Assortativity of an Elastic Network with Implicit Use of **Information about Nodes Degree**

Vadim Shergin^a, Larysa Chala^a, Serhii Udovenko^b and Mariia Pohurska^a

^a Kharkov National University of Radio Electronics, Nauky Ave. 14, Kharkov, 61166, Ukraine

^b Simon Kuznets Kharkov National University of Economics, Nauky Ave. 9-a, Kharkov, 61166, Ukraine

Abstract

The problem of modeling dynamics of complex networks is considered. The network model with non-unit elasticity controlled by mediation-driven attachment rule is studied. According to that rule, new nodes do not need a direct access to node statistics of the network: this data affects the links of new nodes, but is used implicitly. The assortative property of the considered model of complex networks is studied. It was shown that the proposed model generates networks with non-zero assortativity coefficient that is an important feature of this model. The dependence of the assortativity coefficient on the network size and copy-factor is obtained by numerical simulation.

Keywords¹

Assortativity, complex network, scale-free network, mediation-driven attachment, elasticity, power law

1. Introduction

The theory of complex networks is a modern science area. As is known, the most of such networks demonstrate the power law distribution of nodes by degree [1-4]. Networks with the above property are called "scale-free". For instance, most of social networks, collaboration networks, communication networks, and citations of scientific papers, neural and protein networks [5] (but not all of them [6]) are scale-free.

At the heart of most of today's scale-free network models is the model proposed by Barabási and Albert in 1999 (BA-model) [7]. It is based on two powerful but simple concepts: concept of growth and concept of preferential attachment (or preferential linking). According to the latter, the probability for a new vertex to attach with some existing one is proportional to its degree. Although at present there exists many specifications and generalizations for the preferential attachment rule [1-4, 8-10], the proposed network models has some common features inherited from the parental BA-model. In particular, all these models inherit such a drawback of the BA-model as the explicit use of network's node statistic. The concept of implicit use of node degree information has been proposed in the mediation-driven anchor (MDA) model by Hassan [11]. Later the elastic variant of this model (EMDA model [12]) has been also developed.

Apart from the nodes degree distribution, the assortativity property of networks is also very important. It is known [2, 13-15] that the real world networks are essentially assortative (social networks), or disassortative (biological and technical ones), while artificial dynamic network models (such as BA) generate networks that are neutral in assortativity of nodes degree. This fact is a significant drawback of existing models of complex networks. Studying the assortativity properties of the EMDA-model forms the main subject of current interest.

II International Scientific Symposium «Intelligent Solutions» IntSol-2021, September 28–30, 2021, Kyiv-Uzhhorod, Ukraine

EMAIL: vadim.shergin@nure.ua (A. 1); larysa.chala@nure.ua (A. 2); udovenkosg@gmail.com (A. 3); mariia.yerovenko@nure.ua (A. 4) ORCID: 0000-0002-4388-8180 (A. 1); 0000-0002-9890-4790 (A. 2); 0000-0001-5945-8647 (A. 3); 0000-0001-7778-231X (A. 4)



© 2020 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

CEUR Workshop Proceedings (CEUR-WS.org)

2. Models of dynamics of complex networks overview and problem statement

All models describing the networks topology can be classified into dynamic and static, depending on whether they are based on the concept of growth, i.e. increasing the number of nodes and links over time, or not. Static are models "inherited" from graph theory (for instance, the Erdős-Rényi (ER) model), as well as a well-known "small world" model by Watts-Strogatz [15].

In dynamic models, new nodes and links are added to the network step by step. Usually, but not always, the addition occurs one node at a each step. In this case, the size of the network (n) is a measure of time, and the number of the node (i) is the time of its birth.

The growth of the network and its properties are determined by the presence or absence of dying off nodes and links, the source of links and the linking rule. The simplest and most popular class of dynamic models of complex networks is formed by models in which only a new node can be a source of a new edge, and there is no dying off of nodes and links. In such models, the linking rule is the dependence of the probability π_i of connecting a new node with an existing node *i* on its individual properties (*Prop*(*i*)), the network size *n*, and on the general (control) parameters of the network (*Net_Prop*):

$$\pi_i = f(n, Prop(i), Net_Prop).$$
(1)

The most famous dynamic network model is the Barabási-Albert (BA) model [7], for which rule (1) has the form

$$\pi_i = \frac{k_i}{\sum k_i}.$$
 (2)

According to (2) the probability of joining a new node (having a number n+1) to an existing one (*i*) is proportional to its degree k_i . Linking rule (2) is called the preferential attachment rule.

The control parameter of the BA-network is the number of links m = const incident to each incoming node.

The BA-model generates a scale-free network (SFN). It means, that the distribution of nodes by degree (i.e. the number of links p(k)) follows the Yule-Simon (YS) distribution. This distribution is a discrete analogue of a power law and asymptotically tends to it degree tends to infinity:

$$p(k) = C \frac{\Gamma(k)}{\Gamma(k+\gamma)} \propto k^{-\gamma}, \quad k \ge k_0.$$
(3)

The parameter $\gamma \ge 2$ is called the scaling factor. For BA-network $\gamma = 3$.

Currently, there are many different extensions and generalizations of the BA model, which nevertheless fit into the general form (1) of the linking rule (fitness model [8], model with an aging factor [9], with a factor of additional attractiveness [1], based on nonlinear preferential attachment [2], link redirection [13], using the properties of nodes of the second level of neighborhood [10], etc.).

One of such modifications of the BA-model is the elastic network model [12, 16, 17]. This model assumes that the relative growth rates of the number of links and nodes are different. Their ratio is called the elasticity index:

$$\lambda = \frac{\delta L(t)}{\delta n(t)} = \frac{\frac{\Delta L(t)}{L(t)}}{\frac{\Delta n(t)}{n(t)}} = n \frac{L(n+1) - L(n)}{L(n)}.$$
(4)

If $\lambda = const$ $(1 \le \lambda < 2)$, then the total number of links in the network L(n) and, accordingly, the average degree of nodes $\overline{k}(n)$ grow asymptotically according to a Yule-Simon law with exponents λ and $\lambda - 1$ respectively. Thus, using the concept of elasticity make us possible to generalize the original BA-model and allows dense networks to be modeled.

For $\lambda > 1$, each new node brings more links than their current average:

$$\Delta L(n) = L(n+1) - L(n) = \lambda \cdot \overline{k}(n).$$
⁽⁵⁾

For example, the more scientific papers have been written, the more references the author of the new paper should revise.

The most important classifying factor of network growth models is the availability of information about the nodes and / or links of the existing network. In BA-model and its numerous generalizations, this information (for example, the values k_i of the degrees of nodes in (2)) is considered as available.

Obviously, for many networks in the real world, the hypothesis of the availability and explicit use of information is unrealistic. Both because of the closed nature of this information, and because of its huge volumes. Thus, a more natural class of complex network models is mediation-driven attachment (MDA) models. In the original model [11], a new node chooses one of the existing ones as an intermediary (equiprobably) and then connects (also equiprobably) with m of the neighbors of this mediator. The probability of joining π_i will depend on the degree of this node k_i , but this degree value is used implicitly. In the elastic MDA-model (EMDA, [12]), not m is fixed, but the probability (q) of forming a link with an intermediary's neighbor. Parameter q is also called the copy-factor. This model is relatively new, its description and analysis is the first task of the current study.

An important property of networks is assortativity, i.e. the tendency of nodes to connect with similar ones by some property, or with opposite ones. Any node parameter can be used as such a property; in the simplest case, the node degree. It is known [13-15] that the real world networks are essentially assortative (social networks), or disassortative (biological and technical ones), while artificial dynamic network models (BA, ER, and others) generate networks that are neutral in assortativity of nodes degree. This property should be considered as a significant drawback of existing models of complex networks. Studying the assortativity properties of the EMDA-model is the second and main problem of current research.

3. Model of an elastic network with mediation-driven attachment (EMDA)

In the EMDA model [12], the control parameter is the copy-factor (q) - the probability with which a new node binds to an mediator's neighbor node. In addition, it is assumed that the intermediary node itself is always linked with a new node (Figure 1).

Thus, the number of links incident with the new node is not a constant (as in the MDA rule, or in the BA-model), but equal to $m = 1 + q \cdot k_{med}$ (where k_{med} is the degree of the intermediary node). Then the equation for the dynamics of the number of links in the network has the form

$$E\{\Delta L(n)\} = 2(1 + q \cdot E\{k_{med}\}) = 2 + 2q \cdot k(n).$$
(6)

Comparing (5) and (6), it is easy to see that the EMDA model is elastic with the elasticity index

$$\lambda = 2q \,. \tag{7}$$

The average degree of nodes $\overline{k}(n) = L(n)/n$ depends on the network size and on the copy-factor:

$$\overline{k}(n) = \frac{2}{2q-1} \left(\frac{\Gamma(n+2q)}{\Gamma(n+1)\Gamma(1+2q)} - 1 \right).$$
(8)



Figure 1: Probabilities of connecting a new node according to the EMDA rule.

According to (8), for 0 < q < 0.5, the average nodes degree increases to a finite limit $\overline{k}_{\text{lim}}$. While 0.5 < q < 1, the average nodes degree grows unboundedly according to an asymptotically power law. And for q = 0.5 the growth rate is logarithmic:

$$\overline{k}(n) \propto \begin{cases} 2/(1-2q), & 0 < q < 0.5\\ 2\ln(n), & q = 0.5\\ n^{2q-1}, & 0.5 < q < 1 \end{cases}$$
(9)

The dependence of the average degree of nodes on the network size and the copy factor (both theoretical (8)-(9) and numerical) is shown in Fig. 2.



Figure 2: The dependence of the average degree of nodes on the network size and the copy factor

It was shown [12] that the rank distribution of the nodes of the EMDA network asymptotically follows to a power law with the scaling factor

$$\beta = \min\{q, 1-q\}.$$
 (10)

The power law of the rank distribution with exponent (10) corresponds to the power law of frequency distribution (3) with the scaling exponent

$$\gamma = 1 + 1/\beta = 1 + 1/\min\{q, 1 - q\}.$$
(11)

Thus, the EMDA model generates networks that are elastic with the elasticity exponent (7) and scale-free with the scaling factor (11).

We can compare this fact with the properties of non-elastic (MDA) model. In such models factor m (instead of q) used as a control parameter. Because of m is not scale-free, thus the generated MDA – networks should be not. Hassan [11] called MDA models "superpreferencial" in the sense of the attachment rule. Rank distributions of nodes degree for cases m=1, m=5 and m=100 (obtained in [12]) are shown on Fig. 2 – Fig.4.

As one can see, this distributions are not power-law, thus MDA-model is not scale-free. Conversely, the analyzed class of models (EMDA) produces networks, which are scale-free.



Figure 3: Rank distribution of nodes degree for the case m = 1.



Figure 4: Rank distribution of nodes degree for the case m = 5.



Figure 5: Age distribution of nodes degree for the case m = 100.

4. Assortativity of complex networks

An *assortativity coefficient* is defined as correlation coefficient of the nodes by their degrees [13, 15], or by the adjacency matrix:

$$r = \frac{S_1 N_3 - S_2^2}{S_1 S_3 - S_2^2},$$
 (12)

where

$$S_{k} = \sum_{i=1}^{n} d_{i}^{k} , \qquad N_{3} = d^{T} A d = \sum_{i=1}^{n} \sum_{j=1}^{n} A_{ij} d_{i} d_{j} , \qquad (13)$$

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

In general case assortativity coefficient (12) lies in the range $-1 \le r \le 1$; r=0 for random networks, such as BA-model or ER-graph. In boundary cases $r=\pm 1$, the network has [14] a special structure (shown in Fig. 6), which are far from the structure of real networks.



Figure 6: Structure of extremely disassortative (a) and extremely assortative (b) networks

The power law of the nodes degree distribution (3) significantly limits the boundary values of the assortativity index. In [14] the lower and upper bounds of the assortativity coefficient for BA-networks were estimated:

$$\lim_{n \to \infty} r_{\min} \approx \frac{-\ln^2(n)}{7.16\sqrt{n}}, \qquad \lim_{n \to \infty} r_{\max} \approx 1.4n^{-1/8}.$$
 (14)

It is known [15] that the real-world networks are not neutral on its assortativity: social networks are positively assortative while technical and biological ones are disassortative. Based on the correlation nature of the assortativity coefficient, the numerical values given in Table 1 (less than 0.3 in absolute value) do not seem to be very significant. But supposing these networks as follows to BA-model, we can conclude that the assortative property should be considered as strongly significant (Fig. 7). In fact, the networks shown in this example are SFN (3), but do not follow exactly to the BA model (i.e., scaling factor γ is not obviously equal to three and the elasticity index λ also is not necessarily to be unity).

Table 1

Kind of networks	Network	Size	Assortativity
social	physics coauthorship	52 909	0.363
	biology coauthorship	1 520 251	0.127
	mathematics coauthorship	253 339	0.120
	film actor collaborations	449 913	0.208
	company directors	7 673	0.276
	email address books	16 881	0.092
technological	power grid	4 941	-0.003
	Internet	10 697	-0.189
	World-Wide Web	269 504	-0.067
	software dependencies	3 162	-0.016
biological	protein interactions	2 115	-0.156
	metabolic network	765	-0.240
	neural network	307	-0.226
	marine food web	134	-0.263

Assortativity of some real-world networks (from [15])



Figure 7: Estimates of the boundaries of the assortativity coefficient for BA-network

Asymptotic bounds of the assortativity index for the general case of SFN ($^{2 < \gamma \leq 3}$ and $^{1 \leq \lambda < 2}$) are also known [18]:

$$r_{\min} \propto -n^{-(1-\beta)}, \qquad r_{\max} \propto n^{(\lambda-1+\beta)(1-3\beta+\beta^2)},$$
 (15)

where $0.5 < \beta < 1$ is the nodes degree rank distribution scaling factor (10).

In this, general case, the boundaries (15) are wider than in the case of the BA-model (14), but the general properties of the asymptotic boundaries are preserved: the boundaries narrow to zero with increasing the network size, and the lower boundary tends to zero much faster than the upper one.

5. Numerical estimation of the assortativity of the EMDA-network

The seed of the EMDA network model was two nodes connected by an edge. At each step, one node was added and linked to the existing ones according to the EMDA algorithm. Thus, network models having size n = (128,256,512,1024,2048,4096,8192) were built. For each of these networks, the assortativity index was calculated by (12)-(13). The results obtained were averaged over M = 100 ensembles.

The dependence of the assortativity coefficient of the network on its size and control parameter (copy-factor q) is shown in Fig. 8.



Figure 8: The dependence of the assortativity coefficient of the network on its size and copy-factor

In Fig.9 is shown the dependence of the assortativity coefficient of the network with fixed size n = 8192 on the control parameter.

Thus, numerical simulations show that the EMDA algorithm generates networks with significantly positive assortativity. Assortativity coefficient decreases with increasing copying factor. This is consistent with the fact that by increasing the copy factor, the scaling factor of the nodes degree distribution is also increase.



Figure 9: The dependence of the assortativity coefficient of the network on the copy-factor

Last years some new approaches to networks assortativity problem were presented [19-26]. Thus, applying it to the theoretical study of the assortativity of the EMDA model is considered as a direction for further research.

6. Conclusions

The analysis of dynamic models of complex networks and the corresponding attachment rules is provided. It was obtained that significant disadvantages of traditional models are the explicit use of information about the properties of existing nodes and the linear dependence of the number of links in the network on the number of nodes. It is shown that the EMDA model is free of these lacks: it is elastic and does not use node statistics. A brief description of this model is provided.

The analysis of the assortative property of real world networks and dynamic models of such networks is carried out. It was revealed that real networks have an explicit property of assortativity (positive or negative), while existing artificial models generate networks with zero assortativity, i.e. neutral. In this regard, the task of studying the assortativity of networks generated by the EMDA-model was formulated and solved.

As a result of the provided numerical simulation, it was found that the EMDA algorithm generates networks with significantly positive assortativity. Assortativity coefficient decreases with increasing copying factor. Thus, it can be argued that the EMDA model reflects the most important properties of social networks: asymptotic scale invariance, implicit use of information about nodes, high density of links, and significantly expressed positive assortativity.

The theoretical study of the assortativity of the EMDA model is considered as a direction for further research.

7. References

- [1] S. N. Dorogovtsev, J. F. F. Mendes, "Evolution of Networks: From Biological Networks to the Internet and WWW", Oxford, USA: Oxford University Press, 2003. 280 p.
- [2] P. L. Krapivsky, S. Redner; F. Leyvraz, "Connectivity of Growing Random Networks", *Phys. Rev. Lett.*, 2000, 85: 4629–4632.
- [3] K. Choromański, M. Matuszak, J. MięKisz, "Scale-Free Graph with Preferential Attachment and Evolving Internal Vertex Structure." *Journal of Statistical Physics*. (2013), 151 (6): 1175–1183.

- [4] J. Kunegis, M. Blattner, C. Moser, "Preferential Attachment in Online Networks: Measurement and Explanations.", WebSci13 ACM Web Science Conference 2013. 13. doi: 10.1145/2464464.2464514.
- [5] M. E. J. Newman, "Power laws, Pareto distributions and Zipf's law", *Contemporary Physics*, 2005, 46(5). p.323-351.
- [6] A. D. Broido, A. Clauset, "Scale-free networks are rare". *Nat Commun* 10, 1017 (2019). https://doi.org/10.1038/s41467-019-08746-5
- [7] R. Albert, A.-L. Barabási, "Statistical mechanics of complex networks", *Rev. Mod. Phys.*, 2002, V. 74. p. 42-97.
- [8] G. Bianconi, A.-L. Barabási, "Competition and multiscaling in evolving networks", *Europhysics Letters*, 2001, 54 (4):436–442.
- [9] G. Caldarelli, Guido; Catanzaro, Michele (2012). Networks: A Very Short Introduction. Oxford University Press. p. 78
- [10] Ch. Dangalchev, "Generation models for scale-free networks", Physica A, 2004, 338, 659
- [11] M. K. Hassan, L. Islam, S. A. Haque, "Degree distribution, rank-size distribution, and leadership persistence in mediation-driven attachment networks". *Physica A: Stat. Mech. and its Appl.* 2017. 469 (2017): 23–30
- [12] V. Shergin, L. Chala, S. Udovenko, M. Pogurskaya, "Elastic Scale-Free Networks Model Based on the Mediaton-Driven Attachment Rule", 2020 IEEE Third International Conference on Data Stream Mining & Processing (DSMP), Lviv, 2020, in press.
- [13] R. Noldus, P. Van Mieghem, "Assortativity in complex networks", J. Complex Networks, 2015, vol. 3, pp. 507-542.
- [14] V. Shergin, S. Udovenko, L. Chala. "Assortativity Properties of Barabási-Albert Networks," In *Data-Centric Business and Application*. Springer, 2021.
- [15] M. E. J. Newman. "Mixing patterns in networks". *Physical Review E.* 67 (2), 2003, 026126. arXiv:cond-mat/0209450. Bibcode:2003PhRvE..67b6126N. DOI:10.1103/physreve.67.026126. ISSN 1063-651X. PMID 12636767
- [16] V. Shergin, L. Chala, "The concept of elasticity of scale-free networks," 2017 4th International Scientific-Practical Conference Problems of Infocommunications. Science and Technology (PIC S&T), Kharkov, 2017, pp. 257-260.
- [17] V. L. Shergin, L. E. Chala and S. G. Udovenko, "Fractal dimension of infinitely growing discrete sets," 2018 14th International Conference on Advanced Trends in Radioelecrtronics, Telecommunications and Computer Engineering (TCSET), Slavske, 2018, pp. 259-263.
- [18] V. Shergin, L. Chala, S. Udovenko. Assortativity Properties of Scale-Free Networks. In 2019 International Scientific-Practical Conference Problems of Infocommunications. Science and Technology (PIC S&T). IEEE, Kharkiv, 2019, pp. 92-96.
- [19] B. K. Fosdick, D.B. Larremore, J. Nishimura, J. Ugander, "Configuring random graph models with fixed degree sequences.", 2016, arXiv:1608.00607.
- [20] L. Peela, J-C. Delvennea, R. Lambiotted, "Multiscale mixing patterns in networks", https://doi.org/10.1073/pnas.1713019115
- [21] G. Q. Zhang, S. Q. Cheng, (2012) "A universal assortativity measure for network analysis", 2012, arXiv preprint arXiv:1212.6456 1: 1–8.
- [22] N. Meghanathan, "Assortativity Analysis of Real-World Network Graphs based on Centrality Metrics", Computer and Information Science Archives, Vol. 9, No. 3 (2016)
- [23] D. N. Fisher, M. J. Silk, D. W. Franks "The Perceived Assortativity of Social Networks: Methodological Problems and Solutions." In: Trends in Social Network Analysis. Lecture Notes in Social Networks. Springer, Cham. 2017 https://doi.org/10.1007/978-3-319-53420-6_1
- [24] Y. Yuan, J. Yan, P. Zhang, "Assortativity measures for weighted and directed networks", arXiv:2101.05389v1 [stat.AP] 13 Jan 2021
- [25] G. Thedchanamoorthy, M. Piraveenan, D. Kasthuriratna, U. Senanayake, "Node Assortativity in Complex Networks: An Alternative Approach", Procedia Computer Science, Volume 29, 2014, Pages 2449-2461, ISSN 1877-0509, https://doi.org/10.1016/j.procs.2014.05.229.
- [26] M. Murakami, S. Ishikura, D. Kominami, Tetsuya Shimokawa, M. Murata, "Robustness and efficiency in interconnected networks with changes in network assortativity", Applied Network Science, vol. 2, No.6 (2017), DOI 10.1007/s41109-017-0025-4