Scale-free homophilic network

M. L. de Almeida[1], G. A. Mendes[2], G. M. Viswanathan[1] and L. R. da Silva[1]

[1] Departamento de Física Teórica e Experimental, Universidade Federal do Rio Grande do Norte, Natal – RN, Brazil

[2] Departamento de Física , Universidade Federal do Ceará, Fortaleza-CE, Brazil

Abstract

An important aspect governing the growth of complex networks is homophily, which is defined as the tendency of sites to link with others which are similar to themselves. Here, we modify the preferential attachment from Barabási-Albert model by including a homophilic term. Comparisons are made with the Barabási-Albert model, fitness model and our present model considering its topological properties: degree distribution, time dependence of the connectivity, shortest path length and clustering coefficient. We verify the existence of a region where the characteristics of sites play an important role in the rate of gaining links as well as in the number of links between sites with similar and dissimilar characteristics.

Introduction

Complex Network (CNs) play an important role in the understanding of many artificial and natural systems.

The Internet, *World Wide Web (WWW),* road maps and gossip propagation, are examples of systems that can be suitably associated with the theory of complex netrwork.

In its abstract form – set sites connected via links – the networks are the most general and powerful means to represent patterns of connections or interactions between parts of a system.

Recently, it has been given a good deal of attention to the affect that the characteristics of the sites put on structural formation of the network. Here we pay special attention to an important aspect that guides the growth of networks: *homophily,* which is defined as the tendency of sites connect to others who are similar to themselves.

Objective

Analize, via topological properties, the structure of the network model proposed.

Methodology

Through numerical simulations, we obtained data to analyze the topological properties of networks obtained by our model.

Model

Preferential Attachment → $$\Pi_i = \frac{(1 - A_{ij})k_i}{\sum\limits_i (1 - A_{ij})k_i}$$

Similitude → $A_{ij} = |\eta_i - \eta_j|$
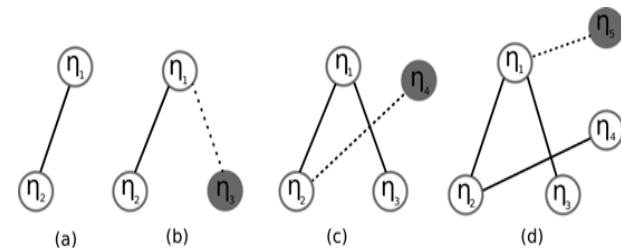
η → represents the characteristic of sites.

Illustration of how the sacle-free homophilic model grows. In (a) the initial number of sites, m_0, is two. In (b), (c) and (d) each site make only one new link. $\eta_1 = 0.5, \eta_2 = 0.1, \eta_3 = 0.6, \eta_4 = 0.2, \eta_5 = 0.4$

Measuring homophily

Network composed of sites that can assume two states *s* with values, *s = 1 or 0*. To compute the degree of homophily, consider n_1 (n_0) the number of sites with s = 1 (0).

$$N = n_1 + n_0 \quad \text{and} \quad M = m_{11} + m_{10}$$

$$m_{11} \longrightarrow (1 -- 1)$$
$$m_{10} \longrightarrow (1 -- 0)$$

If *s* is distributed randomly among the N sites, the expected values of m_{11} and m_{10} are:

$$\overline{m}_{11} = \binom{n_1}{2} \times p = \frac{n_1(n_1 - 1)}{2}p,$$

$$\overline{m}_{10} = \binom{n_1}{1}\binom{n_0}{1} \times p = n_1(N - n_1)p,$$

$$p \equiv 2M/N(N-1)$$

Connectance

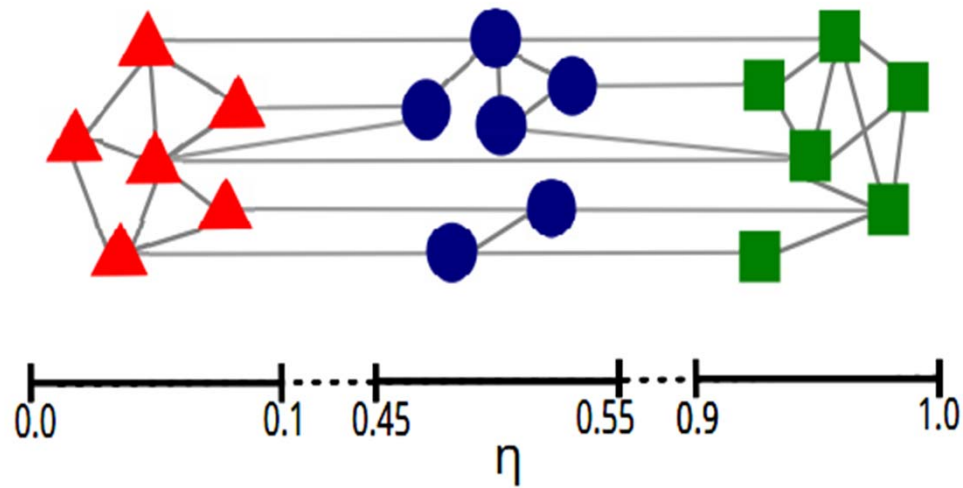If m_{11} and m_{10} calculated exibit a significant deviations their mean values, s = 1 is not random.

$$D_i \equiv \frac{m_{11}}{\overline{m}_{11}} \quad \text{and} \quad H_i \equiv \frac{m_{10}}{\overline{m}_{10}}$$

➡ Where *D* and *H* means, respectively, the dyadicity and the hererophilicity.

Results

Site characteristics distribuitions on scale-free homophilic model.

The figure shows that sites with characteristics , $\eta's$, of same intervals (similar characteristics or same forms in figure), are some tightly connected themselves and more sparsely connected with sites of different intervals (dissimilar characteristics or different forms in figure). Note that in the region, where $\eta \approx 0.5$, the sites with $\eta's$ from here, are also densely linked sites having $\eta's$ outside this region.
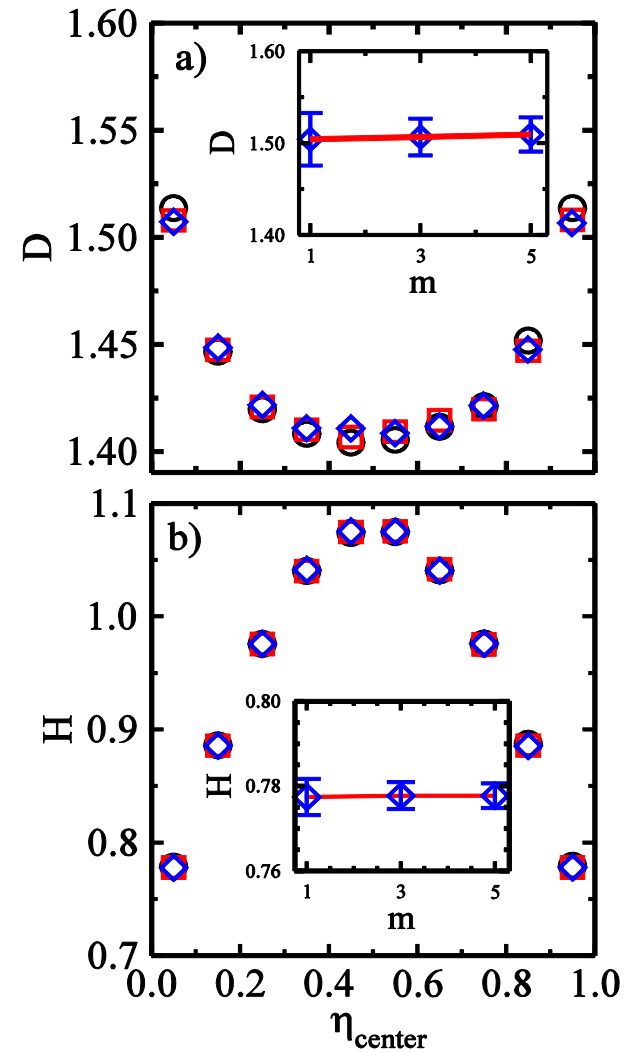
Degree Distribution

The degree distribution follows a power law $P(k) \sim k^{-\gamma}$, showing that our model is scale-free network. It means that few sites have high connectivity while many ones have a small number of nearest neighbors (heterogeneity of degree) .
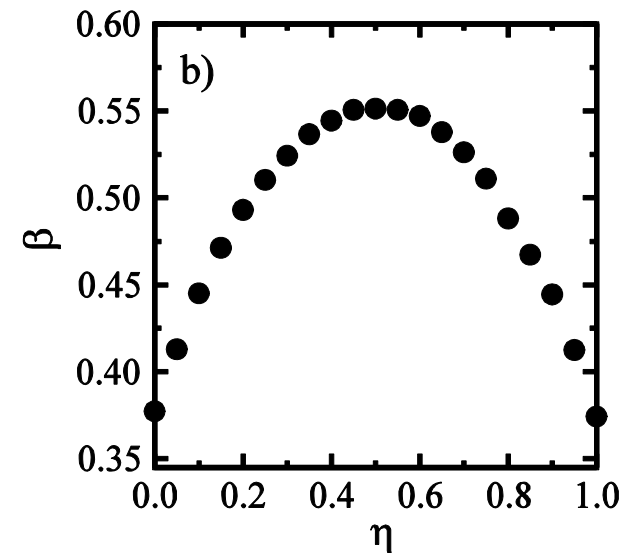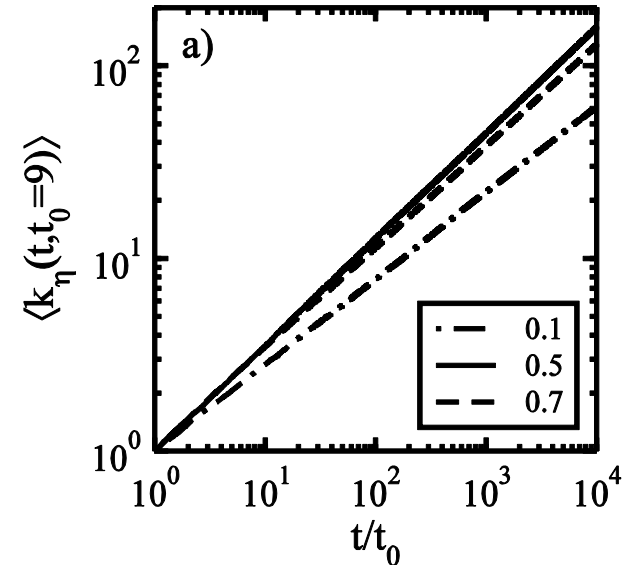

$\gamma = 2.84(1)$
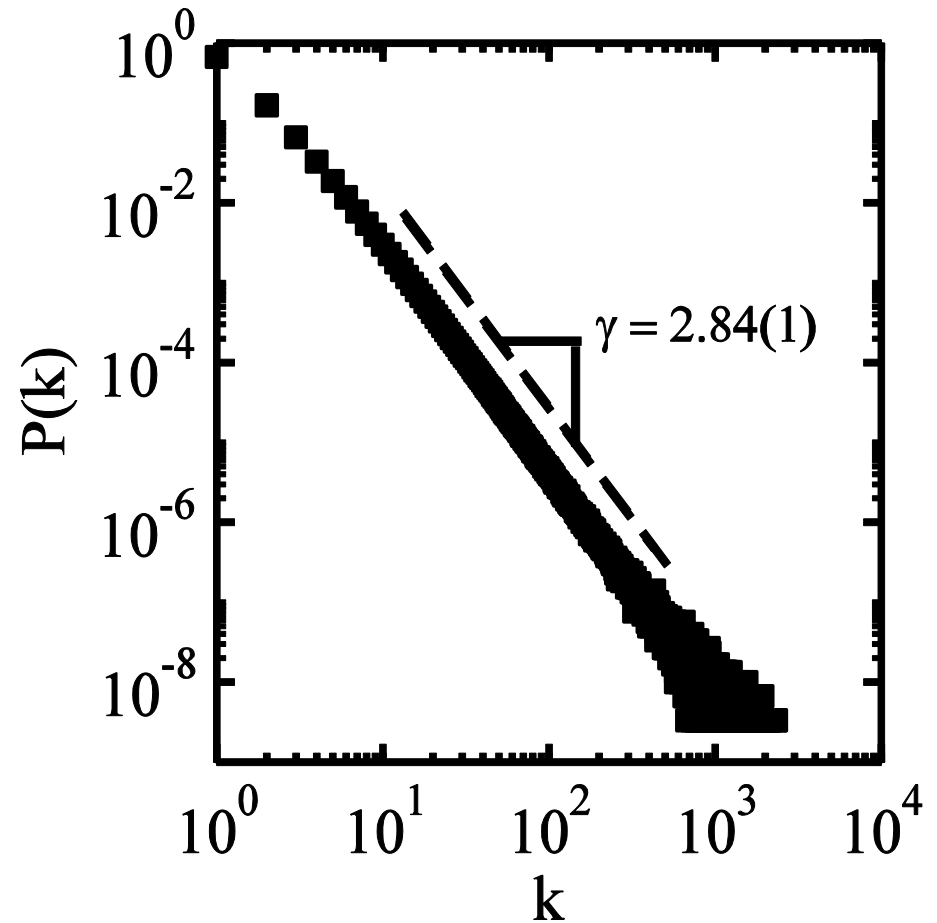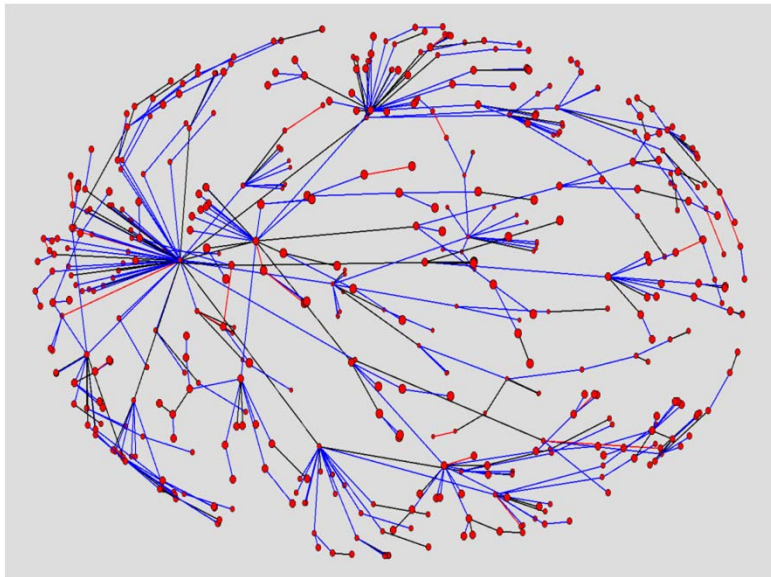
The dynamical of our model

The time dependence of the average connectivity follows a power law.

$$\left\langle k_i\left(t, t_0\right)\right\rangle \propto \left(\frac{t}{t_0}\right)^{\beta}$$

Dynamic exponent β, indicating the rate at which sites get connections.It reveals that the sites with $\eta \approx 0.5$ acquire, on average, more links.

The average path length

Note that the average path length increases approximately logarithmically with N, indicating the small-world character of the scale-free homophilic model in accordance with empirical results real networks.
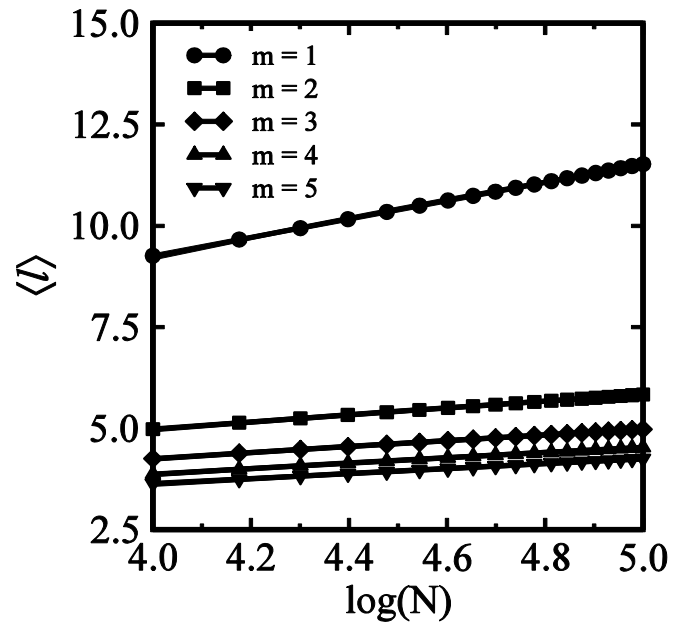
Clustering coefficient

The clustering coefficient follows a power law, meaning that the probability to find two nearest neighbors, of a given site, connected themselves decreases as the network grows.
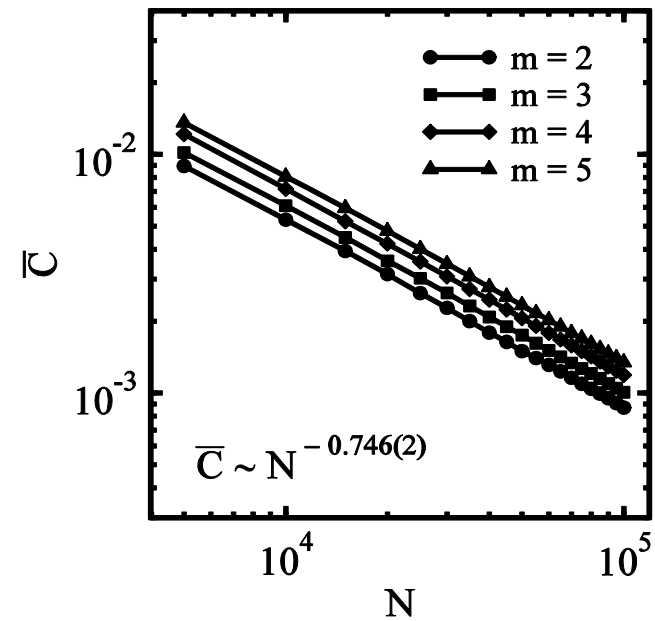
Conclusions

We have shown that the proposed model exhibits five common features of network which were studied in the present work: homophily, heterogeneity in degree, preferential attachment, adding of sites and links and small-world character. Our main finding is that when the value of site characteristic is 0.5, interesting behaviors are observed: dyadicity (D) is minumum, the heterophilicity (H) is maximum and the rate of links that a site specific acquires is maximum.

References

A.-L. Barabási, R. Albert, Science **286**, 509 (1999)

G. Bianconi, A.-L. Barabási, Europhys. Lett. **54**, 036112 (2001)

J. Park, A.-L. Barabási, Proc. Nat. Acad. Sci. USA **104**, 17916 (2007)

M. L. de Almeida, G. A Mendes, G. M. Viswanathan, L. R. da Silva, accepted for EPJ B, (2012)