Assortativity Properties of Scale-Free Networks

Vadim Shergin, Larysa Chala, Serhii Udovenko
Kharkiv National University of Radio Electronics
Kharkiv, Ukraine
vadim.shergin@nure.ua, larysa.chala@nure.ua, serhiy.udovenko@hneu.net

***Abstract*— Nodes distribution by degrees is the most important characteristic of complex networks, but not comprehensive one. While degree distribution is a first order graph metric, the assortativity is a second order one. Assortativity coefficient is a measure of a tendency for nodes in networks to connect with similar or dissimilar ones in some way. As a simplest case, assortative mixing is considered according to nodes degree. In general, degree distribution forms an essential restriction both on the network structure and on assortativity coefficient boundaries. The problem of determining the structure of SF-networks having an extreme assortativity coefficient is considered. The estimates of boundaries for assortativity coefficient have been found. It was found, that these boundaries are as wider as scaling factor of SF-model is far from one of BA-model. In addition, the boundaries are narrowing with increasing the network size.**

Keywords*— *Assortativity, Scale-Free Networks, Boundary Values, Elasticity, Barabasi-Albert Model.

I. Introduction

The most important characteristic of complex networks is the nodes distribution by degrees, i.e. by the number of links. A network is called *Scale-Free* (SF) if this distribution follows to a power law, at least asymptotically. According to the results of many studies [1-4], most of the real-world networks are scale-free. The simplest and most common model of SF-networks is the Barabási-Albert (BA) model.

Despite the importance the degree distribution, it is not comprehensive characteristic of networks. An important one is *assortativity* [4]-[7], that is, the tendency of nodes to connect with similar or dissimilar ones (thus networks called assortative or disassortative respectively). While degree distribution is a first order graph metric, the assortativity is a second order one. In common, assortativity is considered for the degree of the nodes in the network. That consideration is provided in Sec II.

Despite the large number of studies of scaling and assortativity, these characteristics are studied separately that is considered as an essential disadvantage of the current state of the complex networks theory. While the structure of general-type networks with extreme assortativity is widely known [4]-[5], it is unknown what it is for SF-networks. The influence of scaling to assortativity is studied in Sec. III.

Finding the networks structure for extreme cases (extremely assortative and disassortative SF-networks) forms the first problem to deal with, while estimating the bounds of assortativity coefficient for the SF-networks with known power exponent is another one. These problems are studied in Sec. IV.

II. Assortativity Properties of Networks

The term «assortative mixing» originally appeared in epidemiology and sociology, in particular for mixing by race among sexual partners [1].

In network theory assortative mixing characterizes the tendency for vertices in networks to be connected to other vertices that are like (or unlike) them in some way. As a simplest case assortative mixing is considered according to vertex degree. An *assortativity coefficient* is defined as correlation coefficient of the nodes by their degrees [1] and in general case it lies in the range $-1 \le r \le 1$.

It is known [1], [5]-[6], [8]-[9] that social networks have a mainly positive assortativity, while biological and technical ones are disassortative. Artificial network models (BA-model and others) are asymptotically neutral.

More convenient form of this coefficient is given in [5]:

$$r = \frac{S_1 N_3 - S_2^2}{S_1 S_3 - S_2^2}, \qquad (1)$$

where

$$S_k = \sum_{i=1}^{n} d_i^k, \quad N_3 = d^T A d = \sum_{i=1}^{n} \sum_{j=1}^{n} A_{ij} d_i d_j, \qquad (2)$$

A is adjacency matrix of the network, d_i is node degrees.

Without loss of generality, we can assume that the nodes are ordered by a decreasing degree, so index number i of the node d_i is its range.

As it follows from (2), N_3 is the only component of (1) which depends on network structure (i.e. adjacency matrix A). Thus, finding the networks structure (in extreme assortative/disassortative case) is a *binary programming* problem:

$$N_3 = d^T A d = \sum_{i=1}^{n} \sum_{j=1}^{n} A_{ij} d_i d_j \to extr \qquad (3)$$

with restrictions

$$\sum_{j=1}^{n} A_{ij} = d_i, \quad A_{ij} \in \{0,1\}, \quad A_{ij} = A_{ji}, \quad A_{ii} = 0. \qquad (4)$$

In general case (i.e. fixed but arbitrary degree distribution) problem (3)-(4) cannot be solved in analytical form. However, asymptotical solutions of this problem can be found by greedy algorithms of simple structure [4].

978-1-7281-4184-8/19/$31.00 ©2019 IEEE

Links are distributed starting from the richest node. Network became extremely disassortative if on each step current node connects with as poorest one as possible. In the same way, while creating an extremely assortative network current node connects with as richest one as possible. Proof of the asymptotic optimality of given greedy algorithms is also done in [4].

Examples of structure and adjacency matrix of extremely assortative / disassortative networks are shown on Fig.1 and Fig.2.

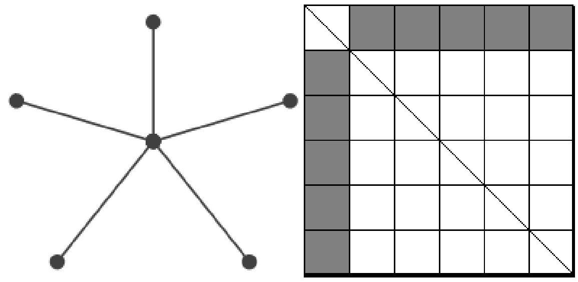

Fig. 1. Structure and adjacency matrix of extremely disassortative network.

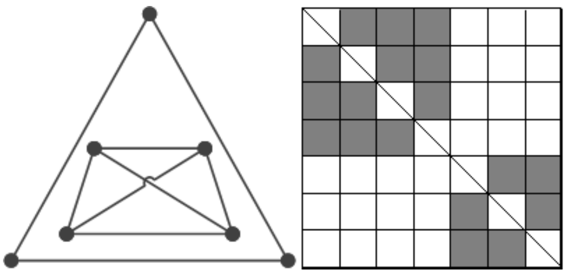

Fig. 2. Structure and adjacency matrix of extremely assortative network.

As one can see, in the boundary case $r=-1$ (an extremely disassortative mixing) network has *biregular* structure [5] and in another boundary case $r=1$ (assortative mixing) network consists of two or more regular connected components.

III. The Effect of Scaling to Assortativity Properties

It follows directly from the definition (1)-(2) that network assortativity coefficient (r) depends on degree distribution d_i, so networks mixing cannot be studied apart from scaling..

One of the key-properties of complex networks is their self-similarity or scale-freeing [2], [10]-[11]. The network is called *scale-free* (SF), if the distribution of vertex degrees follows a power law, at least asymptotically, i.e. the fraction p_k of nodes, having k edges for large k is

$$p_k \sim k^{-\gamma}, \quad 2<\gamma\le 3. \tag{5}$$

In this case the range distribution of nodes asymptotically also follows to a power-family distribution:

$$d_i \approx c\cdot i^{-\beta}, \quad \beta = 1/(\gamma-1), \quad \tfrac{1}{2}\le\beta<1. \tag{6}$$

Actually, factor c in (6), is not a constant, but depends on network size (n). We use c, rather than $c(n)$, for simplicity:

One of the simplest and the most popular particular case of SF-network is Barabási-Albert (BA) model, for which $\gamma=3$ and thus $\beta=1/2$. The structures of extremely assortative / disassortative BA-networks were also studied in [4]. As it was shown, the structure of extremely disassortative BA-network is not strictly biregular, but tends to be *bipartite* (Fig.3). The structure of extremely assortative BA-network is close to be a set of almost isolated clusters all but the largest of which are almost regular (Fig.4).

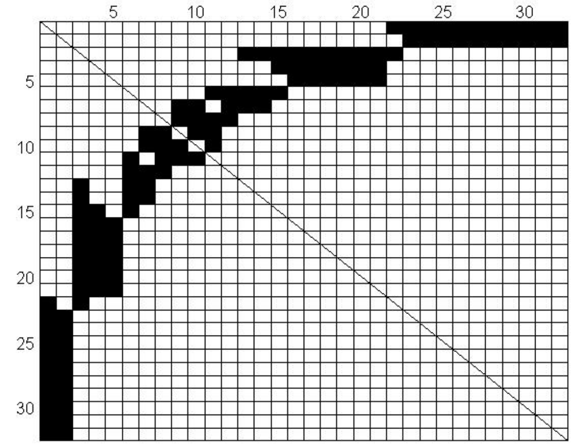

Fig. 3. An adjacency matrix of extremely disassortative BA-network with $n=32$, $\beta=1/2$.

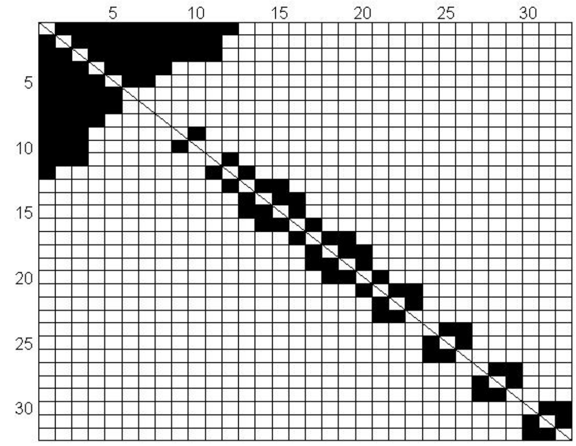

Fig.4. An adjacency matrix of extremely assortative BA-network with $n=32$, $\beta=1/2$.

Asymptotic estimates of the bounds of assortativity coefficient for BA-network were also obtained in [4]. It looks natural, that the achievable bounds of assortativity coefficient are essentially narrower than ± 1:

$$\lim_{n\to\infty} r_{\min} \approx \frac{-\ln^2(n)}{c_3\sqrt{n}}, \tag{7}$$

$$\lim_{n\to\infty} r_{\max} \approx \frac{b}{c_3} n^{-1/8}, \tag{8}$$

with constants $b \approx 10.026$ and $c_3 \approx 7.160$.

Generalization this results to the general case of SF-networks forms a subject of interest of current research.

IV. Estimating the Boundaries of Assortativity Coefficient for SF-Networks

The terms S_1, S_2, S_3 in (1) does not depends on mixing properties (i.e. on adjacency matrix A), but only on scaling properties, i.e. on degree distribution d_i. For SF-model having degree distribution (6) an asymptotic estimates are:

$$S_1 \approx c\left(\zeta(\beta)+\frac{n^{1-\beta}}{1-\beta}\right) \approx \frac{c}{1-\beta} n^{1-\beta}, \qquad (9)$$

$$S_2 \approx c^2\left(\zeta(2\beta)+\frac{n^{1-2\beta}}{1-2\beta}\right) \approx c^2 \cdot \zeta(2\beta), \qquad (10)$$

$$S_3 \approx c^3\left(\zeta(3\beta)+\frac{n^{1-3\beta}}{1-3\beta}\right) \approx c^3 \cdot \zeta(3\beta), \qquad (11)$$

where $\zeta(\,)$ is Riemann-zeta function.

In disassortative case according to generative algorithm each vertex i from "rich" subset is linked with d_i of "poor" vertices which indexes (ranges) sequentially varies from some j_{min} up to $j_{max} = j_{min} + d_i$. Thus, according to mean value theorem, for each i there is some intermediate value $\bar{j} = f(i)$ such that the minimum value of (3) is

$$\begin{aligned} N_3^{\min} &= 2\sum_{i=1}^{k} \sum_{j=j_{\min}}^{j_{\max}} d_i d_j = \\ &= 2\sum_{i=1}^{k} d_i \sum_{j=j_{\min}}^{j_{\max}} d_j = 2\sum_{i=1}^{k} d_i^2 d_{\bar{j}} \end{aligned}. \qquad (12)$$

The value of $\bar{j}$ can be estimated in assumption that vertices having indexes less than i captures all the links with vertices having indexes greater than $\bar{j}$:

$$\sum_{s<i} d_s = \sum_{s>\bar{j}} d_s. \qquad (13)$$

In abstraction from integerness of all indexes and degrees, the solution of (13) for SF-model with degree distribution (6) takes the form

$$\bar{j} = f(i) = \left(n^{1-\beta} - i^{1-\beta}\right)^{1/(1-\beta)}. \qquad (14)$$

From (13)-(14) it also follows that limit of summation in (12) is determined by the condition $i = \bar{j}$, thus,

$$k = 1 + \left\lfloor n \cdot 2^{-1/(1-\beta)} \right\rfloor. \qquad (15)$$

Approximating the sum (12) by the integral, in respect with (6) and (15) we obtain

$$N_3^{\min} \approx \frac{2c^3}{2\beta-1} n^{-\beta}, \qquad (16)$$

By substituting (9)-(11) and (16) into (1), we get an asymptotic estimate of the lower bound of assortativity coefficient:

$$\lim_{n\to\infty} r_{\min} \approx -\chi(\beta) \cdot n^{-(1-\beta)}, \qquad (17)$$

where

$$\chi(\beta) = \frac{(1-\beta) \cdot \zeta^2(2\beta)}{\zeta(3\beta)}. \qquad (18)$$

For the case of maximization the criterion (3) (i.e. an extremely assortative case) the first $d_1 = c$ vertices (i.e. the "richest" ones) linked to each other, forming the largest cluster ("arrow head" on Fig.4). The contribution of this cluster to the objective function (3) is

$$N_{31} = \sum_{i=1}^{d_1+1} \sum_{j=1}^{d_i+1} d_i d_j. \qquad (19)$$

By abstracting from integerness of all parameters, one can approximate this sum. In respect with (6) we get:

$$\begin{aligned} N_{31} &= \sum_{i=1}^{c+1} d_i \sum_{j=1}^{d_i+1} d_j \approx \frac{c}{1-\beta} \sum_{i=1}^{c+1} (d_i)^{2-\beta} \approx \\ &\approx \frac{c^{3-\beta}}{1-\beta} \sum_{i=1}^{c+1} i^{-\beta(2-\beta)} \approx \frac{c^{4-3\beta+\beta^2}}{(1-\beta)^3} \end{aligned}. \qquad (20)$$

The total contribution from all other clusters (from "arrow body" on Fig.4) growth much more slowly than (20). Thus $N_3 \approx N_{31}$ and in accordance with (9)-(11), (20), the upper bound of assortavitity coefficient (1) can be estimated as

$$\lim_{n\to\infty} r_{\max} \approx \eta(\beta) \cdot c^{1-3\beta+\beta^2}, \qquad (21)$$

where

$$\eta(\beta) = \frac{1}{(1-\beta)^3 \cdot \zeta(3\beta)}. \qquad (22)$$

As it was mentioned, factor c in (6) (and therefore in (20)-(21)), *is not a constant*, but depends on network size (n). This dependence can be obtained from (9), since S_1 is a total number of links. Denote as λ $(1 \le \lambda < 2)$ the *elasticity factor* [6] of network, i.e. the scaling factor of links number w.r.t. nodes number:

$$S_1(n) = \frac{2\Gamma(n-1+\lambda)}{\Gamma(n-1)\Gamma(\lambda+1)} \approx const \times n^{\lambda}, \qquad (23)$$

If the SF-network is not strictly elastic (23) holds asymptotically. Comparing (9) and (23), we can express c through n, λ, β:

$$c = \nu(\lambda,\beta)\cdot n^{\lambda-1+\beta}, \quad \nu(\lambda,\beta) = \frac{2(1-\beta)}{\Gamma(\lambda+1)} . \qquad (24)$$

Thereby the asymptotic estimate for upper bound of assortavitity coefficient takes the form

$$\lim_{n\to\infty} r_{\max} \approx \sigma(\lambda,\beta)\cdot n^{(\lambda-1+\beta)(1-3\beta+\beta^2)}, \qquad (25)$$

where

$$\sigma(\lambda,\beta) = \eta(\beta)\cdot(\nu(\lambda,\beta))^{1-3\beta+\beta^2} . \qquad (26)$$

Estimates of the lower (17) and upper (25) boundaries of the assortativity coefficient of general SF-model in comparison with ones of BA-model are shown on Fig.5. Elasticity factor $\lambda = 1$ was used.

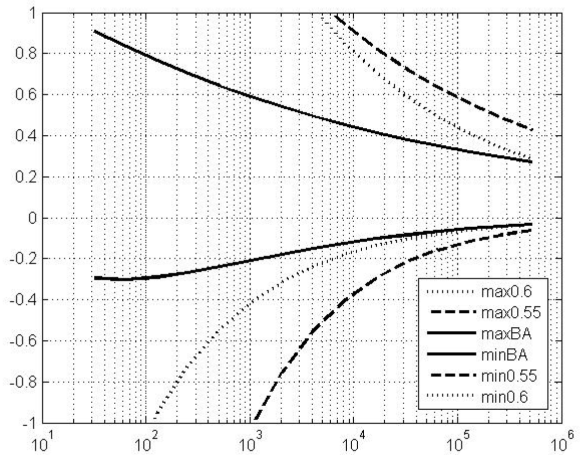

Fig.5. Estimates of the boundaries of the assortativity coefficient for SF-networks.

As one can see, these boundaries are as wider as scaling factor β (6) increase from ½ to 1, i.e. as factor γ (5) decrease from 3 to 2.

One can also conclude that for large networks the boundary from the positive side is much wider than from the negative one. Therefore, the fact that technical and biological networks have assortativity coefficients, which are much smaller in absolute value than for social ones, can be explained by the sign of this coefficient.

V. Conclusions

The most important characteristic of complex networks is the nodes distribution by degrees. Many real-world networks are found to be scale-free, i.e. the nodes degrees follow a power law.

While degree distribution is a first order graph metric, the assortativity is a second order one. The concept of assortativity is extensively using in network analysis.

In general, assortativity coefficient can vary from -1 to +1, but degree distribution forms an essential restriction both on the network structure and on coefficient bounds. The problem of determining the structure and assortativity coefficient boundaries for SF-networks is considered. The mathematical formulation of this problem is done. It was found that the structures of SF-networks with extreme assortativity coincide with ones of BA-networks: an extremely disassortative net is asymptotically bipartite (Fig.3) while an extremely assortative one (Fig.4) is asymptotically a set of complete and disconnected clusters.

The formulation of the problem of finding the boundary values for the assortativity coefficient of SF-networks were made. The estimates of boundaries for assortativity coefficient of SF-networks have been found. These boundaries were compared with ones obtained for BA-model. It was found, that these boundaries are as wider as scaling factor of SF-model is far from one of BA-model. In addition, boundaries for assortativity coefficient are narrowing with increasing the network size.

Obtaining more accurate estimates for the model of elastic SF-networks is treated as a direction for further research.

References

[1] M. Newman. “Mixing patterns in networks”, Phys. Rev. E, vol. 67, 2003

[2] N. Dorogovtsev and J. Mendes. “Evolution of networks”, Adv. in Physics, Vol. 51(4), pp. 1079-1187, 2002.

[3] V. L. Shergin, L. E. Chala and S. G. Udovenko. "Fractal dimension of infinitely growing discrete sets", in 14th Int. Conf. on Adv. Trends in Radioelecrtronics, Telecommunications and Computer Engineering (TCSET), Slavske, 2018, pp. 259-263.

[4] V. L. Shergin, L. E. Chala and S. G. Udovenko. "Assortativity Properties of Barabási-Albert Networks", DCBA-2019, 2019.

[5] R. Noldus and P. Van Mieghem. “Assortativity in complex networks”, J. Complex Networks, vol. 3, pp. 507-542, 2015.

[6] I. Sendiña-Nadal et al., “Assortativity and leadership emerge from anti-preferential attachment in heterogeneous networks”, Scientific reports, vol. 6 21297, 2016, doi:10.1038/srep21297

[7] M. Piraveenan, M. Prokopenko and A. Y. Zomaya. "Local assortativeness in scale-free networks", EPL 84, 28002, 2008.

[8] L. Peel, J. C. Delvenne and R. Lambiotte. “Multiscale mixing patterns in networks”, Proc. of the National Academy of Science, 2017, doi: 115.10.1073/pnas.1713019115.

[9] A. Muscoloni and C. V. Cannistraci. “Rich-clubness test: how to determine whether a complex network has or doesn’t have a rich-club?”, in 6th Int. Conf. on Complex Networks & Their Applications, Lyon (France), 2017, pp.291-293.

[10] V. Shergin and L. Chala. "The concept of elasticity of scale-free networks", in 2017 4th Int. Sci.-Pract. Conf. Problems of Infocommunications, Science and Technology (PIC S&T), Kharkov, 2017, pp. 257-260.

[11] Z. Zhang, L. Fang, S. Zhou and J. Guan. "Effects of accelerating growth on the evolution of weighted complex networks", Physica A: Statistical Mechanics and its Applications, vol. 388, pp. 225-232, 2007, doi: 10.1016/j.physa.2008.10.008.