KC network to test their new algorithm for clustering communities [12,16,17]. Here we notice that the KC network contains ingredients, division and independence, which forms a modular structure, in addition to growth and PA principles as noticed in the BA model. Thus, we introduce a network model evolving by such rules, and perform numerical simulations for large system size. Indeed, we find that the model exhibits a characteristic feature of both modular and hierarchical structure, $C(k) \sim k^{-1}$, as much as those for empirical data.